Psychiatry Research 286 (2020) 112891

Contents lists available at ScienceDirect

Psychiatry Research

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

Psychometric properties of the Observable Social Cognition Rating Scale (OSCARS): Self-report and informant-rated social cognitive abilities in schizophrenia

Tate F. Halverson[a,*], Michal Hajdúk[b,c,d], Amy E. Pinkham[e,f], Philip D. Harvey[g,h], L. Fredrik Jarskog[i], Lana Nye[j], David L. Penn[a,k]

[a] University of North Carolina at Chapel Hill, Department of Psychology and Neuroscience, Chapel Hill, NC, United States
[b] Department of Psychology, Faculty of Arts, Comenius University, Bratislava, Slovakia
[c] Clinic of Psychiatry, Faculty of Medicine, Comenius University, Bratislava, Slovakia
[d] Center for Psychiatric Disorders Research – UK, Science Park, Comenius University in Bratislava, Slovakia
[e] School of Behavioral and Brain Sciences, The University of Texas at Dallas, Richardson, TX
[f] Department of Psychiatry, University of Texas Southwestern Medical School, Dallas, TX
[g] Department of Psychiatry and Behavioral Sciences, University of Miami Miller School of Medicine, Miami, FL
[h] Research Service, Miami VA Healthcare System, Miami, FL
[i] University of North Carolina at Chapel Hill, North Carolina Psychiatric Research Center, Raleigh, NC, United States
[j] College of Social Work, The University of Utah, Salt Lake City, UT
[k] Australian Catholic University, School of Behavioural and Health Sciences, Melbourne, VIC, Australia

ABSTRACT

Individuals with schizophrenia spectrum disorders (SSD) consistently show deficits in social cognition (SC) which is associated with real world outcomes. Psychosocial treatments have demonstrated reliable improvements in SC abilities, highlighting the need for accurate identification of SC deficits for efficient and individualized treatment planning. To this end, the Observable Social Cognition Rating Scale (OSCARS) is an 8-item scale with both self and informant versions. This study investigated psychometric properties of the OSCARS as both a self and informant-reported scale in a large sample of SSD (n = 382) and individuals without a psychiatric diagnosis (n = 289). A two-factor structure (Social Cognitive Bias and Social Cognitive Ability) of the OSCARS demonstrated acceptable model fit with good internal consistency for both self- and informant-report. The OSCARS had adequate convergent, external, and predictive validity. Area Under the Curve (AUC) values suggest the OSCARS has some value in identifying individuals with impaired SC and social competence, although stronger AUC values were demonstrated when identifying individuals with impaired real-world functioning. Overall, psychometric properties indicate the OSCARS may be a useful first-step tool for clinicians to detect functioning deficits in SSD and efficiently identify individuals in need of additional assessment or psychosocial interventions.

1. Introduction

Social cognition (SC) encompasses cognitive domains related to how individuals think about themselves, others, social situations, and social interactions (Penn et al., 1997, 2008). Individuals with schizophrenia spectrum disorders (SSD) consistently show deficits in National Institute of Mental Health - defined domains of SC: theory of mind, social perception, social knowledge, attributional bias, and emotional processing (Green et al., 2008; Savla et al., 2013). SC is strongly associated with functional outcomes, with evidence that SC explains more variance in social outcomes than neurocognition (NC; Fett et al., 2011; Halverson et al., 2019). To date, psychotropic medications do not produce robust effects on SC but psychosocial treatments demonstrate reliable improvements in SC abilities (Grant et al., 2017; Kurtz et al., 2016; Sergi et al., 2007). While SC deficits are consistently reported in SSD, findings regarding the prevalence of intact SC in SSD are scant and less consistent, with estimates that 25% - 42% of individuals with SSD do not show impairment (Hajdúk et al., 2018; Rocca et al., 2016). Thus, there does appear to be a subset of individuals without impairment, and therefore, efficient and accurate identification of SC deficits in SSD is an important first step for effective and individualized treatment planning.

To this end, the Social Cognition and Psychometric Evaluation (SCOPE) Study investigated SC measures in SSD and recommended tasks with good psychometric properties (Pinkham et al., 2014). The SCOPE battery is a comprehensive assessment with rigorously tested psychometric properties (see Pinkham et al., 2018, 2016) for SC assessment, however, an administration length of an hour and the need for trained raters may not be feasible for all referrals or treatment

* Corresponding author.
E-mail address: tate.halverson@unc.edu (T.F. Halverson).

https://doi.org/10.1016/j.psychres.2020.112891
Received 5 November 2019; Received in revised form 19 February 2020; Accepted 20 February 2020
Available online 26 February 2020

settings. An alternative approach to detect SC deficits is a brief rating scale. Observer and interview-based rating scales exist for NC deficits with demonstrated reliability and convergence with performance-based tasks (e.g., Schizophrenia Cognition Rating Scale; Keefe et al., 2006, Clinical Global Impression of Cognition in Schizophrenia; Ventura et al., 2008). The development and validation of SC rating scales may similarly prove to be useful and efficient tools for clinicians and clinical research.

Healey et al. (2015) developed a rating scale of SC, the Observable Social Cognition Rating Scale (OSCARS). The OSCARS is an eight-item assessment available as a self- or informant-report measure with a total administration time of five minutes. OSCARS items were selected and reviewed by internationally recognized experts to probe areas of SC with demonstrated deficits in SSD while items considered to be indicators of social skill, self-awareness, and insight were removed. Items are rated on a Likert scale with higher scores reflecting greater impairment. The OSCARS includes a total summary score of all items and a separate single item "global score." Whereas the OSCARS total score is the sum of individually rated areas of SC difficulties, the global score is a single-item overall assessment of social cognition ("overall impression of difficulty in these areas") completed at the end of the OSCARS (analogous to the single-item General Assessment of Functioning (GAF) completed by clinicians after a clinical interview). In an initial validation study, the OSCARS was administered to informants of 62 individuals with SSD and 50 healthy controls (HC). The OSCARS demonstrated sufficient psychometric properties and significant relationships with functional outcomes (Healey et al., 2015). Differences in the factor structure of the OSCARS emerged based on diagnostic group. A two-factor (i.e., Social Cognitive Bias and Social Cognitive Ability) structure provided the best fit for the SSD group while a three-factor structure (i.e., Social Cognitive Bias, Social Cognitive Ability, and Social Cognitive Flexibility) emerged in the HC group. The Social Cognitive Bias factor comprised items indicating impulsivity, hostility, and rigidity; the Social Cognitive Ability factor comprised items involving reasoning and perceptual abilities; and the Social Cognitive Flexibility factor (found only in HC) comprised items involving subtle theory of mind ability and flexibility in social situations (Healey et al., 2015).

Increasingly, studies are utilizing the OSCARS single-item global score. Silberstein et al. (2018) examined the OSCARS global score generated by individuals with SSD and high-contact informants. Informant reports were significantly correlated with both performance-based assessments of SC abilities and with real-world functional outcomes, while self-reports were not. Additionally, studies have used the difference between self- and informant-reported OSCARS global scores as an indicator of introspective accuracy, a component of metacognition defined as how well individuals can evaluate their own abilities and performance on tasks (Harvey et al., 2019; Jones et al., 2019; Silberstein and Harvey, 2019). Given the recent interest and dissemination of the OSCARS and its relationship with SC and functioning, a thorough psychometric investigation to complement the initial validation study is warranted. A more comprehensive understanding of factors that impact the OSCARS global score is also needed as this score has received recent attention as a potential measure of functioning. For example, informants may weigh certain OSCARS factors (e.g., Cognitive Bias versus Cognitive Ability) differently when making this summary rating compared with self-report.

The present study seeks to replicate findings from the OSCARS initial validation study and investigate the factor structure and psychometric properties in a larger sample of SSD and HC individuals. For the first time, psychometric properties of the OSCARS self-report will also be examined and compared with informant-report. The aims of the present study are 1) evaluate the structure and psychometric properties of the OSCARS self-report and informant-report in HC and SSD, 2) examine relationships between OSCARS factor scores and global score, and 3) examine relationships among the OSCARS, SC tasks, and functional outcomes. We hypothesize that 1) OSCARS informant-report will exhibit stronger relationships with functional outcomes compared with self-report and 2) informant-report will significantly predict impaired SC task performance, validating the OSCARS as a reliable proxy to assess SC compared with more resource-intensive performance tasks.

2. Methods

2.1. Participants

The OSCARS was administered as part of the SCOPE study; a multi-site, multi-phase, National Institute of Mental Health project to identify the best measures for assessing SC in schizophrenia (Pinkham et al., 2014). Participants (N = 671) included individuals with at least one source of OSCARS ratings (i.e., self or informant). SSD (n = 382) and HC (n = 289) participants were recruited from four different academic sites (i.e., Southern Methodist University – initial evaluation phase, University of Miami – all phases, University of North Carolina – second and final validation phases, and The University of Texas at Dallas – second and final validation phases) during either the initial (n = 213) second (n = 88), or final validation phase (n = 370) of the SCOPE Study (Pinkham et al., 2016, 2018).

SSD participants were stable outpatients (i.e., no hospitalizations within two months and a stable medication regimen for at least six weeks, no change in medication dosage for at least two weeks) recruited from community clinics, academic medical centers, and Veterans Affairs medical centers. Schizophrenia diagnoses were confirmed through a structured diagnostic interview, the Mini International Neuropsychiatric Interview (MINI; Sheehan et al., 1998), and the psychosis module of the Structured Clinical Interview for DSM Disorders (SCID; First et al., 2002). HC participants were recruited through community advertisements and were administered the MINI to confirm no Axis I or Axis II diagnoses were present. Exclusion criteria for the entire sample included: IQ <70, presence of a medical condition with the potential to affect neurological functioning, vision or sensory limitations interfering with assessment, lack of proficiency in English, or presence of substance abuse (past month) or dependence (past six months). All study measures included in the present analysis were collected in one study visit while informant ratings were completed within two weeks of this visit (see Table 1 for demographics and clinical characteristics).

2.2. OSCARS administration

HC and SSD individuals completed the OSCARS as a self-report measure. SSD participants were asked to identify an informant with whom they had the most regular contact. Only high-quality informants (i.e., professional role or more than two hours of contact per week; n = 174) were included in the present analyses (see Table 3). Informant report was not available for all SSD individuals. The eight OSCARS items include anchor points of severity and frequency for ratings of 1, 3, 5, and 7 to guide raters. The OSCARS global score is on a 10-point scale. See Appendices A and B for full versions of the OSCARS self-report and informant-report.

2.3. Social cognition tasks

Five task-based SC measures collected during all phases of the SCOPE Study with good psychometric properties were examined in relation to the OSCARS. The Bell Lysaker Emotion Recognition Task (BLERT; Bell et al., 1997; range 0–21) and the Penn Emotion Recognition Task (ER-40; Kohler et al., 2003; range 0–40), measured emotion perception. The Reading the Mind in the Eyes task (Eyes; Baron-Cohen et al., 2001; range 0–36), the Hinting Task (Hinting; Corcoran et al., 1995; range 0–20), and The Awareness of Social Inferences Test (TASIT; McDonald et al., 2003; range 0–64) measured

Table 1
Sample characteristics.

	SSD (*n* = 382)	HC (*n* = 289)	*p*-value[a]
Age, years	41.7 ± 12.0	41.1 ± 12.6	.51
Male,% (*n*)	66.0 (252)	57.4 (166)	.02
Education, years	13.0 ± 2.4	13.9 ± 1.9	<.01
WRAT-3 Standard Score	94.6 ± 14.9	98.8 ± 12.3	<.01
Race/Ethnicity[b],% (*n*)			
White	49.0 (187)	47.1 (136)	.63
Black	44.5 (170)	46.4 (134)	.63
Other	6.5 (25)	6.6 (19)	.99
Hispanic/Latinx	18.1 (69)	19.4 (56)	.67
Diagnosis,% (*n*)			
Schizophrenia	49.2 (188)		
Schizoaffective	49.7 (190)		
Psychosis NOS	1.0 (4)		
Medication Type[c],% (*n*)			
Typical	12.6 (48)		
Atypical	73.0 (279)		
Combination	4.7 (18)		
No Antipsychotic	7.6 (29)		
PANSS			
Positive	16.0 ± 5.3		
Negative	13.8 ± 5.2		
General	31.5 ± 8.0		
Total	61.3 ± 14.8		

[a] Chi-squared for categorical variables (sex, race/ethnicity), *t*-test for continuous variables (age, education, WRAT standard score);
[b] individuals able to identify more than one race/ethnicity.
[c] Medication information unavailable for 7 participants. WRAT - Wide Range Achievement Test – 3rd Edition, PANSS = Positive and Negative Syndrome Scale, SSD = schizophrenia spectrum disorder, HC = healthy control; all values presented are *M* ± SD unless otherwise noted.

theory of mind and mental state attribution. Performance on all measures was indexed as the total number of correct responses.

Different measures of attributional bias were administered. The Ambiguous Intentions Hostility Questionnaire (AIHQ; Combs et al., 2007) was administered during the initial and second phases. The AIHQ generates three indices: an Aggression Index and Hostility Index based on rater-scored open-ended responses to vignettes and a Blame Index which is an average composite of Likert items completed by participants. The AIHQ was classified as "not recommended" in the SCOPE initial validation study due to limited relation with functional outcomes (Pinkham et al., 2016). However, recent research suggests the Blame Index exhibits improved psychometric properties and significant relationships with functional capacity and was therefore included in the present analyses (Buck et al., 2017). The Intentionality Bias Task (IBT; Rosset, 2008) was administered during the final validation phase to replace the AIHQ. The IBT generates an Intentionality Bias Index which is the percentage of actions from vignettes that participants endorse as intentional.

Detailed descriptions of all SC tasks are described in previous publications (Pinkham et al., 2016, 2018).

2.4. Neurocognition measures

A subset of the MATRICS Consensus Cognitive Battery (Nuechterlein et al., 2008) was administered to assess NC. Category Fluency - Animal Naming, Symbol Coding, and Trail Making Test - Part A (Trails A) assessed processing speed. Letter-Number Span assessed working memory, and verbal learning was assessed with the Hopkins Verbal Learning Test – Revised (HVLT – R). Due to high correlations between tasks, a composite score was calculated and used for analyses. The Wide Range Achievement Test – 3 Reading subscale (WRAT-3) estimated premorbid IQ (Weickert et al., 2000).

2.5. Functional measures

Three different measures assessed functioning in competence or performance contexts. The Social Skills Performance Assessment (SSPA; Patterson et al., 2001) is a role play task that assesses social competence and consists of two role plays yielding a 1–5 score for each role play and an average total score with higher scores reflecting better functioning. The SSPA was only collected during the initial and final validation phases. Another performance-based measure, the UCSD Performance-Based Skills Assessment, Brief (UPSA-B; Mausbach et al., 2007) was administered to assess functional capacity. The Specific Levels of Functioning Scale (SLOF; Schneider and Struening, 1983) is a 31-item measure of domains of real-world functioning (i.e., interpersonal relationships, social acceptability, participation in activities, and work skills). The SLOF was administered as both a self-report and informant-report measure in the current study. The SLOF yields an average score across all domains (range 1–5) as well four specific domain scores (range 1–5) with higher scores reflecting better functioning.

2.6. Symptoms

The Positive and Negative Syndrome Scale (PANSS; Kay et al., 1987) assessed current symptomatology in the SSD group with higher scores reflecting more severe symptoms. PANSS Positive, Negative, and General scores as well as total score were calculated.

2.7. Statistical approach

The structure of the OSCARS was assessed by fitting confirmatory factor analysis (CFA) models from the two and three-factor models presented by Healey et al. (2015) and comparing model fit statistics across nested models. First, unidimensional models were fit separately to each informant source (i.e., self and informant) within each diagnostic group. Next, one-factor then two-factor models were first fit, followed by three-factor models. Model indices were inspected and chi-square likelihood difference tests were applied to assess optimal model fit. Model fit was evaluated based on the following guidelines: non-significant chi-square (X^2) statistic (Barrett, 2007), comparative fit index (CFI) and Tucker Lewis Index (TLI) values close to or greater than .95 (Hu and Bentler, 1999), standardized root square mean residual (SRMR) value less than .08 (Hu and Bentler, 1999), and a RMSEA value with a 90% confidence interval that does not include values greater than .10 (MacCallum et al., 1996). Judgment of optimal model fit was informed by multiple model fit statistics and theoretical considerations rather than a single criterion (Bentler, 2007). Multi-group CFAs were fit with optimal models to assess measurement invariance (e.g., potential differences in OSCARS factor structure and loadings) across sites. Once an appropriate model fit was established for each informant source within each diagnostic group, internal consistency statistics were calculated.

Correlations between the OSCARS and task-based measures of SC, NC, and functional capacity and real-world outcomes were calculated to assess convergent, divergent, and external validity. Regressions predicting the OSCARS general score from factor scores were examined to assess relative influence of specific domains of observerable SC. Hierarchical regressions predicting functional outcomes and SC task performance examined predictive validity of the OSCARS. Finally, SC composite z-scores normed to average HC performance were calculated and Receiver Operating Characteristic (ROC) analyses assessed utility of the OSCARS to predict impaired SC (i.e., at least 1.5 standard deviations below HC performance). All analyses were done in *R* using the *lavaan* package (Rosseel, 2012) for factor analyses using maximum likelihood estimation.

Table 2
Social cognition tasks and functional outcome measures by diagnostic group.

	SSD	HC	*p*
Social Cognition Tasks			
AIHQ – Blame[a]	8.8 ± 2.9	7.0 ± 2.4	<.01
BLERT	13.6 ± 4.1	15.9 ± 2.8	<.01
ER40	30.7 ± 4.9	33.0 ± 3.3	<.01
Eyes	21.0 ± 5.6	24.3 ± 4.4	<.01
Hinting	13.4 ± 3.7	16.1 ± 2.4	<.01
IBT[b]	44.0 ± 17.7	40.2 ± 14.7	.03
TASIT	44.8 ± 7.6	51.6 ± 6.2	<.01
Neurocognition Tasks[c]			
Animal Naming	19.3 ± 5.7	22.7 ± 6.0	<.01
HVLT	20.9 ± 5.6	25.2 ± 4.6	<.01
Letter Number Sequence	11.8 ± 4.2	14.6 ± 3.8	<.01
Symbol Coding	42.8 ± 11.4	53.3 ± 12.4	<.01
Trails A	39.6 ± 17.4	31.3 ± 10.7	<.01
Functional Outcome Measures			
SSPA Average	4.1 ± 0.5	4.6 ± 0.4	<.01
SSPA1	4.2 ± 0.6	4.7 ± 0.4	<.01
SSPA2	4.0 ± 0.6	4.5 ± 0.5	<.01
SLOF Self-Report Average	4.1 ± 0.6	4.6 ± 0.4	<.01
Interpersonal Relationships	3.6 ± 0.9	4.1 ± 0.7	<.01
Social Acceptability	4.4 ± 0.6	4.7 ± 0.4	<.01
Activities of Community Living	4.4 ± 0.9	4.8 ± 0.5	<.01
Work Skills	4.0 ± 0.8	4.7 ± 0.5	<.01
SLOF Informant Average	4.0 ± 0.6		
Interpersonal Relationships	3.4 ± 0.9		
Social Acceptability	4.4 ± 0.6		
Activities of Community Living	4.5 ± 0.9		
Work Skills	3.7 ± 0.9		
UPSA Total	70.3 ± 14.4		

Note:.
[a] AIHQ only collected during initial and second phases.
[b] IBT only collected during final phase.
[c] Neurocognition composite score used for analyses.
AIHQ – Blame = Ambiguous Intentions and Hostility Questionnaire Blame Index, BLERT = Bell Lysaker Emotion Recognition Task, ER40 = Penn Emotion Recognition Task, Eyes = Reading the Mind in the Eyes Task, Hinting = Hinting Task, IBT = Intentionality Bias Task, TASIT = The Awareness of Social Inference Test, HVLT = Hopkins Verbal Learning Test, SSPA = Social Skills Performance Assessment, SLOF = Specific Levels of Functioning, UPSA = UCSD Performance-Based Skills Assessment, HC = healthy control, SSD = schizophrenia spectrum disorder; all values presented are *M* ± SD.

Table 3
OSCARS descriptive statistics by diagnostic group.

	SSD		HC	
OSCARS	*N*	*M ± SD*	*n*	*M ± SD*
Self-Report Total Score	*289*	19.7 ± 9.1	*289*	12.1 ± 4.9**
Social Cognitive Ability		9.4 ± 4.6		5.8 ± 2.5**
Social Cognitive Bias		10.2 ± 5.4		6.3 ± 2.9**
Informant-Report Total Score	*174*	23.6 ± 10.4		
Social Cognitive Ability		11.5 ± 5.3		
Social Cognitive Bias		12.0 ± 5.9		
Informant Role				
Psychologist/Psychiatrist	*62*			
Parent	*28*			
Significant other	*22*			
Social worker	*18*			
Sibling	*14*			
Friend	*13*			
Child	*9*			
Group home manager	*4*			
Other	*3*			
Aunt	*1*			

** Significant betwen-group differences, $p < .01$; SSD = schizophrenia spectrum disorder, HC = healthy control.

3. Results

3.1. Sample characteristics and descriptives

HC and SSD groups did not significantly differ in age, race, or ethnicity, but the HC group completed more years of education, had a higher IQ, and included more females (see Table 1) which were included as covariates in subsequent analyses. Table 2 presents descriptives by diagnostic group for SC, NC, and functional outcomes and Table 3 presents OSCARS scores. As expected, the HC group demonstrated significantly better performance on all measures of SC, NC, and functional outcomes compared with the SSD group ($p < .03$ for all comparisons). Across sites, SSD groups showed differences on some demographic and clinical variables (see Supplementary Tables 1 and 2).

3.2. Construct validity

Chi-square likelihood difference tests comparing one-, two-, and three-factor CFA models and model fit indices suggested the two-factor model proposed by Healey et al. (2015) was the optimal fit, yielding a Social Cognitive Bias Factor and a Social Cognitive Ability Factor across diagnostic group and informant source (see Table 4).

Multi-group two-factor CFAs were fit to test for configural invariance (i.e., same overall factor structure) and factor loadings or metric invariance (i.e., similar associations between individual items and latent factors) across sites. The OSCARS self-report configural invariance model demonstrated acceptable fit (RMSEA = .09, CFI = .90) and results of a X^2 difference test demonstrated no difference with the metric invariance model ($X^2(12) = 20.2$, $p = .06$) supporting full metric invariance. Likewise, the OSCARS informant-report configural model demonstrated acceptable fit (RMSEA = .11, CFI = .95) with no difference observed with the metric invariance model ($X^2(12) = 20.5$, $p = .06$). Overall, results demonstrate relative stability of the OSCARS factor structure and loadings across sites.

Factor loadings were acceptable to high (.48–.88). Correlations between factors within self-report ($r_{SSD} = .64$ and $r_{HC} = .60$, $ps < 0.01$) and informant-report ($r_{SSD} = .74$, $p < .01$) were also high (see Supplementary Table 3). In the SSD group, medium correlations (*rs* .29–.42, $ps < .01$) were observed between OSCARS self-report and informant-report (i.e., total, global, and factor scores), consistent with previous work (Silberstein et al., 2018). The SSD group, as expected, had significantly higher self-reported OSCARS scores (i.e., more impaired SC) compared with the HC group (see Table 3). Supplementary Table 4 presents regressions predicting OSCARS global scores from factor scores with results suggesting both SC Bias and SC Ability factor scores are significantly related to the OSCARS global scores. Slightly larger estimates for the SC Ability factor score were observed across informant source and diagnostic group.

3.3. Internal consistency

Cronbach's alphas were calculated as a measure of internal consistency for the OSCARS total and factor scores for both self-report and informant-reported outcomes (αs .72–.92). McDonald's hierarchical omegas were also calculated (ωs .64–.91) and presented (see Table 5) as a more conservative and robust measure and demonstrated good internal consistency (Dunn et al., 2014; Peters, 2014).

3.4. Convergent and divergent validity

Convergent and divergent validity were assessed through correlations between OSCARS scores and measures of SC (convergent validity), NC (divergent validity; see Table 5), and symptoms (divergent validity; see Supplementary Table 5). OSCARS total and factor scores demonstrated convergent validity with significant small to medium correlations (*rs* −.12 to −.38) with at least one measure of SC (with the exception of SSD self-reported SC Bias scores). The OSCARS did not show strong support for divergent validity with significant small to medium

Table 4
OSCARS confirmatory factor analysis.

OSCARS Item	SSD Self-Report SC Ability	SSD Self-Report SC Bias	SSD Informant-Report SC Ability	SSD Informant-Report SC Bias	HC Self-Report SC Ability	HC Self-Report SC Bias
1. Recognizing other people's emotions, particularly negative emotions (sadness, fear and anger) based on facial expression, body language and/or vocal tone and rate?	.52		.75		.61	
6. Understanding subtle jokes, sarcasm and insults in conversation?	.60		.74		.50	
7. Seeing things from the perspective of others (i.e., putting themselves in other people's shoes)?	.67		.88		.48	
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)	.61		.84		.65	
2. Interpreting social interactions in a malevolent, hostile manner?		.60		.79		.59
3. Making decisions quickly (i.e., jumps to conclusions) without examining other evidence?		.67		.75		.63
4. Being flexible in interpreting social situations?		.67		.70		.66
5. Can change or correct their interpretation of social interactions when wrong?		.78		.73		.73
Model Fit Statistics	X^2 (19) = 37.2, $p < .01$ CFI = .97 TLI = .96 SRMR = .03 RMSEA = .06 [.03, .09]		X^2 (19) = 54.5, $p < .01$ CFI = .96 TLI = .94 SRMR = .04 RMSEA = .10 [.07, .14]		X^2 (19) = 69.3, $p < .01$ CFI = .91 TLI = .86 SRMR = .06 RMSEA = .10 [.07, .12]	

Note: SC = social cognition, SSD = schizophrenia spectrum disorder, HC = healthy control.

correlations (*r*s −.13 to −.38), essentially the same as those for SC, observed between NC composite scores and most OSCARS scores. OSCARS scores showed significant correlations with all PANSS scores except the negative symptoms subscale.

3.5. External validity

Overall, the OSCARS demonstrated adequate external validity with measures assessing functional outcomes. Most OSCARS scores were significantly correlated with at least one performance-based measures

Table 5
OSCARS, social cognition, neurocognition, and functional outcome correlations.

		SSD Self-Report				SSD Informant-Report				HC Self-Report			
	OSCARS Factors	Total	SC Ability	SC Bias	Global	Total	SC Ability	SC Bias	Global	Total	SC Ability	SC Bias	Global
	ω	.83	.69	.78	–	.91	.83	.88	–	.80	.64	.75	–
Divergent Validity	NC Composite	−.13*	−.14*	−.09	−.18**	−.09	−.13	−.04	−.19**	−.12*	−.14*	−.08	−.16**
Convergent Validity	AIHQ – Blame[a]	–	–	–	–	.15	.12	.15	.11	.37**	.29**	.36**	.16
	BLERT	−.06	−.07	−.04	−.12*	−.20**	−.19*	−.18*	−.28**	−.20**	−.21**	−.16**	−.19**
	ER40	−.09	−.14*	−.04	−.10	−.02	−.02	−.01	−.10	−.17**	−.15*	−.15*	−.11
	Eyes	−.15**	−.19**	−.10	−.19**	−.02	−.06	.03	−.11	−.15*	−.13*	−.13*	−.19**
	Hinting	−.07	−.06	−.07	−.10	−.08	−.12	−.03	−.14	−.06	−.07	−.03	−.01
	IBT[b]	.14*	.14*	.12	.10	.07	.08	.05	.14	.07	.00	.11	−.01
	TASIT	−.07	−.11	−.03	−.05	−.04	−.07	.00	−.13	−.15*	−.21**	−.07	−.24**
External Validity	<u>SSPA Average</u>	−.15*	−.17*	−.11	−.06	−.09	−.17*	−.02	−.14	−.12	−.13*	−.09	−.14*
	SSPA1	−.17*	−.19**	−.12	−.08	−.12	−.15	−.07	−.14	−.17**	−.19**	−.12	−.17**
	SSPA2	−.09	−.10	−.06	−.03	−.05	−.14	.05	−.11	−.05	−.05	−.04	−.08
	<u>SLOF Self-Report Total</u>	−.40**	−.35**	−.36**	−.39**	−.40**	−.39**	−.37**	−.38**	−.47**	−.43**	−.41**	−.31**
	Interpersonal Relationships	−.38**	−.33**	−.36**	−.26**	−.24**	−.26**	−.19*	−.26**	−.49**	−.42**	−.46**	−.36**
	Social Acceptability	−.30**	−.19**	−.34**	−.24**	−.28**	−.21*	−.31**	−.26**	−.28**	−.17**	−.33**	−.25**
	Community Living	−.20**	−.20**	−.16*	−.30**	−.34**	−.33**	−.31**	−.29**	−.24**	−.27**	−.18**	−.11
	Work Skills	−.28**	−.26**	−.25**	−.26**	−.24**	−.24**	−.21*	−.27**	−.33**	−.34**	−.26**	−.19**
	<u>SLOF Informant Total</u>	−.19*	−.14	−.20**	−.21**	−.60**	−.54**	−.57**	−.67**				
	Interpersonal Relationships	−.20**	−.14	−.21**	−.19*	−.53**	−.50**	−.49**	−.53**				
	Social Acceptability	−.29**	−.21**	−.30**	−.18*	−.56**	−.42**	−.62**	−.59**				
	Community Living	−.05	.01	−.08	−.13	−.30**	−.28**	−.29**	−.38**				
	Work Skills	−.15*	−.17*	−.11	−.15*	−.48**	−.46**	−.44**	−.58**				
	<u>UPSA Total</u>	−.14*	−.12*	−.13*	−.18**	−.13	−.20**	−.06	−.14				

Note:.
* $p < .05$.
** $p < .01$.
[a] AIHQ only collected during initial and second phases.
[b] IBT only collected during final phase.
AIHQ – Blame = Ambiguous Intentions and Hostility Questionnaire Blame Index, BLERT = Bell Lysaker Emotion Recognition Task, ER40 = Penn Emotion Recognition Task, Eyes = Reading the Mind in the Eyes Task, Hinting = Hinting Task, IBT = Intentionality Bias Task, TASIT = The Awareness of Social Inference Test, HVLT = Hopkins Verbal Learning Test, SSPA = Social Skills Performance Assessment, SLOF = Specific Levels of Functioning, UPSA = UCSD Performance-Based Skills Assessment, SSD = schizophrenia spectrum disorder, SC = social cognition; ω = McDonald's hierarchical omegas.

Table 6
Regressions predicting functional outcomes from OSCARS in SSD.

Outcome	Adjusted R^2	Change in R^2	Model Statistic	p
SSPA Average				
OSCARS Self-Report				
Total	.08	.01[a]	F (4, 222) = 5.8	<.01
Global	.08	<0.01	F (5, 221) = 4.7	<.01
OSCARS Informant-Report				
Total	.09	<.01[a]	F (4, 148) = 4.8	<.01
Global	.09	<0.01	F (5, 147) = 4.0	<.01
SLOF Self-Report Total				
OSCARS Self-Report				
Total	.17	.10 [a]*	F (4, 182) = 10.5	<.01
Global	.19	.02*	F (5, 181) = 10.0	<.01
OSCARS Informant-Report				
Total	.15	.07[a]*	F (4, 105) = 6.0	<.01
Global	.15	<0.01	F (5, 104) = 5.0	<.01
SLOF Informant Total				
OSCARS Self-Report				
Total	.07	.01[a]	F (4, 171) = 4.5	<.01
Global	.08	.01	F (5, 170) = 3.8	<.01
OSCARS Informant-Report				
Total	.36	.28[a]*	F (4, 165) = 24.8	<.01
Global	.44	.08*	F (5, 164) = 27.9	<.01
UPSA Total				
OSCARS Self-Report				
Total	.24	<0.01 [a]	F (4, 273) = 22.9	<.01
Global	.24	<0.01	F (5, 272) = 18.7	<.01
OSCARS Informant-Report				
Total	.20	<.01[a]	F (4, 161) = 11.5	<.01
Global	.20	<0.01	F (5, 160) = 9.2	<.01

Note: [a]change in R^2 from covariate (sex, years of education, WRAT3 score) model.
* indicates significant model improvement, p <.05; SSD = schizophrenia spectrum disorders; Due to significant correlations between predictors, variance inflation factor (VIF) statistics were calculated, VIF values for all predictors across models < 5 indicating acceptable multicollinearity; SSPA = Social Skills Performance Assessment, SLOF = Specific Levels of Functioning, UPSA = UCSD Performance-Based Skills Assessment, SSD = schizophrenia spectrum disorder.

of functioning (i.e., SSPA or UPSA) with small to medium correlations (*rs* −.12 to −.33, see Table 5), with the exception of SSD total informant-report, SSD informant-report SC Bias, and SSD informant-report global score. All OSCARS informant-reported scores also demonstrated external validity with self and informant-reported measures (i.e., SLOF) of functioning (*rs* −.67 to −.19, *ps* < .05). Most of the OSCARS self-report scores demonstrated external validity with self and informant-reported SLOF scores (*rs* −.39 to −.15), with the notable exception of non-significant relationships between all OSCARS scores and SLOF informant-reported Community Living Subscales, and some OSCARS scores with SLOF Interpersonal Relationships and Work Skills.

3.6. Predictive validity

Predictive validity was examined through a series of hierarchical regressions predicting functional outcomes from OSCARS total scores after accounting for the variance explained by sex, years of education, and WRAT scores (see Table 6). For each analysis, predictor variables were entered into the model in the following order 1) sex, years of education, and WRAT scores 2) OSCARS total score 3) OSCARS global score. The OSCARS self-report and informant-based total scores were shown to significantly predict functioning on both the SLOF self-report (OSCARS self-report total, $F(4, 182) = 10.5$, $p < .01$, $\Delta R^2 = 0.10$; OSCARS self-report global score, $F(5, 181) = 10.0$, $p < .01$, $\Delta R^2 = 0.02$; OSCARS informant total $F(4, 105) = 6.0$, $p < .01$, $\Delta R^2 = 0.07$), and SLOF informant-report total (OSCARS informant total $F(4, 165) = 24.8$, $p < .01$, $\Delta R^2 = 0.28$; OSCARS global score $F(5, 164) = 27.9$, $p < .01$, $\Delta R^2 = 0.08$). Self- and informant-reported OSCARS scores did not predict UPSA or SSPA performance.

Table 7
Regressions predicting social cognition task performance from OSCARS in SSD.

Outcome	Adjusted R^2	Change in R^2	Model Statistic	p
SC Composite				
OSCARS Self-Report				
Total	.29	<0.01[a]	F(4, 275) = 29.0	<.01
Global	.29	<0.01	F (5, 274) = 23.3	<.01
OSCARS Informant-Report				
Total	.29	<0.01[a]	F(4, 160) = 17.9	<.01
Global	.30	.01	F (5, 159) = 14.9	<.01

[a] change in R^2 from covariate (sex, years of education, WRAT3 score) model; *indicates significant model improvement, $p < 0.05$; SC = social cognition; Due to significant correlations between predictors, variance inflation factor (VIF) statistics were calculated, VIF values for all predictors across models < 5 indicating acceptable multicollinearity.

Predictive validity of the OSCARS was also examined through a series of hierarchical regressions predicting SC task performance from OSCARS total scores (see Table 7). Self- and informant-reported OSCARS scores did not predict SC task performance.

3.7. Predictive validity identifying social cognition impairment

ROC analyses yielding AUC values examined the diagnostic utility of the OSCARS to identify individuals with SSD for comparison with Healey et al. (2015). Generally, AUCs of .70 −.79 reflect fair, .80 - .89 good, and .90 – 1.00 excellent values (Swets, 1988). AUC values for predicting SSD diagnoses were fair for OSCARS self-reported (AUC = .77 [95% CI .73, .81]) and informant-reported (AUC = .76 [95% CI .68, .85]) total scores.

However, AUC values may be inflated when distinguishing a clinical sample from a HC sample (Whiting et al., 2003). Therefore, predictive utility of the OSCARS was also examined in a more conservative manner by distinguishing individuals with impaired SC task performance (i.e., SC composite score 1.5 standard deviations below average HC performance) from the rest of the sample. AUC values for the OSCARS self-reported total (AUC = .69 [95% CI .61, .77]) and global scores (AUC = .68 [95% CI .60, .76]) were fair. OSCARS informant-reported total (AUC = .58 [95% CI .46, .70]) and global scores (AUC = .64 [95% CI .51, .76]) were less impressive, albeit not significantly lower than self-report (Venkatraman's test *ps* > .05, see Table 8). OSCARS total and global score cut points for identifying impaired SC based on optimal sensitivity and specificity values are presented in Table 8.

Table 8
OSCARS predicting impaired social cognition task performance.

OSCARS Scale	AUC [95% CI]	Score Range Low Risk	Score Range Elevated Risk	Sensitivity	Specificity
OSCARS Self-Report					
Total*	.69 [.61, 0.77]	≤ 15	> 15	.59	.72
Global*	.68 [.60, 0.76]	≤ 3	> 3	.56	.70
OSCARS Informant-Report					
Total	.58 [.46, 0.70]	–	–	–	–
Global*	.64 [.51, 0.76]	≤ 2	> 2	.21	.90

Note: AUC = Area under the curve; Impaired social cognition task performance indicated by performance 1.5 standard deviations or more below average performance of healthy control sample;.
* $p < 0.05$.

Table 9
OSCARS predicting impaired functioning.

Functional Measure	AUC [95% CI]	Score Range		Sensitivity	Specificity
		High Risk	Low Risk		
SLOF Self-Report					
OSCARS Self-Report					
Total*	.82 [.77, 0.87]	≤ 16	>16	.82	.74
Global*	.81 [.75, 0.86]	≤ 3	>3	.75	.77
OSCARS Informant-Report					
Total*	.75 [.66, 0.84]	≤ 14	>14	.92	.45
Global*	.74 [.65, 0.83]	≤ 4	>4	.70	.77
SLOF Informant-Report					
OSCARS Self-Report					
Total*	.62 [.54, 0.70]	≤ 17	>17	.64	.56
Global*	.64 [.56, 0.72]	≤ 2	>2	.80	.41
OSCARS Informant-Report					
Total*	.85 [.79, 0.90]	≤ 22	>22	.84	.76
Global*	.85 [.79, 0.91]	≤ 4	>4	.80	.80
SSPA					
OSCARS Self-Report					
Total*	.65 [.59, 0.71]	≤ 18	>18	.49	.75
Global*	.64 [.58, 0.71]	≤ 2	>2	.59	.59
OSCARS Informant-Report					
Total*	.60 [.52, 0.69]	≤ 19	>19	.69	.54
Global*	.61 [.52, 0.70]	≤ 2	>2	.76	.45
UPSA[a]					
OSCARS Self-Report					
Total*	.61 [.54, 0.67]	≤ 21	>21	.47	.75
Global*	.59 [.52, 0.66]	≤ 3	>3	.69	.53
OSCARS Informant-Report					
Total	.59 [.50, 0.67]	–	–	–	–
Global	.58 [.50, 0.67]	–	–	–	–

Note: AUC = Area under the curve, SLOF = Specific Levels of Functioning Scale, SSPA = Social Skills Performance Assessment, UPSA = UCSD Performance-Based Skills Assessment; Impaired functioning indicated by performance 1.5 standard deviations or more below average performance of healthy control sample;.

[a] No healthy control comparison sample so impaired functioning indicated by published cut-off score of 75 (Mausbach et al., 2008);.

* $p < 0.05$.

3.8. Exploratory predictive validity identifying functional impairment

A review of correlations between the OSCARS and other measures revealed roughly equivalent associations between the OSCARS, SC, and NC with strong correlations observed with functional outcomes (Table 5). Given this finding, as well as recent work demonstrating relationships between the OSCARS and functioning (e.g., Oliveri et al., 2020; Silberstein et al., 2018), exploratory analyses examined the predictive validity of the OSCARS identifying impaired functioning. Impaired functioning was defined as 1.5 standard deviations or more below average HC performance (i.e., SSPA and SLOF). A published cut score for UPSA impairment (i.e., a score of 75; Mausbach et al., 2008) was used since this measure was only administered to the SSD group. All AUC values for the OSCARS predicting functional impairment were significant with the exception of impairment on the UPSA. Significant AUC values were fair to good (i.e., .60 - .82; see Table 9 for AUC values and suggested OSCARS cut scores to identify impairment).

4. Discussion

The present study is the first to demonstrate the psychometric properties of the OSCARS when administered as a self-report measure. It also replicated findings from Healey et al. (2015) demonstrating the psychometric reliability of the OSCARS when administered as an informant-based measure in an independent sample over five times the size of the original validation sample. Overall, results indicate the OSCARS is a reliable, easily administered, and time-efficient measure. Surprisingly, the predictive validity of the OSCARS predicting task-based SC performance and identifying SC impairments was only fair. Follow-up analyses revealed improved predictive validity identifying impairments in real-world functioning (i.e., SLOF), suggesting the OSCARS may be better conceptualized as a measure of functioning.

The OSCARS exhibited good construct validity (Aim 1) with factor analyses confirming a two-dimensional model with interpretable factors across informant-report and self-report administration. This two-factor structure (i.e., SC Ability and SC Bias) replicates previous findings for separate constructs of skill and bias in SSD (Buck et al., 2016). Good construct validity was also evidenced by significantly better HC group scores compared with the SSD group on all OSCARS scores.

The OSCARS displayed good internal consistency and modest convergent validity with significant small to medium correlations with all SC tasks (Aim 3). The SCOPE study suggested SC tasks were not impressive in their prediction of everyday outcomes and instead exhibited stronger correlations with performance-based measures of NC (Pinkham et al., 2018). Consistent with this idea, small to medium correlations between the OSCARS and some functional outcomes demonstrated external validity. Divergent validity of the OSCARS was equivocal with significant correlations observed with measures of NC but no significant correlations with negative symptoms. Recent findings suggest NC and SC may be less distinct than previously thought, but roughly equivalent relationships between NC and SC with the OSCARS is surprising and suggests potential low divergent validity (Deckler et al., 2018). Alternately, previous work demonstrates negative symptoms and SC are distinct domains and non-significant relationships observed between the OSCARS and negative symptoms offers some support for divergent validity (Bell et al., 2013; Sergi et al., 2007).

Finally, the OSCARS also demonstrated equivocal predictive validity dependent on the type of outcome measure. Self-report and informant-report showed incremental validity predicting functional outcomes above and beyond demographic characteristics on self and informant-reported SLOF scores but not UPSA or SSPA performance. This null finding may reflect modest rather than strong correlations with functional outcomes, a replication of previous studies examining relationships between report-style cognition measures (e.g., Cognitive Assessment Interview) and functional outcomes (Ventura et al., 2008). SC Ability and SC Bias each demonstrated significant relationships with the OSCARS global score (Aim 2) suggesting both factors influence self and informant-reported overall assessment of SC.

After establishing psychometric properties, three hypotheses investigated the performance and utility of the OSCARS. The first hypothesis received modest support. Informant-reported OSCARS scores significantly predicted informant and self-reported SLOF scores but not SSPA or UPSA performance. Results also modestly support the second hypothesis: OSCARS informant-report significantly predicted impaired SC task performance and ROC analyses predicting SSD diagnostic group and impaired SC demonstrated fair predictive validity. Interestingly, OSCARS self-report exhibited stronger AUC values for identifying impaired SC performance compared with informant-report, although this difference was not statistically significant. One explanation for this finding may be the inclusion of attribution bias in our SC composite score. Self-assessment has been shown more effective in evaluating effects of attribution style on outcomes (Vidarsdottir et al., 2019).

Exploratory analyses predicting functional *impairment* were conducted since the OSCARS demonstrated roughly equivalent correlations with SC and NC measures but noticeably stronger relationships with functional outcomes. The OSCARS demonstrated better utility predicting real-world functional impairments compared with SC impairments (i.e., the highest AUC values were observed when using the OSCARS to predict functional impairment measured by the SLOF). This finding provides preliminary support for the OSCARS as a measure of real-world functioning in SSD. However, the utility of the OSCARS as a screener for impairments in functioning was exploratory and future validation research is needed.

Good psychometric properties, modest correlations with performance-based SC tasks and NC tasks, and strong relationships with functional outcomes suggest the OSCARS may be best utilized as an efficient screening tool for clinicians to identify individuals with impaired functioning to receive psychosocial treatments or additional assessment. The OSCARS is a free instrument with an administration time of five minutes, is available as a paper questionnaire, and does not require a trained rater for administration or scoring. Additionally, results support the validity of both self- and informant-report versions. These qualities make the OSCARS an attractive and easily disseminated first-step measure to be used early in treatment initiation and planning. Early detection of impairment in functioning may facilitate enrollment in psychosocial interventions yielding improvements in treatment outcomes (e.g., Mueser et al., 2013).

Informants included in the present study were heterogenous with significant differences in informant roles observed between SSD and HC groups. Informant role likely influences the quality of relationship and subsequent informant-reported abilities. Future work should collect quality ratings of informant-individual relationships, in addition to quantity of time spent together, to investigate potential influences on OSCARS performance. Finally, some validity analyses (e.g., external and predictive validity) utilized the same informant source (e.g., informant-reported OSCARS and informant-reported SLOF) and significant findings may reflect shared measurement variance. However, method variance does not account for incremental validity in predicting functional outcomes and impairment from different informant sources (e.g., informant-reported OSCARS explains additional variance in self-reported SLOF outcomes). Another limitation is the use of an abbreviated MATRICS Battery to assess divergent validity. Significant relationships observed between the OSCARS and NC are congruent with recent models demonstrating significant and complex relationships between NC and SC, especially as they relate to functioning (e.g., Green et al., 2019; Hasson-Ohayon et al., 2018). However, the similar correlation coefficients observed across both NC and SC is surprising as stronger correlations were expected with SC. These findings suggest potential low divergent validity of the OSCARS and highlight the need to identify stronger measures of divergent validity (e.g., domains where no associations are expected) for assessing measures of SC.

Notwithstanding these limitations, the present study provides support for the OSCARS as an adequate brief measure of SC deficits in SSD with preliminary support that the OSCARS may be a good measure to identify impairment in functioning. Additionally, the OSCARS can be flexibly administered in either an informant or self-report format. Future work is needed to investigate discrepancies between self-reports and informant impressions as these may provide useful information about limitations in self-assessment of functioning and SC abilities. Both formats demonstrated acceptable psychometric properties and good validity with some support for use as a screening tool to detect impairment in SC and more promising preliminary support to detect impairment in real-world functioning. When making administration decisions, clinicians are encouraged to consider time-constraints and availability of informants given roughly equivalent psychometric properties of self and informant-report. Overall, findings best support that the OSCARS may be a useful first-step tool for clinicians to identify real-world functioning deficits and efficiently identify individuals in need of additional assessment or psychosocial interventions.

Contributors

AEP, PDH and DLP designed the study and wrote the protocol. TFH wrote the first draft of the manuscript. TFH performed statistical analyses and certifies the accuracy of the results. All authors contributed to the data interpretation, meaningful manuscript revision, and all authors have approved the final manuscript.

Declaration of Competing interest

MH received a fee from Lundbeck as a speaker at an education grant conference and travel support from Angelini. In the last year, AEP received consulting fees and travel reimbursement from Roche. PDH reports the following conflicts of interest within the past year: consulting fees or travel reimbursements from Alkermes, Boehringer Ingelheim, Intra-Cellular Therapies, Minerva Pharma, Otsuka America, Regeneron Pharma, Roche Pharma, Sunovion Pharma, Takeda Pharma, and Teva; Royalties from the Brief Assessment of Cognition in Schizophrenia; He is chief scientific officer of i-Function, Inc.; Research grants from Takeda and from the Stanley Medical Research Foundation. LFJ has received grant support from Auspex/Teva, Boehringer Ingelheim, Otsuka, NIH and has served as a consultant for Bracket and UpToDate. TFH, LN, and DLP report no conflicts of interest.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.psychres.2020.112891.

References

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., Plumb, I., Hill, J., ... Wheelwright, S., 2001. The "Reading the mind in the eyes" test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism. J. Child Psychol. Psychiatry 42 (2), 241–251. https://doi.org/10.1111/1469-7610.00715.

Barrett, P., 2007. Structural equation modelling: adjudging model fit. Pers Individ Dif 42 (5), 815–824. https://doi.org/10.1016/j.paid.2006.09.018.

Bell, M., Bryson, G., Lysaker, P., 1997. Positive and negative affect recognition in schizophrenia: a comparison with substance abuse and normal control subjects. Psychiatry Res. 73 (1–2), 73–82.

Bell, M.D., Corbera, S., Johannesen, J.K., Fiszdon, J.M., Wexler, B.E., 2013. Social cognitive impairments and negative symptoms in schizophrenia: are there subtypes with distinct functional correlates? Schizophr. Bull. 39 (1), 186–196. https://doi.org/10.1093/schbul/cbr125.

Bentler, P.M., 2007. On tests and indices for evaluating structural models. Pers. Individ. Dif. 42 (5), 825–829. https://doi.org/10.1016/j.paid.2006.09.024.

Buck, B.E., Healey, K.M., Gagen, E.C., Roberts, D.L., Penn, D.L., 2016. Social cognition in schizophrenia: factor structure, clinical and functional correlates. J. Mental Health 25 (4), 330–337. https://doi.org/10.3109/09638237.2015.1124397.

Buck, B., Kern, R.S., Marder, S.R., Penn, D.L., Healey, K.M., Green, M.F., ... Penn, D.L., 2017. Improving measurement of attributional style in schizophrenia; a psychometric evaluation of the ambiguous intentions hostility questionnaire (AIHQ). Journal of Psychiatric Research 89, 48–54. https://doi.org/10.1016/j.jpsychires.2017.01.004.

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. Cogn Neuropsychiatry 12 (2), 128–143. https://doi.org/10.1080/13546800600787854.

Corcoran, R., Mercer, G., Frith, C.D., 1995. Schizophrenia, symptomatology and social inference: investigating "theory of mind" in people with schizophrenia. Schizophr. Res. 17 (1), 5–13.

Deckler, E., Hodgins, G.E., Pinkham, A.E., Penn, D.L., Harvey, P.D., 2018. Social cognition and neurocognition in schizophrenia and healthy controls: intercorrelations of performance and effects of manipulations aimed at increasing task difficulty. Front. Psychiatry 9 (AUG), 1–10. https://doi.org/10.3389/fpsyt.2018.00356.

Dunn, T.J., Baguley, T., Brunsden, V., 2014. From alpha to omega: a practical solution to the pervasive problem of internal consistency estimation. Brit. J. Psychol. (London, England : 1953) 105 (3), 399–412. https://doi.org/10.1111/bjop.12046.

Fett, A.K.J., Viechtbauer, W., Dominguez, M., de, G., Penn, D.L., van Os, J., Krabbendam, L., 2011. The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. Neurosci. Biobehav. Rev. 35 (3),

573–588. https://doi.org/10.1016/j.neubiorev.2010.07.001.
First, M.B., Spitzer, R.L., Gibbon, M., Williams, J.B.W., 2002. Structured Clinical Interview For DSM-IV-TR Axis I Disorders, Research version, Patient Edition. SCID-I/P.
Grant, N., Lawrence, M., Preti, A., Wykes, T., Cella, M., 2017. Social cognition interventions for people with schizophrenia: a systematic review focussing on methodological quality and intervention modality. Clin. Psychol. Rev. 56 (November 2016), 55–64. https://doi.org/10.1016/j.cpr.2017.06.001.
Green, M.F., Horan, W.P., Lee, J., 2019. Nonsocial and social cognition in schizophrenia: current evidence and future directions. World Psychiatry 18 (2), 146–161 https:doi.org/10.1002/wps.20624.
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. https://doi.org/10.1093/schbul/sbm145.
Hajdúk, M., Harvey, P.D., Penn, D.L., Pinkham, A.E., 2018. Social cognitive impairments in individuals with schizophrenia vary in severity. J. Psychiatr. Res. 104 (June), 65–71. https://doi.org/10.1016/j.jpsychires.2018.06.017.
Halverson, T.F., Orleans-Pobee, M., Merritt, C., Sheeran, P., Fett, A.-K., Penn, D.L., 2019. Pathways to functional outcomes in schizophrenia spectrum disorders: meta-analysis of social cognitive and neurocognitive predictors. Neurosci. Biobehav. Rev. 105 (July), 212–219. https://doi.org/10.1016/j.neubiorev.2019.07.020.
Harvey, P.D., Deckler, E., Jones, M.T., Jarskog, L.F., Penn, D.L., Pinkham, A.E., 2019. Autism symptoms, depression, and active social avoidance in schizophrenia: association with self-reports and informant assessments of everyday functioning. J. Psychiatr. Res. 115 (April), 36–42. https://doi.org/10.1016/j.jpsychires.2019.05.010.
Hasson-Ohayon, I., Goldzweig, G., Lavi-Rotenberg, A., Luther, L., Lysaker, P.H., 2018. The centrality of cognitive symptoms and metacognition within the interacting network of symptoms, neurocognition, social cognition and metacognition in schizophrenia. Schizophr. Res. 202, 260–266. https://doi.org/10.1016/j.schres.2018.07.007.
Healey, K.M., Combs, D.R., Gibson, C.M., Keefe, R.S.E., Roberts, D.L., Penn, D.L., 2015. Observable social cognition - A Rating scale: an interview-based assessment for schizophrenia. Cogn. Neuropsychiatry 20 (3), 198–221. https://doi.org/10.1080/13546805.2014.999915.
Hu, L., Bentler, P.M., 1999. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct. Equ. Model. 6 (1), 1–55. https://doi.org/10.1080/10705519909540118.
Jones, M.T., Deckler, E., Laurrari, C., Jarskog, L.F., Penn, D.L., Pinkham, A.E., Harvey, P.D., 2019. Confidence, performance, and accuracy of self-assessment of social cognition: a comparison of schizophrenia patients and healthy controls. Schizophr. Res. (November 2018), 0–1. https://doi.org/10.1016/j.scog.2019.01.002.
Kay, S.R., Fiszbein, A., Opler, L.A., 1987. The positive and negative syndrome scale (PANSS) for schizophrenia. Schizophr. Bull. 13 (2), 261–276.
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. Am. J. Psychiatry 163 (3), 426–432. https://doi.org/10.1176/appi.ajp.163.3.426.
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. https://doi.org/10.1176/appi.ajp.160.10.1768.
Kurtz, M.M., Gagen, E., Rocha, N.B.F., Machado, S., Penn, D.L., 2016. Comprehensive treatments for social cognitive deficits in schizophrenia: a critical review and effect-size analysis of controlled studies. Clin. Psychol. Rev. 43, 80–89. https://doi.org/10.1016/j.cpr.2015.09.003.
MacCallum, R.C., Browne, M.W., Sugawara, H.M., 1996. Power analysis and determination of sample size for covariance structure modeling. Psychol. Methods 1 (2), 130–149. https://doi.org/10.1037/1082-989X.1.2.130.
Mausbach, B.T., Bowie, C.R., Harvey, P.D., Twamley, E.W., Goldman, S.R., Jeste, D.V., Patterson, T.L., 2008. Usefulness of the ucsd performance-based skills assessment (UPSA) for predicting residential independence in patients with chronic schizophrenia. J Psychiatr. Res. 42 (4), 320–327. https://doi.org/10.1016/j.jpsychires.2006.12.008.
Mausbach, B.T., Harvey, P.D., Goldman, S.R., Jeste, D.V., Patterson, T.L., 2007. Development of a brief scale of everyday functioning in persons with serious mental illness. Schizophr. Bull. 33 (6), 1364–1372. https://doi.org/10.1093/schbul/sbm014.
McDonald, S., Flanagan, S., Rollins, J., Kinch, J., 2003. TASIT: a new clinical tool for assessing social perception after traumatic brain injury. J. Head Trauma Rehabilit. 18 (3), 219–238.
Mueser, K., Deavers, F., Penn, D., Cassisi, J., 2013. Psychosocial treatments for schizophrenia. Annu. Rev. Clin. Psychol. 9, 465–497. https://doi.org/10.1177/0963721410377743.
Nuechterlein, K.H., Green, M.F., Kern, R.S., Baade, L.E., Barch, D.M., Cohen, J.D., ... Marder, S.R., 2008. The matrics consensus cognitive battery, part 1: Test selection, reliability, and validity. The American Journal of Psychiatry 165 (2), 203–213. https://doi.org/10.1176/appi.ajp.2007.07010042.
Oliveri, L.N., Awerbuch, A.W., Jarskog, L.F., Penn, D.L., Pinkham, A., Harvey, P.D., 2020. Depression predicts self-assessment of social function in both patients with schizophrenia and healthy people. Psychiatry Res. 284. https://doi.org/10.1016/j.psychres.2019.112681.
Patterson, T.L., Moscona, S., McKibbin, C.L., Davidson, K., Jeste, D.V., 2001. Social skills performance assessment among older patients with schizophrenia. Schizophr. Res. 48 (2–3), 351–360.
Penn, D.L., Corrigan, P.W., Bentall, R.P., Racenstein, J.M., Newman, L., 1997. Social cognition in schizophrenia. Psychol. Bull. 121 (1), 114–132. https://doi.org/10.1037/0033-2909.121.1.114.
Penn, D.L., Sanna, L.J., Roberts, D.L., 2008. Social cognition in schizophrenia: an overview. Schizophr. Bull. 34 (3), 408–411. https://doi.org/10.1093/schbul/sbn014.
Peters, G.Y., 2014. The alpha and the omega of scale reliability and validity. Eur. Health Psychol. 1 (2), 54–72. https://doi.org/10.31234/osf.io/h47fv.
Pinkham, A.E., Harvey, P.D., Penn, D.L., 2016a. Paranoid individuals with schizophrenia show greater social cognitive bias and worse social functioning than non-paranoid individuals with schizophrenia. Schizophr. Res. 3, 33–38. https://doi.org/10.1016/j.scog.2015.11.002.
Pinkham, A.E., Harvey, P.D., Penn, D.L., 2018. Social cognition psychometric evaluation: results of the final validation study. Schizophr. Bull. 44 (4), 737–748. https://doi.org/10.1093/schbul/sbx117.
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. Schizophr. Bull. 40 (4), 813–823. https://doi.org/10.1093/schbul/sbt081.
Pinkham, A.E., Penn, D.L., Green, M.F., Harvey, P.D., 2016b. Social cognition psychometric evaluation: results of the initial psychometric study. Schizophr. Bull. 42 (2), 494–504. https://doi.org/10.1093/schbul/sbv056.
Rocca, P., Galderisi, S., Rossi, A., Bertolino, A., Rucci, P., Gibertoni, D., ... De Capua, A., 2016. Social cognition in people with schizophrenia: A cluster-analytic approach. Psychological Medicine 46 (13), 2717–2729. https://doi.org/10.1017/S0033291716001100.
Rosseel, Y., 2012. lavaan: an r package for structural equation modeling human forearm during rythmic exercise. J. Stat. Softw. 48 (2), 1–36. https://doi.org/10.18637/jss.v048.i02.
Rosset, E., 2008. It's no accident: our bias for intentional explanations. Cognition 108 (3), 771–780. https://doi.org/10.1016/j.cognition.2008.07.001.
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. Schizophr. Bull. 39 (5), 979–992. https://doi.org/10.1093/schbul/sbs080.
Schneider, L.C., Struening, E.L., 1983. SLOF: a behavioral rating scale for assessing the mentally ill. Soc. Work. Res. Abstr. 19 (3), 9–21.
Sergi, M.J., Green, M.F., Widmark, C., Reist, C., Erhart, S., Braff, D.L., ... Mintz, J., 2007. Social cognition and neurocognition: Effects of risperidone, olanzapine, and haloperidol. American Journal of Psychiatry 164 (10), 1585–1592. https://doi.org/10.1176/appi.ajp.2007.06091515.
Sheehan, D.V., Lecrubier, Y., Sheehan, K.H., Amorim, P., Janavs, J., Weiller, E., ... Dunbar, G.C., 1998. The mini-international neuropsychiatric interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. The Journal of Clinical Psychiatry 59 (Suppl 2), 22–57.
Silberstein, J., Harvey, P.D., 2019. Cognition, social cognition, and self-assessment in schizophrenia: prediction of different elements of everyday functional outcomes. CNS Spectr. 1–6. https://doi.org/10.1017/S1092852918001414.
Silberstein, J.M., Pinkham, A.E., Penn, D.L., Harvey, P.D., 2018. Self-assessment of social cognitive ability in schizophrenia: association with social cognitive test performance, informant assessments of social cognitive ability, and everyday outcomes. Schizophr. Res. 199, 75–82. https://doi.org/10.1016/j.schres.2018.04.015.
Swets, J.A., 1988. Measuring the accuracy of diagnostic systems. Science 240 (4857), 1285–1293.
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. Schizophr. Res. 106 (1), 59–69. https://doi.org/10.1016/j.schres.2007.07.025.
Vidarsdottir, O.G., Twamley, E.W., Roberts, D.L., Gudmundsdottir, B., Sigurdsson, E., Magnusdottir, B.B., 2019. Social and non-social measures of cognition for predicting self-reported and informant-reported functional outcomes in early psychosis. Scand. J. Psychol. 60, 295–303. https://doi.org/10.1111/sjop.12549.
Weickert, T.W., Goldberg, T.E., Gold, J.M., Bigelow, L.B., Egan, M.F., Weinberger, D.R., 2000. Cognitive impairments in patients with schizophrenia displaying preserved and compromised intellect. Arch. Gen. Psychiatry 57 (9), 907–913.
Whiting, P., Rutjes, A.W.S., Reitsma, J.B., Bossuyt, P.M.M., Kleijnen, J., 2003. The development of quadas a tool for quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med. Res. Methodol. 3, 1–13. https://doi.org/10.1186/1471-2288-3-25.