Schizophrenia

ARTICLE OPEN

Harmonizing cross-cultural and transdiagnostic assessment of social cognition by expert panel consensus

Amy E. Pinkham[1]✉, Michal Hajdúk[2,3,4], Tim Ziermans[5] and the SIRS Social Cognition Research Harmonization Group*

Social cognition, the perception and processing of social information, is adversely affected in multiple psychiatric, neurological, and neurodevelopmental disorders, and these impairments negatively impact quality of life for individuals across the globe. Despite the clear importance of social cognition, efforts to advance research via harmonization of data across cultures and diagnoses has been stymied by the lack of uniformly used and suitable assessments. To address this issue, the current study conducted an expert survey and consensus process to identify social cognitive assessments that are best suited for cross-cultural and transdiagnostic use among adults. A large group of experts in social cognition were surveyed to gather nominations for cross-culturally and transdiagnostically appropriate measures. These measures were then critically evaluated by a smaller group of experts using a Delphi consensus process to identify the best existing tasks for each use. Ninety-eight experts, representing 25 countries, responded to the initial survey and nominated a total of 81 tasks. Initial rounds of the Delphi process identified 50 tasks with adequate psychometric properties that were then subdivided into social cognition domains. For each domain, members ranked the five best tasks, once for cross-cultural use and once for transdiagnostic use, and rated the suitability of those tasks for the intended use. No tasks were identified as ideally suited for either use; however, within each domain, 4–5 tasks emerged as the most consistently selected, and all were ranked as having "good" or better suitability for use. While there is still a critical need for social cognitive assessments that are specifically designed for cross-cultural and transdiagnostic use, there does appear to be a handful of existing tasks that are currently available and likely informative. Caution is warranted however, as these still require comprehensive evaluation in cross-cultural and transdiagnostic studies.

Schizophrenia (2025)11:25 ; https://doi.org/10.1038/s41537-024-00540-7

INTRODUCTION

Social-cognitive deficits and biases are common among individuals with schizophrenia spectrum disorders[1–5]. These difficulties are related to various negative functional outcomes such as poorer community functioning, underdeveloped social skills, and less effective social problem-solving[6]. Recently, two important expert reviews supported a key role of social cognition in the assessment and treatment of psychosis. First, the consensus statement from a group of world leading researchers and clinicians identified social cognition as one of the domains that is crucial for precise clinical characterization and treatment planning for individuals with schizophrenia[7]. In addition, the schizophrenia section of the European Psychiatric Association also recently formulated a guidance paper with recommendations for optimal assessment of, among others, social cognition[8]. In their paper, they recommend assessment of social cognition mainly for the characterization of patients, as well as for personalized treatment planning.

Despite these calls for the routine assessment of social cognition, measurement continues to be a significant challenge, particularly regarding which constructs to consider and how to best measure them[9]. Many social cognition tasks were originally developed in the context of autism spectrum disorder research (involving false beliefs and verbal/visual mentalizing tasks) but subsequent research has demonstrated that these tasks show poor psychometric properties and/or construct validity in adult populations (e.g.,[10,11]). Likewise, assessments derived from social neuroscience studies in healthy adults have also shown relatively weak psychometric properties in schizophrenia[12]. Most notably, the Social Cognition Psychometric Evaluation project (SCOPE[13,14]), a NIMH-funded, multi-round study focusing on the identification of sound social cognitive measures for use in clinical trials of schizophrenia, identified just three measures with sufficient psychometric properties that could be recommended for further use: the Hinting Task[15], the Penn Emotion Recognition Task (ER-40)[16], and the Bell-Lysaker Emotion Recognition Task (BLERT)[17].

While the SCOPE recommended tasks have been heavily used in the United States and United Kingdom, two of them show limited utility for large-scale international trials. The Hinting Task, for example, is a verbal task strongly influenced by social norms and knowledge, with vignettes that may not be applicable across cultures[18]. The BLERT utilizes low-quality videos of a single white male depicting various emotions that may not accurately represent diverse cultural contexts, particularly in light of the well-established other-race effect, which is also evident among individuals with schizophrenia and which may contribute to poorer performance in non-white individuals[19]. As these examples demonstrate, the role of culture in social cognitive performance is well established[20], and thus individuals may appear to have more impaired social cognitive functioning when tasks are not matched to culture. Although SCOPE did not consider cross-cultural applicability in its evaluation criteria, the results underscore both the paucity of high-quality social cognitive tasks as well as the broad lack of tasks that can be used cross-culturally. The need for

[1]Department of Psychology, School of Behavioral and Brain Sciences, The University of Texas at Dallas, Richardson, TX, USA. [2]Department of Psychology, Faculty of Arts, Comenius University in Bratislava, Bratislava, Slovak Republic. [3]The Centre for Psychiatric Disorders Research, Science Park, Comenius University in Bratislava, Bratislava, Slovak Republic. [4]Department of Psychiatry, Faculty of Medicine, Comenius University in Bratislava, Bratislava, Slovak Republic. [5]Dutch Autism & ADHD Research Center, Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands. *A list of authors and their affiliations appears at the end of the paper. ✉email: amy.pinkham@utdallas.edu

Published in partnership with the Schizophrenia International Research Society

culturally sensitive social cognitive assessments has long been emphasized[8,21,22], and the unavailability of such tasks renders harmonization efforts impossible.

In addition, interest in understanding social cognitive impairments is not limited to schizophrenia spectrum illnesses. Social cognition is adversely affected in numerous psychiatric (e.g., bipolar disorder, post-traumatic stress disorder, anorexia nervosa), neurological (e.g., traumatic brain injury, stroke, frontotemporal dementia, Parkinson's Disease, Alzheimer's Disease), and neurodevelopmental conditions (e.g., Autism, Attention Deficit Hyperactivity Disorder)[23,24]. By and large, these impairments are of moderate to large effect sizes and in many cases exceed the magnitude of other cognitive impairments that are also seen in these conditions. As such, experts are increasingly calling for the incorporation of social cognitive assessment into clinical practice and note that patterns of social cognitive impairment may inform differential diagnosis (e.g., parsing frontotemporal dementia from primary psychiatric disorder) and disease progression[25]. Harmonization efforts such as these require measures that can be utilized transdiagnostically, in which task psychometric properties remain similar across different clinical groups and meaningful comparisons between clinical groups can be made. Successful identification of such tasks would facilitate efforts to examine the possibility of shared vs. distinct etiologies of social cognitive impairments across disorders and to identify disorder specific difficulties that would inform clinical decisions and treatment planning.

To address these measurement limitations and challenges within social cognitive research, a group of experts was convened to form the Schizophrenia International Research Society (SIRS) Social Cognition Research Harmonization Group (RHG). The goals of the RHG were twofold: (1) conduct a wide-ranging expert survey to gather nominations for tasks of social cognition that may be well-suited to cross-cultural and transdiagnostic use among adults, and (2) use the Delphi Method within our RHG to identify a consensus set of social cognitive measures from these nominations for use in future data collection. Although cross-cultural and transdiagnostic applicability are separate concepts and usually examined individually, the current study sought to leverage the expertise of the field and the RHG to advance social cognitive assessment on both fronts. In doing so, it was also hoped that some tasks would be identified as suitable for both cross-cultural *and* transdiagnosic use, which could provide an immediate foundation for collaborative projects. This paper reports the outcomes of these efforts.

RESULTS

Expert survey

Fifty-two tasks were nominated for cross-cultural use, and 77 tasks were nominated for transdiagnostic use. Forty-seven tasks were nominated as potentially suitable for use pending further development. Many tasks were nominated across uses, resulting in a total of 81 unique tasks across all categories. These tasks are listed alphabetically in Table 1.

Delphi Round 1

Twenty-one members of the RHG provided complete ratings for each of the 81 tasks. Forty-seven tasks received mean ratings ≥2 (adequate or better). Eleven of these tasks had an IQR > 1 and were included in Round 2. The remaining 36 tasks receiving consensus ratings of ≥2 were retained for further evaluation in Round 3. Thirty-three tasks scored in the inadequate range, and of these, 26 had consensus average ratings of <2 (i.e., average score of <2 and IQR of ≤1) and were therefore removed from further consideration. The remaining 7 were carried forward to Round 2. IQR could not be calculated for one task due to the lack of non-zero scores, and this task was also omitted from further consideration. Average Round 1 ratings and IQR values for each task are provided in Table 1.

During this round, 10 additional tasks were also nominated, resulting in 28 tasks that still required consensus ratings and a total of 64 tasks remaining under consideration. New tasks are listed at the bottom of Table 1, and a flow chart of the complete rating process is depicted in Fig. 1.

Delphi Round 2

Seventeen members of the RHG provided complete ratings. Consensus was reached for 16 of the original tasks and 8 of the additional tasks, resulting in 4 tasks for which consensus was not reached. In either Round 1 or 2, 50 tasks received ratings of ≥2 indicating "adequate" or better quality, and these tasks advanced to Round 3. Tasks that failed to achieve a rating of 2 or more in either Round 1 or 2 were dropped from consideration ($n = 14$). Average ratings and IQR values from Round 2 are presented in Table 1.

Delphi Round 3

For this round, tasks were categorized according to social cognitive domain, and tasks assessing multiple domains (e.g., OSCARS) were listed within each applicable domain. This process resulted in 23 tasks for emotion processing, 6 for social perception, 17 for mental state attribution, 4 for attributional style/bias, and 6 for empathy.

Eighteen members of the RHG provided task rankings and ratings for suitability of use. Within each domain, a top tier of tasks emerged as the most consistently selected, and all were rated as having "good" or better suitability for use. The top six ranked tasks for cross-cultural use are listed in Table 2, and the top six ranked tasks for transdiagnostic use are listed in Table 3. Average suitability ratings for these tasks are also presented in the corresponding tables. Information for the remaining emotion processing and mental state attribution tasks are provided in Supplementary Tables 3, 4.

DISCUSSION

The current study utilized an expert survey and Delphi consensus process to identify social cognitive tasks that may be appropriate for cross-cultural and/or transdiagnostic use. Within the domain of emotion processing, the ER-40 ranked as the top task for both cross-cultural and transdiagnostic use. This task shows 40 static images of individual's faces and asks participants to label the emotion shown from the following choices: happy, sad, anger, fear, no emotion. The ER-40 uses age, gender, and ethnically diverse stimuli and has minimal verbal demand, which likely contributed to its high ranking.

Far fewer tasks assessing social perception were available; however, the OSCARS emerged as the most promising for cross-cultural use and the second most promising for transdiagnostic use. The OSCARS can be used as a self- and/or informant-report assessment and asks how much difficulty someone has decoding verbal cues. While an early study indicated that informant reports showed stronger convergent and external validity than self-reports[26], a more recent, larger study suggests equivalent validity between the two modalities[27]. The interview-based format provides the option to combine across sources of information, including self- and informant-report or multiple informants, and may provide a more sensitive indicator of real-world change than performance-based tasks, thus potentially serving as a valuable coprimary measure for clinical trials[28]. Translation of this task is likely necessary for cross-cultural use; however, the general, non-performance based assessment approach may make it widely applicable and more easily adaptable to other cultures.

Table 1. Task Ratings from Delphi Rounds 1 and 2.

Task name	Number of Times Nominated for Each Use			Round 1 Quality Rating	Round 1 IQR	Round 2 Quality Rating	Round 2 IQR	Included in Round 3
	Cross-Cultural	Trans-diagnostic	Further Development					
Affective Responsiveness Task[53]			1	2.05	.50			X
Ambiguous Intentions Hostility Questionnaire (AIHQ)[54]	5	6	1	2.47	1.00			X
Bell-Lysaker Emotion Recognition Task (BLERT)[17]	6	3	1	2.86	2.00	2.44	1.00	X
Benton Face Recognition Test (BFRT)[55]		1		1.90	2.00	1.35	1.00	
Brune's Picture Sequencing Test[56]	1	1		1.89	1.00			
Combined Stories task (COST)[57]	1		1	2.44	1.00			X
Continuous Facial Sequence Task (CFST)[58]		1		2.07	.50			X
Davos Assessment of Cognitive Biases Scale (DACOBS)[59]			1	1.89	1.00			
Dynamic Interactive Social Cognition Training in Virtual Reality (DiSCoVR)[60]			1	1.47	1.00			
Dynamic Virtual Faces(DVFs)[61]	1			2.16	.00			X
Emotion in Biological Motion[12]	5	2	1	2.24	1.00			X
Ekman Faces Test (FEEST)[62]	2	3		2.67	1.00			X
Emotion Recognition Index (ERI)[63]	1	5		1.85	1.50	1.40	1.00	
Emotion Recognition Task (ERT) - CANTAB version[64,65]	3	4	1	2.43	1.00			X
Emotional Perspective Taking Task(EPTT)[53]	2	2	2	2.61	1.00			X
Emotional Stroop, e.g.,[66]		1		1.25	.50			
Empathic Accuracy Task[67]		1	1	2.26	1.00			X
Empathy Test[68]		1		2.33	1.00			X
Face Emotion Discrimination Test (FEDT)[69]		1		2.28	.00			X
Facial Emotion Identification Task (FEIT)[69]		1		2.16	.75			X
False Belief Theory of Mind Stories[70]		1		2.19	.50			X
Faux Pas Test[71]	2	2	1	2.43	1.00			X
Florida Affect Battery (FAB)[72]		1		2.22	1.00			X
Geneva Emotion Recognition Test (GERT)	2	3	2	2.39	1.00			X
Half Profile of Nonverbal Sensitivity (Half-PONS)[73]			1	1.63	1.00			
Happé's Cartoon Task[74]		3		2.05	2.00	1.82	1.00	X
Heider and Simmel Triangle Task[75]	1			1.71	1.25	1.28	.25	
High-Risk Social Challenge (HiSoC)[76]			1	1.83	2.00	1.39	1.00	
Hinting Task[15,77]	7	9	2	2.65	1.00			X
Intentionality Bias Task (IBT)[78]			2	2.13	1.50	1.63	1.00	X
Internal, Personal and Situational Attributions Questionnaire (IPSAQ)[79]	2	2		2.22	1.25	1.75	1.00	X
Interpersonal Reactivity Index (IRI)[80]	1	1		2.14	1.50	1.94	1.25	X
Levels of Emotional Awareness Scale (LEAS)[81]	1	1		2.12	1.00			X

Table 1 continued

Task name	Number of Times Nominated for Each Use			Round 1 Quality Rating	Round 1 IQR	Round 2 Quality Rating	Round 2 IQR	Included in Round 3
	Cross-Cultural	Trans-diagnostic	Further Development					
Lecture Intentionnelle en Situation – Versailles (LIS-V)[82]	1	1		2.33	n/a			
Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT)[83]	6	2	1	2.32	1.00			X
Metacognition Assessment Scale-Abbreviated (MAS-A)[84]			1	2.00	1.75	1.35	.50	X
Metaphor and irony stories[85]		1		1.79	1.00			
MINI Social Cognition & Emotional Assessment (MINI-SEA)[86]		1	1	2.26	1.00			X
Mini-Profile Of Nonverbal Sensitivity (Mini-PONS)[87]	1	1		2.35	1.00			X
Modified Reversal Dictator Game[88]	1	1		1.38	1.00			
Movie for the Assessment of Social Cognition (MASC)[89]	1	2	2	2.58	1.00			X
Multi-faceted Empathy Test (MET)[90]		2		2.26	1.00			X
Multilingual Communicative Interaction Database[91]		1		1.78	1.00			
Multimodal Emotion Recognition Test (MERT)[92]	1	1		2.22	.25			X
NIH Toolbox Adult Social Relationship Scales[93]		1		1.39	2.00	1.06	.00	
Observable Social Cognition Task (OSCARS)[94]	1	1		2.11	1.25	1.88	.50	X
Penn Emotion Differentiation Test (PEDT)[95]			1	1.83	1.25	1.88	1.00	
Penn Emotion Recognition Test (ER-40)[16]	5	7	1	2.90	.75			X
Perceptual animacy[96]	1	1		1.38	1.00			
Reading the Mind in the Eyes Test (RMET)[97]	7	9	1	2.15	2.00	1.50	.00	X
Reading the Mind in the Voice (RMVT)[98]		1		2.28	1.00			X
Self Referential Memory, Ownership[99]		1		1.24	0.0			
Self Referential Memory, Self-reference Effect (SRE)[100]	1			1.31	.75			
Short Story Task (SST)[101]			1	1.53	1.00			
Social and Occupational Functioning Assessment Scale (SOFAS)[102]		1		1.12	0.0			
Social Attribution Task – Multiple Choice (SAT-MC)[103]	1	2	1	2.25	1.00			X
Social Communication Disorders Checklist[104]	1			1.59	1.00			
Social Cue Recognition Test (SCRT)[105]	2	2		1.94	.5			
Social Functioning Questionnaire (SFQ)[106]		1		1.18	0.0			
Social Functioning Scale (SFS)[107]	2	1		1.18	0.0			
Social Instrumental Learning[108]	1			1.45	1.00			
Social Judgment Task[109]		1		1.81	0.75			
Social Learning Task[110,111]			1	1.46	1.00			
Social Norms Questionnaire[112]		1		2.00	0.0			X

Table 1 continued

Task name	Number of Times Nominated for Each Use			Round 1 Quality Rating	Round 1 IQR	Round 2 Quality Rating	Round 2 IQR	Included in Round 3
	Cross-Cultural	Trans-diagnostic	Further Development					
Social Perception and Interaction Database (SoPID)[113]			1	2.07	1.25	1.87	1.00	X
Social Skills Performance Assessment (SSPA)[114]			2	1.83	2.00	1.24	1.00	
Strange Stories Short Film Task (SSFT)[115]	1			2.27	1.00			X
Strange Stories Task[116]	2	4		2.33	1.00			X
Test de Reconnaissance des Emotions Faciales (TREF)[117]	1	2		2.11	1.00			X
The Awareness of Social Inference Test (TASIT)[118]	4	6	1	2.80	0.75			X
The Interview Task[119]			1	1.71	1.00			
The Penn Emotion Acuity Test (PEAT)[120]		1	1	2.07	.25			X
The Personality Pairs Task (PPT)[119]			1	1.25	.75			
The Self-other Differentiation Paradigm using Perceptual Matching[121]	1	1		1.27	1.00			
The Social Attribution Task[121]		1		2.24	2.00	1.94	.50	X
The Social Knowledge Test (SKT)[121]		1		2.5	1.00			X
Theory of Mind by Humor Comprehension and Appreciation Test ToM-HCAT)[122]	2			1.80	1.00			
Theory of Mind-15 (TOM-15)[123]		1		1.77	1.00			
Triangles Task/Frith-Happé Animations Test[124]	1	1	2	2.11	2.00	2.00	2.00	X
Virtual Assessment of Mentalising Ability (VAMA)[125]			1	2.65	1.00			X
Whodunit[126]	1			1.60	1.00			
Tasks Introduced at Round 2 by RHG Members								
Animated Shape Videos[127]						1.75	1.75	
Brief Assessment of Social Skills (BASS)[128]						1.73	1.0	
Cartoon Intentions Task[129]						1.92	1.75	
Contextual Assessment of Social Skills (CASS)[130]						1.40	1.0	
Edinburgh Social Cognition Test (ESCoT)[131]						2.44	1.00	X
Emotion Recognition Test[64,132]						2.38	1.00	X
Empathy Quotient (EQ)[133]						1.93	1.0	
Interpersonal Perception Task (IPT)[134]						1.57	1.0	
La Trobe Communication Questionnaire[135]						1.40	1.0	
Questionnaire of Cognitive and Affective Empathy (QCAE)[136]						2.20	1.0	X

Quality ratings were made using a 5-point scale, ranging from 0–4: 0 = not possible to rate or insufficient information, 1 = inadequate, 2 = adequate, 3 = good, 4 = excellent. Designations of "n/a" indicate tasks for which IQR could not be calculated due to a lack of non-zero scores.

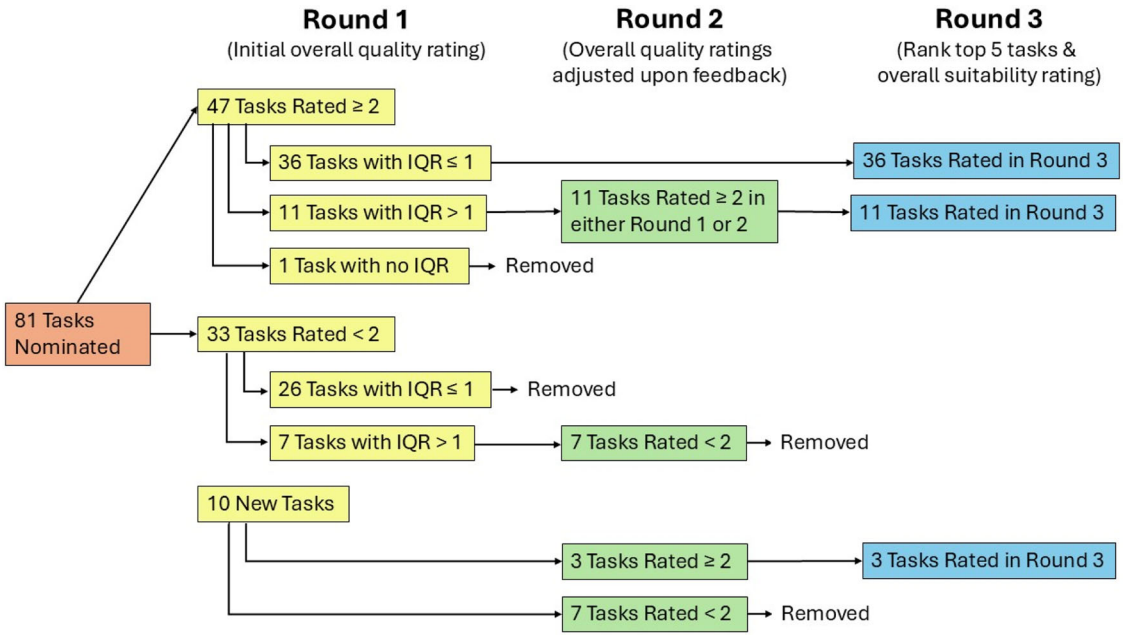

Fig. 1 Flow Chart of the Delphi Rating Process.

Several tasks evaluating mental state attribution were nominated, with both the SAT-MC and TASIT being ranked highest for cross-cultural use and the TASIT ranked highest for transdiagnostic use. The SAT-MC shows short animations of geometrical shapes enacting a social drama, and participants answer multiple choice questions about the actions and intentions of the shapes. The minimal verbal and memory demands of this task make it highly feasible for both uses, and it is noteworthy that it has one of the highest suitability ratings of all tasks for cross-cultural use. The TASIT involves short videos of everyday social interactions involving two or more people and evaluates the ability to detect lies and sarcasm in these interactions. The videos include English-speaking actors with Australian accents, which may somewhat limit its cross-cultural applicability; however, this should not impede transdiagnostic use, and an abbreviated format and some translations are available (e.g.,[29,30]).

For attributional style/bias, the AIHQ received the highest rankings for both uses. This task presents hypothetical, negative situations with ambiguous causes and asks participants why the scenario occurred. Participants also rate whether the other person acted intentionally, how angry it would make them feel, and how much they would blame the other person. Finally, they are asked how they would respond to the situation. The high ranking of this task for cross-cultural use is somewhat surprising given the high dependence on verbal language; however, the situations used are likely to be considered negative and ambiguous across most cultures, which may explain its high ranking. Further, only four tasks were nominated in this category, and the modest suitability ranking for cross-cultural use, as compared to transdiagnostic use, likely reflects some hesitancy regarding the task.

Finally, within the domain of empathy, the QCAE was highly ranked for both uses. This self-report scale asks participants to rate how much statements pertaining to cognitive and affective empathy pertain to them. Importantly, despite being ranked highly, suitability ratings for this task were modest, again suggesting that while it is a "good" task, additional development would be needed for optimal use.

Overall, there was considerable overlap in the top ranked tasks within each domain for cross-cultural and transdiagnostic use. This may be due to the somewhat limited number of generally "good" tasks from which to choose or to the possibility that the same characteristics make a task appealing for either transdiagnostic or cross-cultural use (e.g., sound basic psychometrics like validity and reliability). Nevertheless, there were a few instances, for example the Hinting Task, in which a task was ranked highly for one use and not another. Thus, we suggest care when selecting tasks for each potential use.

Additionally, while these measures, and others, exhibited potential and were rated relatively favorably, no task emerged as ideally suited for the intended use. Of tasks considered in Round 3, suitability ratings were primarily in the "good" range, with no tasks scoring in the highest range, and from Rounds 1 and 2, almost half of the tasks received ratings in the inadequate range. This may be because very few social cognitive tasks have been specifically designed with cross-cultural considerations in mind. As excellently reviewed by Bourdage and colleagues[31] who focused on tasks appropriate for Global South communities, there is a considerable lack of multicultural assessment tools, and attempts to modify existing tasks have been primarily of low quality. Further, many existing modifications have been conducted with a single culture in mind (i.e., translating a measure to be used in a specific location) rather than with the goal of making a truly cross-cultural tool that could be readily used across several cultures.

It is also noteworthy that several tasks scoring below adequate in Round 1 or 2 still ranked highly in Round 3. This phenomenon was most common for domains with fewer nominated tasks and may reflect a lack of viable options. Alternatively, this may also indicate a willingness to use less than ideal tasks, either out of necessity as just mentioned, or because of easy availability and increased familiarity. The RMET is a good example of this. It is among the most widely used and translated social cognitive tasks[32] and received the overall highest rating of suitability for cross-cultural use. However, translation does not guarantee cultural appropriateness, and recent years have seen a sharp increase in concern regarding its validity[11] and growing evidence of cultural bias[33–35]. As such, we encourage critical evaluation of existing measures and thoughtful consideration of measurement when designing studies.

We also encourage continued development of novel tasks that are proactively designed to be used cross-culturally. Sensitive cultural adaptation of any cognitive measure is difficult and

Table 2. Top Ranked Tasks for Cross-Cultural Use from Delphi Round 3.

Task	Number of Selections	Weighted Mean Rank	Suitability Rating
Emotion Processing			
Penn Emotion Recognition Test (ER-40)	14	2.89 (1.38)	6.23 (1.42)
Emotion in Biological Motion	12	1.83 (1.60)	6.82 (1.47)
Bell-Lysaker Emotion Recognition Task (BLERT)	9	1.50 (1.58)	5.11 (2.03)
Reading the Mind in the Eyes Test (RMET)[a]	4	0.89 (1.16)	7.33 (0.58)
Dynamic Virtual Faces (DVFs)	6	0.89 (1.21)	6.20 (1.92)
Levels of Emotional Awareness	4	0.89 (1.41)	5.33 (2.52)
Social Perception			
Observable Social Cognition Task (OSCARS)[a]	13	2.72 (1.48)	5.15 (1.99)
Social Perception and Interaction Database (SoPID)[a]	10	2.33 (0.79)	6.00 (0.94)
Mini-Profile of Nonverbal Sensitivity (Mini-PONS)	11	2.22 (1.69)	5.91 (2.12)
Edinburgh Social Cognition Test (ESCoT)	11	1.89 (1.22)	5.27 (1.56)
The Social Knowledge Test (SKT)	8	1.33 (1.20)	5.13 (1.73)
Social Norms Questionnaire	5	0.72 (1.14)	4.40 (1.67)
Mental State Attribution			
Social Attribution Task – Multiple Choice (SAT-MC)	7	1.56 (1.16)	7.29 (1.25)
The Awareness of Social Inference Test (TASIT)	10	1.50 (1.25)	4.80 (1.93)
Triangles Task/ Frith-Happé Animations Test[b]	6	1.44 (1.21)	6.67 (0.82)
False Belief Theory of Mind Stories	6	1.11 (0.82)	6.00 (1.55)
Observable Social Cognition Task (OSCARS)[a]	4	1.00 (1.00)	5.00 (2.16)
Hinting Task	6	0.94 (1.60)	4.00 (1.67)
Attributional Style			
Ambiguous Intentions Hostility Questionnaire (AIHQ)	15	3.56 (0.88)	5.13 (2.10)
Internal, Personal, and Situational Attributions Questionnaire (IPSAQ)[a]	11	2.67 (0.81)	5.73 (2.24)
Intentionality Bias Task (IBT)[a]	13	2.33 (1.09)	5.08 (2.22)
Observable Social Cognition Task (OSCARS)	9	1.61 (0.97)	5.00 (2.06)
Empathy			
Empathic Accuracy Task	11	2.33 (1.40)	4.91 (2.02)
Questionnaire of Cognitive and Affective Empathy (QCAE)	12	2.22 (1.56)	5.33 (1.78)
Interpersonal Reactivity Index (IRI)[a]	9	1.83 (1.23)	4.67 (1.67)
Multi-faceted Empathy Test (MET)	8	1.61 (1.30)	5.63 (1.50)
Empathy Test	7	1.33 (1.40)	5.71 (1.60)
Emotional Perspective Tasking Task (EPTT)	6	0.89 (1.37)	6.50 (1.64)

The top six ranked tasks from each social cognitive domain are listed. As the domain of Attributional Style included only four tasks, all are listed. Suitability ratings are on a scale from 1–9 (1 = poor, 3 = fair, 5 = good, 7 = very good, 9 = superb).
[a]Denotes tasks for which consensus was not reached on Round 1 and mean quality ratings dropped below 2 at Round 2.
[b]Denotes tasks for which consensus was not reached in either Round 1 or 2.

requires due process, in which researchers need to pay close attention to concept equivalence across cultures. For example, results from previous studies on functional capacity and interview-based measures of cognitive impairment showed that these adaptations usually require substantial edits and that the content of the tasks needed to be adapted[36]. Thus, applying these same principles to social cognitive tasks, that often use complex and nuanced social stimuli, is an even greater challenge but one that should not be neglected (see refs. [37,38] for suggested guidelines). It also bears noting that work of this type is resource intensive, and that limited funding has previously been cited as a primary barrier[31]. We therefore encourage investment on the part of funding agencies and foundations.

In addition to the general challenge of developing and validating culturally sensitive tasks, another important element will be to ensure measurement invariance across adaptations. Most social cognitive measure validation studies analyze only basic psychometric properties, such as internal consistency or concurrent and criterion validity. Relatively few studies have applied advanced statistical modeling to estimate measurement properties (see ref. [39] for an exception). It may therefore be helpful to use archival data to test whether some widely used measures are invariant across cultures and to consider measurement invariance in future validation efforts.

Moving forward, it may also be fruitful to emphasize continued development of paradigms rather than specific tasks. For example, the basic structure of the TASIT or the Hinting task, in which participants must interpret interactions between characters, is quintessential to social cognition and could be retained while the specific stimuli or scoring criteria could be adapted to apply more broadly to multiple cultures. A recently developed multiracial version of the RMET also provides a good example of this idea. Here, Kim and colleagues retained the structure of the RMET but updated the stimuli and answer choices to produce a more inclusive version of the task that may mitigate some of the bias introduced by only using white, European faces. Future work may

Table 3. Top Ranked Tasks for Transdiagnostic Use from Delphi Round 3.

Task	Number of Selections	Weighted Mean Rank	Suitability Rating
Emotion Processing			
Penn Emotion Recognition Test (ER-40)	11	2.17 (1.64)	7.09 (1.14)
Bell-Lysaker Emotion Recognition Task (BLERT)	8	1.50 (1.51)	6.50 (1.60)
The Awareness of Social Inference Test (TASIT)	5	0.83 (1.41)	6.80 (1.10)
Emotion in Biological Motion	7	0.78 (1.53)	6.00 (1.00)
Geneva Emotion Recognition Test (GERT)	3	0.67 (0.00)	7.33 (0.58)
Emotion Recognition Task (ERT)	4	0.67 (1.41)	6.50 (1.73)
Social Perception			
Mini-Profile of Nonverbal Sensitivity (Mini-PONS)	10	2.28 (1.29)	6.00 (1.33)
Observable Social Cognition Task (OSCARS)[a]	9	2.11 (1.09)	6.22 (0.97)
Edinburgh Social Cognition Test (ESCoT)	9	1.94 (1.05)	5.67 (1.12)
Social Perception and Interaction Database (SoPID)[a]	9	1.67 (1.00)	5.67(1.41)
Social Norms Questionnaire	7	0.94 (1.40)	5.00 (1.73)
The Social Knowledge Test (SKT)	5	0.72 (1.34)	5.20 (1.48)
Mental State Attribution			
The Awareness of Social Inference Test (TASIT)	12	2.44 (1.37)	6.00 (1.71)
Hinting Task	8	1.67 (1.28)	6.50 (1.31)
Reading the Mind in the Eyes Task (RMET)[a]	4	0.94 (0.96)	6.75 (0.96)
False Belief Theory of Mind Stories	4	0.89 (1.41)	6.25 (0.96)
Triangles Task/ Frith-Happé Animations Test[b]	5	0.83 (1.23)	6.20 (0.84)
Social Attribution Task – Multiple Choice (SAT-MC)	6	0.78 (1.21)	7.17 (0.75)
Attributional Style			
Ambiguous Intentions Hostility Questionnaire (AIHQ)	14	3.44 (0.76)	6.14 (1.23)
Intentionality Bias Task (IBT)[a]	10	2.17 (0.88)	5.80 (1.23)
Observable Social Cognition Task (OSCARS)[a]	8	1.56 (1.41)	5.13 (1.46)
Internal, Personal and Situational Attributions Questionnaire (IPSAQ)[a]	7	1.50 (0.90)	5.71 (1.60)
Empathy			
Questionnaire of Cognitive and Affective Empathy (QCAE)	10	2.00 (1.43)	5.70 (1.34)
Multi-faceted Empathy Test (MET)	8	1.83 (0.99)	6.00 (1.20)
Interpersonal Reactivity Index (IRI)[a]	8	1.72 (1.13)	5.63 (1.77)
Empathic Accuracy Task	8	1.67 (1.58)	6.25 (1.28)
Empathy Test	7	1.39 (0.98)	6.29 (1.70)
Emotional Perspective Taking Task (EPTT)	5	0.56 (0.71)	6.20 (1.30)

The top six ranked tasks from each social cognitive domain are listed. As the domain of Attributional Style included only four tasks, all are listed. Suitability ratings are on a scale from 1–9 (1 = poor, 3 = fair, 5 = good, 7 = very good, 9 = superb).
[a]Denotes tasks for which consensus was not reached on Round 1 and mean quality ratings dropped below 2 at Round 2.
[b]Denotes tasks for which consensus was not reached in either Round 1 or 2.

also benefit by adapting novel paradigms from social and experimental psychology. Social cognitive research in the general population has been moving from static paradigms based on the perception and interpretation of social stimuli to more active, dynamic paradigms[40] such as dyadic interactions (e.g., [41]) that may allow for the assessment of social cognitive processes, like emotion recognition, in real time and in more naturalistic ways. These paradigms have substantial technical and analytical demands but might bring novel insights about the nature of social cognitive difficulties observed across clinical conditions.

As our study represents a consensus-based effort to identify cross-cultural and transdiagnostic social cognitive assessments, some limitations require consideration. First, our RHG included individuals with varying expertise including early career individuals, industry representatives, and individuals with lived experience. While this significantly increased the diversity of perspectives in the RHG, these forms of expertise were not equally represented within the RHG and not all members were as familiar with the breadth of existing social cognitive assessments, which could have skewed our results toward tasks that are more widely used and therefore more familiar. Services users were also underrepresented in the expert survey and RHG, which prevented us from broadly capturing their viewpoints. Accessibility of social cognitive tasks, particularly for service users, should be prioritized in future consensus-based work. Second, just over 25% of the experts invited to the initial survey responded. These experts represented a wide range of countries, but the majority were from North America and Europe and expertise in schizophrenia was most common. Broader representation may have yielded a different set of tasks for consideration by the RHG. Likewise, our RHG lacked representation from Spanish- and Arabic-speaking countries as well as African countries, which account for significant portions of the global population. Future work should strive for broader representation, and when possible, consider multiple languages, including indigenous languages. Third, the data generated here may be viewed as being primarily

schizophrenia-focused. Expertise in schizophrenia was disproportionally represented compared to other specific disorders (e.g., autism); however, over half of the experts who responded to the initial survey (53%) identified as having primary expertise in areas other than schizophrenia. Similarly, of the 22 academic and clinical RHG members, 40% worked in populations other than schizophrenia. Thus, we believe this work is still widely applicable and relevant to fields outside psychosis but acknowledge that the results of the Delphi Process may reflect perspectives more heavily weighted by schizophrenia spectrum disorders. Finally, we focused on assessments that are appropriate for adults. Additional work will be needed to identify the most suitable assessments for children and adolescents.

CONCLUSION

Harmonization of social cognitive research has been significantly limited by a lack of cross-culturally and transdiagnostically valid assessment tools. Results of the global expert survey and subsequent consensus process reported here underscore the relative dearth of suitable measures but do identify a small selection of assessments that appear to be appropriate for current use. These tasks should be explicitly evaluated in cross-cultural and transdiagnostic studies. Additional efforts should also be made to continue adapting existing measures and to develop novel measures that can be used in each of these capacities. Notwithstanding these issues, the tasks identified here represent multiple social cognitive domains beyond the traditionally considered core processes of mental state attribution and emotion processing. In addition to guiding future research, identification of these tasks may provide a much needed springboard for increasing assessment of social cognition in clinical practice, where it remains under-utilized[42].

METHODS

This study has been approved by the Institutional Review Board (IRB) of The University of Texas at Dallas (IRB-23-177).

Expert survey for task nomination

Survey content was drafted by the conveners of the RHG (AP, MH, and TZ) and then further refined and augmented with input from RHG-members. The final survey consisted of two primary parts: (1) Background information; and (2) Task nominations for future use. For part 2, respondents were prompted to nominate social cognition tasks that they believed were suitable for: (a) International and/or cross-cultural studies; and/or (b) Transdiagnostic studies. In a separate question, this section also allowed respondents to nominate tasks that may not currently be suitable for either of the intended uses but that may show promise with continued adaptation and further development (see Supplementary Table 1). A final portion of the survey queried current use of social cognitive assessments and perceived measurement-related barriers within social cognitive research, the results of which will be reported elsewhere. A copy of the survey is available in Supplementary Materials.

In parallel, the definition of "expert" was formulated first by the conveners and then edited via group discussion with the RHG.

The term "expert" for academic researchers was defined as follows:

- research experience (either academic or industry) in the field of psychology, psychiatry, social neuroscience, or an allied discipline for at least 4 years and currently active in one of those fields, AND at least 2 peer-reviewed publications on social cognition, of which at least 1 is as first, second, or senior author, and of which at least 1 has been published in the last 5 years.
- for researchers from non-English speaking countries, articles written in languages other than English qualified if they were published in a peer-reviewed journal.

Non-academic expertise (e.g., clinicians, students, industry team members, or service users) was defined as:

- hands-on experience with, or intricate knowledge of, at least 2 social cognition paradigms.

The online survey was implemented in REDCap[43] (hosted at The University of Texas at Dallas) for data collection and management, and subsequently distributed via emailed invitations through the RHG-members' familiarity with experts, as well as supplementary literature searches by graduate students. All RHG members qualified as experts according to the definitions above and were therefore encouraged to complete the survey as well. The estimated duration time for filling out the survey was 10 min. Data for the expert survey were collected between late October 2022—early January 2023.

In total, 381 experts were invited to participate anonymously via an emailed survey-link and asked to share the link with other potential experts meeting the criteria. Ninety-eight experts residing in 25 countries across five continents responded to the invitation and provided digital consent. 50% of respondents identified as men, and age was normally distributed (range: 20–70+; mode = 40–44 y). Most experts (70%) identified professionally as professor/lecturer (any level), 36% as researcher, 18% as a clinician, and 6% as service user/other. Schizophrenia/psychosis or high-risk for psychosis was checked by 46% of experts as their main study population of interest, followed by general population (13%), and autism, bipolar disorder and neurodegenerative disease (all 7%). For additional detail on respondent characteristics, see Supplementary Table 2.

The Delphi methodology

The Delphi method is a structured communication technique used to achieve consensus among a group of experts by soliciting their opinions through an iterative series of questionnaires and providing them with controlled feedback[44]. The method is based on the concept of collective wisdom, which assumes that the combined opinion of multiple people is closer to the truth than a single individual's perspective[45]. To obtain consensus, group members complete a series of online, anonymous questionnaires from which results are aggregated in a systematic manner and then presented back to the larger group. This process of responding and receiving/incorporating feedback (i.e., a "round") is repeated until group consensus is reached. New information can be introduced at any point or during any round, and ensuring anonymity is thought to reduce undue influence from any (especially more influential) group members and the pressure to conform. This process has previously been used in psychological assessment research (e.g., [46]), including studies focused on cross-cultural assessment[47] and one study that sought to identify social cognitive assessments for use in Japanese individuals with schizophrenia[48]. The Delphi portion of this study consisted of 3 consecutive rounds of online questionnaires, further outlined below.

Delphi expert panel selection

Panels with 10 to 50 members are recommended for Delphi studies[49]. As such, the original RHG membership, comprised of 13 international members of the Schizophrenia International Research Society, as well as two service users and two industry representatives (note: one of the industry partners was unable to continue their participation due to time constraints leaving only one industry representative in the final RHG.), was expanded to include experts from clinical specializations other than

Table 4. Characteristics of Research Harmonization Group Members.

	Mean	SD
Age	42.62	9.97
	N	%
Country of Residence		
Australia	2	7.69
Canada	2	7.69
China	1	3.85
France	1	3.85
India	1	3.85
Italy	1	3.85
Netherlands	4	15.38
Republic of Korea	2	7.69
Slovak Republic	1	3.85
United Kingdon	3	11.54
United States	8	30.77
Gender Identity		
Man	10	38.46
Woman	15	57.69
Gender Variant/Non-Conforming	1	3.85
Racial/Ethnic Background		
Asian	4	15.38
Black	1	3.85
White/Caucasian	15	57.69
Latinx	1	3.85
Field left blank	5	19.23
Organizational Affiliation(s)[a]		
University	18	69.23
University Medical Center	8	30.77
Mental health care organization (non-academic)	1	3.85
Research Institute (non-academic)	1	3.85
Commercial Industry	1	3.85
Non-profit	1	3.85
Other	3	11.54
Position(s)[a]		
Professor/Lecturer (any level)	16	61.54
Researcher	11	42.31
Clinician	3	11.54
Service User	1	3.85
Other	3	11.54
Population(s) of Study		
Schizophrenia/psychosis	18	72.0
High-risk for psychosis	5	20.0
Depression	4	16.0
Bipolar disorder	5	20.0
Autism and neurodevelopmental disorders	10	40.0
Neurodegenerative diseases	6	24.0
Traumatic brain injuries	4	16.0
Cerebrovascular accident	3	12.0
General population	8	32.0
Other	4	16.0

[a]Respondents could select multiple options, and thus percentages may add up to >100.

schizophrenia/psychosis based on RHG member recommendations. Clinical fields of expertise included: autism, bipolar disorder, neurodegenerative disorders, pediatrics and acquired brain injury. The final group of 26 experts consisted of residents from: USA (8), the Netherlands (4), United Kingdom (3), Australia (2), Canada (2), South Korea (2), China (1), France (1), Italy (1) Slovakia (1), and India (1). Members of the RHG were diverse in gender, race, ethnicity, and career stage (see Table 4).

Delphi procedure and data analysis

After the expert survey was completed, task nominations were collated and used to build a database containing a description of each task as well as a summary of the currently available psychometric data for that task. To the extent possible, information pertaining to reliability (e.g., test-retest reliability, internal consistency), validity (e.g., convergent and discriminant [including consideration of overlap with cognitive performance assessments], criterion, distribution of scores, sensitivity to group differences), practicality of administration, and tolerability (e.g., pleasantness or unpleasantness of completing the task) was included. This database was distributed to RHG members at each round to aid their evaluations.

For each nominated task the overarching goal was to reach a consensus score and establish its utility for cross-cultural and transdiagnostic research. At the beginning of each round, relevant information (i.e., the database) and/or feedback from the previous round was provided to allow the experts to modify their opinions with the aim of reaching group consensus. All tasks were rated on a 5-point scale, ranging from 0–4: 0=not possible to rate or insufficient information, 1 = inadequate, 2 = adequate, 3 = good, 4 = excellent. Our operationalization of consensus was an interquartile range (IQR) ≤ 1. For a four- to five-point Likert scale, an IQR of 1 or less is considered a high level of consensus[50,51]. Ratings were conducted anonymously while unique user IDs of participants were collected to monitor variation in expert responsiveness across all three rounds. As noted above, off-line and new information could be suggested at any point and/or during any round. After each round, the tasks for which consensus was not achieved moved into the subsequent round for re-rating. Data collection for round 1 started in May 2023 and round 3 was finalized in January 2024.

Round 1. Eighty-one different social cognition tasks were nominated in the expert survey and included in Round 1. The objective of the first round was to rate the overall quality of each nominated task based on the three considerations listed below:

1. Does the task really measure social cognition, and does it tap into at least one social cognitive domain?
2. Is the task generally a "good" task given what you know about its psychometric properties?
3. Is this task relatively easy to administer and take or it is too onerous (e.g., burdensome) to be useful?

Individual scores of 0 ("not possible to rate or insufficient information") were excluded before calculating the mean and IQR for each task. To maintain high responsivity, it was deemed necessary to limit the number of included tasks for subsequent rating rounds. Thus, tasks receiving a consensus average score of less than 2, indicating a rating of less than "adequate" were omitted from further consideration.

Round 2. Eighteen tasks that did not reach consensus in Round 1 were rated once more in a similar manner as for Round 1. These tasks were presented along with information on the average panel rating, each expert's own previous rating (observable with the

unique, anonymous user ID for each participant), and an overview of comments that were offered by experts in support of their ratings in the first round. In addition, 10 newly suggested tasks from round 1 were included to collect an initial rating. As only two old tasks and two newly suggested tasks did not reach a consensus score after Round 2, it was determined ad hoc not to request our RHG members for an additional rating for these tasks.

Round 3. After the consensus procedure, the goal of the third round was to identify the best tasks for cross-cultural and transdiagnostic use. Before this final round, tasks were categorized according to social cognitive domain by the conveners, after independent classification and a consensus meeting. The domains adhered to the four domains distilled from the SCOPE study[52]: emotion processing, social perception, theory of mind/mental state attribution, and attributional style/bias. A fifth domain was added specifically for empathy tasks, which were not included in the SCOPE study. Tasks assessing multiple domains were included in each applicable domain.

In Round 3, RHG members were then asked to identify and rank their top 5 tasks within each social cognitive domain and provide a rating of the overall suitability of that task for the intended use (1 = poor, 3 = fair, 5 = good, 7 = very good, 9 = superb). This was first requested for cross-cultural use and then for transdiagnostic use. To identify top tasks, mean rank was reverse coded so that higher scores indicated better tasks and then weighted by the number of times a task was selected within its social cognitive domain (e.g., (mean rank × number of times the task was selected)/18 raters). Thus, a task with a mean rank of 3.0 based on 10 rankings would have a weighted rank of 1.67 and would be preferred to a task with mean rank of 5.0 (the highest possible) based on just two rankings, which would have a weighted rank of 0.55.

DATA AVAILABILITY

The datasets generated and analyzed for the current study are available from the corresponding author upon reasonable request at amy.pinkham@utdallas.edu.

Received: 1 October 2024; Accepted: 18 November 2024;
Published online: 21 February 2025

REFERENCES

1. Hajdúk, M., Harvey, P. D., Penn, D. L. & Pinkham, A. E. Social cognitive impairments in individuals with schizophrenia vary in severity. *J. Psychiatr. Res.* **104**, 65–71 (2018).
2. Rocca, P. et al. Social cognition in people with schizophrenia: a cluster-analytic approach. *Psychol. Med.* **46**, 2717–2729 (2016).
3. Vaskinn, A., Sundet, K. & Haatveit, B. Social cognitive heterogeneity in schizophrenia: a cluster analysis. *Schizophr. Res. Cogn.* **30**, 100264 (2022).
4. Lewandowski, K. E., Pinkham, A. E. & Van Rheenen, T. E. Social cognition across the schizophrenia–bipolar disorder spectrum. *Nat. Rev. Psychol.* **3**, 91–107 (2024).
5. Green, M. F., Horan, W. P. & Lee, J. Social cognition in schizophrenia. *Nat. Rev. Neurosci.* **16**, 620–631 (2015).
6. Halverson, T. F. et al. Pathways to functional outcomes in schizophrenia spectrum disorders: Meta-analysis of social cognitive and neurocognitive predictors. *Neurosci. Biobehav. Rev.* **105**, 212–219 (2019).
7. Maj, M. et al. The clinical characterization of the patient with primary psychosis aimed at personalization of management. *World Psychiatry* **20**, 4–33 (2021).
8. Vita, A. et al. European Psychiatric Association guidance on assessment of cognitive impairment in schizophrenia. *Eur. Psychiatry* **65**, e58 (2022).
9. Etchepare, A. & Prouteau, A. Toward a two-dimensional model of social cognition in clinical neuropsychology: a systematic review of factor structure studies. *J. Int. Neuropsychol. Soc.* **24**, 391–404 (2018).
10. Higgins, W. C., Kaplan, D. M., Deschrijver, E. & Ross, R. M. Construct validity evidence reporting practices for the reading the mind in the eyes test: a systematic scoping review. *Clin. Psychol. Rev.* **108**, 102378 (2024).
11. Higgins, W. C., Ross, R. M., Langdon, R. & Polito, V. The "Reading the Mind in the Eyes" test shows poor psychometric properties in a large, demographically representative US sample. *Assessment* **30**, 1777–1789 (2023).
12. Kern, R. S. et al. Adapting social neuroscience measures for schizophrenia clinical trials, part 2: trolling the depths of psychometric properties. *Schizophr. Bull.* **39**, 1201–1210, https://doi.org/10.1093/schbul/sbt127 (2013).
13. Pinkham, A. E., Harvey, P. D. & Penn, D. L. Social cognition psychometric evaluation: results of the final validation study. *Schizophr. Bull.* **44**, 737–748 (2018).
14. Pinkham, A. E., Penn, D. L., Green, M. F. & Harvey, P. D. Social cognition psychometric evaluation: results of the initial psychometric study. *Schizophr. Bull.* **42**, 494–504 (2016).
15. Corcoran, R., Mercer, G. & Frith, C. D. Schizophrenia, symptomatology and social inference: investigating "theory of mind" in people with schizophrenia. *Schizophr. Res.* **17**, 5–13 (1995).
16. Kohler, C. G. et al. Facial emotion recognition in schizophrenia: intensity effects and error pattern. *Am. J. Psychiatry* **160**, 1768–1774 (2003).
17. Bryson, G., Bell, M. & Lysaker, P. Affect recognition in schizophrenia: a function of global impairment or a specific cognitive deficit. *Psychiatry Res.* **71**, 105–113 (1997).
18. Lim, K., Lee, S.-A., Pinkham, A. E., Lam, M. & Lee, J. Evaluation of social cognitive measures in an Asian schizophrenia sample. *Schizophr. Res. Cogn.* **20**, 100169 (2020).
19. Pinkham, A. E. et al. The other-race effect in face processing among African American and Caucasian individuals with schizophrenia. *Am. J. Psychiatry* **165**, 639–645 (2008).
20. Koelkebeck, K., Uwatoko, T., Tanaka, J. & Kret, M. E. How culture shapes social cognition deficits in mental disorders: a review. *Soc. Neurosci.* **12**, 102–112 (2017).
21. Hajdúk, M., Achim, A. M., Brunet–Gouet, E., Mehta, U. M. & Pinkham, A. E. How to move forward in social cognition research? Put it into an international perspective. *Schizophr. Res.* **215**, 463 (2020).
22. Mehta, U. M., Thirthalli, J., Gangadhar, B. N. & Keshavan, M. S. Need for culture specific tools to assess social cognition in schizophrenia. *Schizophr. Res.* **133**, 255–256 (2011).
23. Cotter, J. et al. Social cognitive dysfunction as a clinical marker: a systematic review of meta-analyses across 30 clinical conditions. *Neurosci. Biobehav. Rev.* **84**, 92–99 (2018).
24. Henry, J. D., Von Hippel, W., Molenberghs, P., Lee, T. & Sachdev, P. S. Clinical assessment of social cognitive function in neurological disorders. *Nat. Rev. Neurol.* **12**, 28–39 (2016).
25. Van den Stock, J. et al. Current potential for clinical optimization of social cognition assessment for frontotemporal dementia and primary psychiatric disorders. *Neuropsychol. Rev.* **33**, 544–550 (2023).
26. Silberstein, J. M., Pinkham, A. E., Penn, D. L. & Harvey, P. D. 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 (2018).
27. Halverson, T. F. et al. Psychometric properties of the Observable Social Cognition Rating Scale (OSCARS): self-report and informant-rated social cognitive abilities in schizophrenia. *Psychiatry Res.* **286**, 112891 (2020).
28. Keefe, R. S., Poe, M., Walker, T. M., Kang, J. W. & Harvey, P. D. The Schizophrenia Cognition Rating Scale: an interview-based assessment and its relationship to cognition, real-world functioning, and functional capacity. *Am. J. Psychiatry* **163**, 426–432 (2006).
29. Jarsch, M. et al. Development of the Basel Version of the Awareness of Social Inference Test–Theory of Mind (BASIT-ToM) in healthy adults. *J. Neuropsychol.* **17**, 125–145 (2023).
30. Westerhof-Evers, H. J., Visser-Keizer, A. C., McDonald, S. & Spikman, J. M. Performance of healthy subjects on an ecologically valid test for social cognition: The short, Dutch Version of The Awareness of Social Inference Test (TASIT). *J. Clin. Exp. Neuropsychol.* **36**, 1031–1041 (2014).
31. Bourdage, R., Narme, P., Neeskens, R., Papma, J. & Franzen, S. An evaluation of cross-cultural adaptations of social cognition testing: a systematic review. *Neuropsychol. Rev.* **34**, 1048–1094 (2024).
32. Eddy, C. M. What do you have in mind? Measures to assess mental state reasoning in neuropsychiatric populations. *Front. Psychiatry* **10**, 425 (2019).
33. Dodell-Feder, D., Ressler, K. J. & Germine, L. T. Social cognition or social class and culture? On the interpretation of differences in social cognitive performance. *Psychol. Med.* **50**, 133–145 (2020).
34. Van Staden, J. G. & Callaghan, C. W. An evaluation of the reading the mind in the eyes test's psychometric properties and scores in South Africa—cultural implications. *Psychol. Res.* **86**, 2289–2300 (2022).
35. Alvarez, R. et al. Reading the mind in the eyes and cognitive ability in schizophrenia-and autism spectrum disorders. *Psychol. Med.* **53**, 7913–7922 (2023).

36. Gonzalez, J. M., Rubin, M., Fredrick, M. M. & Velligan, D. I. A qualitative assessment of cross-cultural adaptation of intermediate measures for schizophrenia in multisite international studies. *Psychiatry Res.* **206**, 166–172 (2013).
37. Gregoire, J. ITC guidelines for translating and adapting tests. *Int. J. Test.* **18**, 101–134 (2018).
38. Cruchinho, P. et al. Translation, cross-cultural adaptation, and validation of measurement instruments: a practical guideline for novice researchers. *J. Multidiscip. Healthc.* **17**, 2701–2728 (2024).
39. Quesque, F. et al. Does culture shape our understanding of others' thoughts and emotions? An investigation across 12 countries. *Neuropsychology* **36**, 664 (2022).
40. Bolis, D., Dumas, G. & Schilbach, L. Interpersonal attunement in social interactions: from collective psychophysiology to inter-personalized psychiatry and beyond. *Philos. Trans. R. Soc. B* **378**, 20210365 (2023).
41. Kramer, M. et al. Video analysis of interaction in schizophrenia reveals functionally relevant abnormalities. *Schizophr. Res.* **274**, 24–32 (2024).
42. Quesque, F. et al. Social cognition in neuropsychology: a nationwide survey revealing current representations and practices. *Appl. Neuropsychol. Adult* **31**, 689–702 (2024).
43. Harris, P. A. et al. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inform.* **42**, 377–381 (2009).
44. Dalkey, N. & Helmer, O. An experimental application of the Delphi method to the use of experts. *Manag. Sci.* **9**, 458–467 (1963).
45. Habibi, A., Sarafrazi, A. & Izadyar, S. Delphi technique theoretical framework in qualitative research. *Int. J. Eng. Sci.* **3**, 8–13 (2014).
46. Hutten, J. C. et al. Neuropsychological assessment of aggressive offenders: a Delphi consensus study. *Front. Psychol.* **15**, 1328839 (2024).
47. Franzen, S., Papma, J. M., van den Berg, E. & Nielsen, T. R. Cross-cultural neuropsychological assessment in the European Union: a Delphi expert study. *Arch. Clin. Neuropsychol.* **36**, 815–830 (2021).
48. Okano, H. et al. Evaluation of social cognition measures for Japanese patients with schizophrenia using an expert panel and modified Delphi method. *J. Personal. Med.* **11**, 275 (2021).
49. Turoff, M. The design of a policy Delphi. *Technol. Forecast. Soc. Change* **2**, 149–171 (1970).
50. Raskin, M. S. The Delphi study in field instruction revisited: Expert consensus on issues and research priorities. *J. Soc. Work Educ.* **30**, 75–89 (1994).
51. Rayens, M. K. & Hahn, E. J. Building consensus using the policy Delphi method. *Policy, Politi. Nurs. Pract.* **1**, 308–315 (2000).
52. Pinkham, A. E. et al. The social cognition psychometric evaluation study: results of the expert survey and RAND panel. *Schizophr. Bull.* **40**, 813–823 (2014).
53. Smith, M. J. et al. Performance-based empathy mediates the influence of working memory on social competence in schizophrenia. *Schizophr. Bull.* **40**, 824–834, https://doi.org/10.1093/schbul/sbt084 (2013).
54. Combs, D. R., Penn, D. L., Wicher, M. & Waldheter, E. The Ambiguous Intentions Hostility Questionnaire (AIHQ): a new measure for evaluating hostile social-cognitive biases in paranoia. *Cogn. Neuropsychiatry* **12**, 128–143 (2007).
55. Benton, A. L., Sivan, A., deS, K., Varney, N. R. & Spreen, O. *Facial Recognition: Stimulus and Multiple Choice Pictures: Contributions to Neuropsychological Assessment* (Oxford University Press, Incorporated, 1983).
56. Brüne, M. Theory of mind and the role of IQ in chronic disorganized schizophrenia. *Schizophr. Res.* **60**, 57–64 (2003).
57. Thibaudeau, É., Cellard, C., Legendre, M., Villeneuve, K. & Achim, A. M. Reliability of two social cognition tests: the combined stories test and the social knowledge test. *Psychiatry Res.* **262**, 63–69 (2018).
58. Moritz, S., Woznica, A., Andreou, C. & Köther, U. Response confidence for emotion perception in schizophrenia using a Continuous Facial Sequence Task. *Psychiatry Res.* **200**, 202–207 (2012).
59. van der Gaag, M. et al. Development of the Davos assessment of cognitive biases scale (DACOBS). *Schizophr. Res.* **144**, 63–71 (2013).
60. Van Pelt, B. et al. Dynamic interactive social cognition training in virtual reality (DiSCoVR) for adults with autism spectrum disorder: a feasibility study. *Res. Autism Spectr. Disord.* **96**, 102003 (2022).
61. Muros, N. I. et al. Facial affect recognition by patients with schizophrenia using human avatars. *J. Clin. Med.* **10**, 1904 (2021).
62. Young, A., Perrett, D. I., Calder, A., Sprengelmeyer, R. H. & Ekman, P. Facial expressions of emotion: Stimuli and Test (FEEST). (2002).
63. Scherer, K. R. & Scherer, U. Assessing the ability to recognize facial and vocal expressions of emotion: construction and validation of the Emotion Recognition Index. *J. Nonverbal Behav.* **35**, 305–326 (2011).
64. Kessels, R. P. C., Montagne, B., Hendriks, A. W., Perrett, D. I. & de Haan, E. H. F. Assessment of perception of morphed facial expressions using the Emotion Recognition Task: normative data from healthy participants aged 8–75. *J. Neuropsychol.* **8**, 75–93, https://doi.org/10.1111/jnp.12009 (2014).
65. Fray, P. J., Robbins, T. W. & Sahakian, B. J. Neuorpsychiatyric applications of CANTAB. *Int. J. Geriatr. Psychiatry* **11**, 329–336 (1996).
66. Phillips, L. K., Deldin, P. J., Voglmaier, M. M. & Rabbitt, S. Emotional Stroop performance predicts disorganization in schizophrenia. *Schizophr. Res.* **77**, 141–149 (2005).
67. Zaki, J., Bolger, N. & Ochsner, K. It takes two: the interpersonal nature of empathic accuracy. *Psychol. Sci.* **19**, 399–404 (2008).
68. Derntl, B. et al. Generalized deficit in all core components of empathy in schizophrenia. *Schizophr. Res.* **108**, 197–206 (2009).
69. Kerr, S. L. & Neale, J. M. Emotion perception in schizophrenia: specific deficit or further evidence of generalized poor performance? *J. Abnorm. Psychol.* **102**, 312 (1993).
70. Mehta, U. M. et al. Validation of Social Cognition Rating Tools in Indian Setting (SOCRATIS): a new test-battery to assess social cognition. *Asian J. Psychiatry* **4**, 203–209 (2011).
71. Stone, V. E., Baron-Cohen, S. & Knight, R. T. Frontal lobe contributions to theory of mind. *J. Cogn. Neurosci.* **10**, 640–656 (1998).
72. Bowers, D., Blonder, L. & Heilman, K. *Florida Affect Battery* (Center for Neuropsychological Studies, Department of Neurology Florida, USA, 1998).
73. Ambady, N., Hallahan, M. & Rosenthal, R. On judging and being judged accurately in zero-acquaintance situations. *J. Personal. Soc. Psychol.* **69**, 518 (1995).
74. Happé, F., Brownell, H. & Winner, E. Acquiredtheory of mind'impairments following stroke. *Cognition* **70**, 211–240 (1999).
75. Heider, F. & Simmel, M. An experimental study of apparent behavior. *Am. J. Psychol.* **57**, 243–259 (1944).
76. Gibson, C. M., Penn, D. L., Prinstein, M. J., Perkins, D. O. & Belger, A. Social skill and social cognition in adolescents at genetic risk for psychosis. *Schizophr. Res.* **122**, 179–184 (2010).
77. Klein, H. S. et al. Measuring mentalizing: a comparison of scoring methods for the hinting task. *Int. J. Methods Psychiatr. Res.* **29**, e1827 (2020).
78. Rosset, E. It's no accident: our bias for intentional explanations. *Cognition* **108**, 771–780 (2008).
79. Kinderman, P. & Bentall, R. P. A new measure of causal locus: the internal, personal and situational attributions questionnaire. *Personal. Individ. Differ.* **20**, 261–264 (1996).
80. Davis, Mark H. A multidimensional approach to individual differences in empathy. JSAS Catalog of Selected Documents in Psychology (1980).
81. Lane, R. D., Quinlan, D. M., Schwartz, G. E., Walker, P. A. & Zeitlin, S. B. The Levels of Emotional Awareness Scale: A cognitive-developmental measure of emotion. *J. Personal. Assess.* **55**, 124–134 (1990).
82. Bazin, N. et al. Quantitative assessment of attribution of intentions to others in schizophrenia using an ecological video-based task: a comparison with manic and depressed patients. *Psychiatry Res.* **167**, 28–35 (2009).
83. Mayer, J. D., Salovey, P. & Caruso, D. R. Mayer-Salovey-Caruso emotional intelligence test (MSCEIT) users manual. (2002).
84. Lysaker, P. H. et al. Metacognition amidst narratives of self and illness in schizophrenia: associations with neurocognition, symptoms, insight and quality of life. *Acta Psychiatr. Scand.* **112**, 64–71, (2005).
85. Drury, V. M., Robinson, E. & Birchwood, M. Theory of mind'skills during an acute episode of psychosis and following recovery. *Psychol. Med.* **28**, 1101–1112 (1998).
86. Bertoux, M., Delavest, M., de Souza, LC., Funkiewiez, A., L,pine, JP., Fossati, P., Dubois, B. & Sarazin, M. Social Cognition and Emotional Assessment differentiates frontotemporal dementia from depression. *J. Neurol. Neurosurg. Psychiatry* **83**, 411–416 (2012).
87. Bänziger, T., Scherer, K. R., Hall, J. A. & Rosenthal, R. Introducing the MiniPONS: A short multichannel version of the Profile of Nonverbal Sensitivity (PONS). *J. Nonverbal Behav.* **35**, 189–204 (2011).
88. Barnby, J. M., Mehta, M. A. & Moutoussis, M. The computational relationship between reinforcement learning, social inference, and paranoia. *PLoS Comput. Biol.* **18**, e1010326 (2022).
89. Dziobek, I. et al. Introducing MASC: a movie for the assessment of social cognition. *J. Autism Dev. Disord.* **36**, 623–636 (2006).
90. Dziobek, I. et al. Dissociation of cognitive and emotional empathy in adults with Asperger syndrome using the Multifaceted Empathy Test (MET). *J. Autism Dev. Disord.* **38**, 464–473 (2008).
91. Manera, V. et al. The multilingual CID-5: a new tool to study the perception of communicative interactions in different languages. *Front. Psychol.* **6**, 152558 (2015).
92. Bänziger, T., Grandjean, D. & Scherer, K. R. Emotion recognition from expressions in face, voice, and body: the Multimodal Emotion Recognition Test (MERT). *Emotion* **9**, 691 (2009).
93. Cyranowski, J. M. et al. Assessing social support, companionship, and distress: National Institute of Health (NIH) toolbox adult social relationship scales. *Health Psychol.* **32**, 293 (2013).

94. Healey, K. M. et al. Observable social cognition–a rating scale: an interview-based assessment for schizophrenia. *Cogn. Neuropsychiatry* **20**, 198–221 (2015).
95. Gur, R. C. et al. Age group and sex differences in performance on a computerized neurocognitive battery in children age 8–21. *Neuropsychology* **26**, 251 (2012).
96. Powers, K. E., Worsham, A. L., Freeman, J. B., Wheatley, T. & Heatherton, T. F. Social connection modulates perceptions of animacy. *Psychol. Sci.* **25**, 1943–1948 (2014).
97. Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y. & Plumb, I. 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 Allied Discip.* **42**, 241–251 (2001).
98. Rutherford, M. D., Baron-Cohen, S. & Wheelwright, S. Reading the mind in the voice: A study with normal adults and adults with Asperger syndrome and high functioning autism. *J. Autism Dev. Disord.* **32**, 189–194 (2002).
99. Grisdale, E., Lind, S. E., Eacott, M. J. & Williams, D. M. Self-referential memory in autism spectrum disorder and typical development: Exploring the ownership effect. *Conscious. Cogn.* **30**, 133–141 (2014).
100. Lombardo, M. V., Barnes, J. L., Wheelwright, S. J. & Baron-Cohen, S. Self-referential cognition and empathy in autism. *PloS ONE* **2**, e883 (2007).
101. Dodell-Feder, D., Lincoln, S. H., Coulson, J. P. & Hooker, C. I. Using fiction to assess mental state understanding: a new task for assessing theory of mind in adults. *PloS ONE* **8**, e81279 (2013).
102. American Psychiatric Association, D. & Association, A. P. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*, Vol. 5 (American Psychiatric Association Washington, 2013).
103. Bell, M. D., Fiszdon, J. M., Greig, T. C. & Wexler, B. E. Social attribution test—multiple choice (SAT-MC) in schizophrenia: comparison with community sample and relationship to neurocognitive, social cognitive and symptom measures. *Schizophr. Res.* **122**, 164–171 (2010).
104. Skuse, D. H., Mandy, W. P. & Scourfield, J. Measuring autistic traits: heritability, reliability and validity of the Social and Communication Disorders Checklist. *Br. J. Psychiatry* **187**, 568–572 (2005).
105. Corrigan, P. W., Davies-Farmer, R. M. & Stolley, M. R. Social cue recognition in schizophrenia under variable levels of arousal. *Cogn. Ther. Res.* **14**, 353–361 (1990).
106. Tyrer, P. et al. The Social Functioning Questionnaire: a rapid and robust measure of perceived functioning. *Int. J. Soc. Psychiatry* **51**, 265–275 (2005).
107. Birchwood, M., Smith, J., Cochrane, R., Wetton, S. & Copestake, S. The social functioning scale the development and validation of a new scale of social adjustment for use in family intervention programmes with schizophrenic patients. *Br. J. Psychiatry* **157**, 853–859 (1990).
108. Lin, A., Adolphs, R. & Rangel, A. Social and monetary reward learning engage overlapping neural substrates. *Soc. Cogn. Affect. Neurosci.* **7**, 274–281, https://doi.org/10.1093/scan/nsr006 (2011).
109. Langdon, R., Connors, M. H. & Connaughton, E. Social cognition and social judgment in schizophrenia. *Schizophr. Res. Cogn.* **1**, 171–174 (2014).
110. Behrens, T. E., Hunt, L. T., Woolrich, M. W. & Rushworth, M. F. Associative learning of social value. *Nature* **456**, 245–249 (2008).
111. Diaconescu, A. O., Hauke, D. J. & Borgwardt, S. Models of persecutory delusions: a mechanistic insight into the early stages of psychosis. *Mol. Psychiatry* **24**, 1258–1267 (2019).
112. Kramer, J. H. et al. NIH EXAMINER: conceptualization and development of an executive function battery. *J. Int. Neuropsychol. Soc.* **20**, 11–19 (2014).
113. Okruszek, Ł. & Chrustowicz, M. Social perception and interaction database—a novel tool to study social cognitive processes with point-light displays. *Front. Psychiatry* **11**, 496448 (2020).
114. Patterson, T. L., Moscona, S., McKibbin, C. L., Davidson, K. & Jeste, D. V. Social skills performance assessment among older patients with schizophrenia. *Schizophr. Res.* **48**, 351–360 (2001).
115. Murray, K. et al. A new test of advanced theory of mind: the "Strange Stories Film Task" captures social processing differences in adults with autism spectrum disorders. *Autism Res.* **10**, 1120–1132 (2017).
116. Happé, F. G. An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *J. Autism Dev. Disord.* **24**, 129–154 (1994).
117. Gaudelus, B. et al. Measuring impairment of facial affects recognition in schizophrenia. Preliminary study of the facial emotions recognition task (TREF). *Encephale* **41**, 251–259 (2014).
118. McDonald, S., Flanagan, S., Rollins, J. & Kinch, J. TASIT: a new clinical tool for assessing social perception after traumatic brain injury. *J. Head. Trauma Rehabil.* **18**, 219–238 (2003).
119. Long, E. L., Cuve, H. C., Conway, J. R., Catmur, C. & Bird, G. Novel theory of mind task demonstrates representation of minds in mental state inference. *Sci. Rep.* **12**, 21133 (2022).
120. Erwin, R. J. et al. Facial emotion discrimination: I. Task construction and behavioral findings in normal subjects. *Psychiatry Res.* **42**, 231–240 (1992).
121. Sui, J., He, X. & Humphreys, G. W. Perceptual effects of social salience: evidence from self-prioritization effects on perceptual matching. *J. Exp. Psychol. Hum. Percept. Perform.* **38**, 1105 (2012).
122. Aykan, S. & Nalçacı, E. Assessing theory of mind by humor: the humor comprehension and appreciation test (ToM-HCAT). *Front. Psychol.* **9**, 382586 (2018).
123. Desgranges, B. et al. TOM-15: A false-belief task to assess cognitive theory of mind. *Rev. Neuropsychol.* **4**, 216–220 (2012).
124. Abell, F., Happe, F. & Frith, U. Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cogn. Dev.* **15**, 1–16 (2000).
125. Canty, A. L., Neumann, D. L., Fleming, J. & Shum, D. H. Evaluation of a newly developed measure of theory of mind: the virtual assessment of mentalising ability. *Neuropsychol. Rehabil.* **27**, 834–870 (2017).
126. Moritz, S. et al. Whodunit–a novel video-based task for the measurement of jumping to conclusions in the schizophrenia spectrum. *Psychiatry Res.* **317**, 114862 (2022).
127. Ratajska, A., Brown, M. I. & Chabris, C. F. Attributing social meaning to animated shapes: a new experimental study of apparent behavior. *Am. J. Psychol.* **133**, 295–312 (2020).
128. Kelly, M. & McDonald, S. Assessing social cognition in people with a diagnosis of dementia: development of a novel screening test, the Brief Assessment of Social Skills (BASS-D). *J. Clin. Exp. Neuropsychol.* **42**, 185–198 (2020).
129. Sarfati, Y., Hardy-Baylé, M.-C., Besche, C. & Widlöcher, D. Attribution of intentions to others in people with schizophrenia: a non-verbal exploration with comic strips. *Schizophr. Res.* **25**, 199–209 (1997).
130. Ratto, A. B., Turner-Brown, L., Rupp, B. M., Mesibov, G. B. & Penn, D. L. Development of the contextual assessment of social skills (CASS): a role play measure of social skill for individuals with high-functioning autism. *J. Autism Dev. Disord.* **41**, 1277–1286 (2011).
131. Baksh, R. A., Abrahams, S., Auyeung, B. & MacPherson, S. E. The Edinburgh Social Cognition Test (ESCoT): examining the effects of age on a new measure of theory of mind and social norm understanding. *PloS ONE* **13**, e0195818 (2018).
132. Montagne, B., Kessels, R. P., De Haan, E. H. & Perrett, D. I. The emotion recognition task: a paradigm to measure the perception of facial emotional expressions at different intensities. *Percept. Mot. Skills* **104**, 589–598 (2007).
133. Baron-Cohen, S. & Wheelwright, S. The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *J. Autism Dev. Disord.* **34**, 163–175 (2004).
134. Costanzo, M. & Archer, D. Interperting the expressive behavior of others: the Interpersonal Perception Task. *J. Nonverbal Behav.* **13**, 225–245 (1989).
135. Douglas, J. M., O'Flaherty, C. A. & Snow, P. C. Measuring perception of communicative ability: the development and evaluation of the La Trobe Communication Questionnaire. *Aphasiology* **14**, 251–268 (2000).
136. Reniers, R. L., Corcoran, R., Drake, R., Shryane, N. M. & Völlm, B. A. The QCAE: a questionnaire of cognitive and affective empathy. *J. Personal. Assess.* **93**, 84–95 (2011).

ACKNOWLEDGEMENTS

We would like to thank Linlin Fan, Ph.D., Sara Carrico, Lauren Bennett-Leleux, M.S., and Sarah Berretta for their assistance with various stages of the project. The study was supported by Schizophrenia International Research Society – Research Harmonization Award. TVR (group member) was supported by an AI and Val Rosenstrauss Fellowship from the Rebecca L Cooper Medical Research Foundation.

AUTHOR CONTRIBUTIONS

A.E.P., M.H., and T.Z. contributed to study conceptualization, methodological design, project administration, acquisition of study resources and funding, and reviewing and editing drafts of the manuscript. A.E.P. conducted the analyses and drafted the manuscript. Research Harmonization Group members participated in the expert survey and Delphi process, provided feedback at all stages of the study, edited drafts of the manuscript, and provided approval of the final version.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41537-024-00540-7.

Correspondence and requests for materials should be addressed to Amy E. Pinkham.

THE SIRS SOCIAL COGNITION RESEARCH HARMONIZATION GROUP

Minji Bang[6], Maxime Bertoux[7], Bodi Bodenhamer[8], Kelsey A. Bonfils[9], A. M. Buunk[10], Raymond C. K. Chan[11,12], Beshaun Davis[13], Clare M. Eddy[14], Anne-Kathrin Fett[15,16], Ana Flores[17], Michal Hajdúk[2,3,4], Taeyoung Lee[18], Lucy Livingston[19,20], Monica Mazza[21], Skye McDonald[22], Urvakhsh Meherwan Mehta[23], Arundati Nagendra[24], Lindsay D. Oliver[25], David L. Penn[26], Amy Pinkham[1], Bram-Sieben Rosema[27], Julia Sheffield[28], Andrew Spink[29], Tamara Tavares[30], Tamsyn Van Rheenen[31,32] and Tim Ziermans[33]

[6]Department of Psychiatry, CHA Bundang Medical Center, CHA University School of Medicine, Seongnam, Republic of Korea. [7]Lille Neuroscience and Cognition, Inserm, University Lille, CHU Lille, Lille, France. [8]Department of Psychiatry, School of Medicine, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [9]School of Psychology, College of Education and Human Sciences, The University of Southern Mississippi, Hattiesburg, MS, USA. [10]Department of Neurology, Unit Neuropsychology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [11]Neuropsychology and Applied Cognitive Neuroscience Laboratory, CAS Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing, China. [12]Department of Psychology, The University of Chinese Academy of Sciences, Beijing, China. [13]Department of Psychiatry, University of Maryland School of Medicine, Baltimore, MD, USA. [14]Research and Development, Birmingham and Solihull Mental Health NHS Foundation Trust, and College of Medicine and Health, University of Birmingham, Birmingham, UK. [15]Department of Psychology, City St George's, University of London, School of Health & Medical Sciences, London, UK. [16]Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK. [17]Silver School of Social Work, New York University, New York, NY, USA. [18]Department of Psychiatry, Kyungpook National University School of Medicine, Daegu, Republic of Korea. [19]Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, Birmingham, UK. [20]Neuroscience and Mental Health Innovation Institute, Cardiff University, Birmingham, UK. [21]Department of Biotechnological and Applied Clinical Sciences, University of L'Aquila, L'Aquila, Italy. [22]School of Psychology, University of New South Wales, Sydney, NSW, Australia. [23]Department of Psychiatry, National Institute of Mental Health and Neuro-Sciences (NIMHANS), Bangalore, India. [24]Schizophrenia & Psychosis Action Alliance, Alexandria, VA, USA. [25]Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, ON, Canada; Department of Psychiatry, University of Toronto, Toronto, ON, Canada. [26]Department of Psychiatry, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. [27]HAMLETT study and "Kenniswerkplaats Onbegrepen Gedrag", NHL Stenden Leeuwarden, and Expert by Experience for Anoiksis, Patiënt Association for Susceptibility for Psychosis, Leeuwarden, The Netherlands. [28]Department of Psychiatry and Behavioral Sciences, Vanderbilt University Medical Center, Nashville, TN, USA. [29]Noldus Information Technology BV, Wageningen, The Netherlands. [30]Department of Psychology, The Hospital for Sick Children, York University, Toronto, ON, Canada. [31]Department of Psychiatry, University of Melbourne, Faculty of Medicine, Dentistry and Health Sciences, Melbourne, VIC, Australia. [32]Centre for Mental Health and Brain Sciences, Swinburne University of Technology, Sydney, Melbourne, VIC, Australia. [33]Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands.