



City Research Online

City, University of London Institutional Repository

Citation: Ziermans, T. B., Hajdúk, M., Pinkham, A. E., SIRS Social Cognition Research Harmonization Group & Fett, A-K. (2026). Call to action on social cognition measures in clinical research. *Schizophrenia Research: Cognition*, 43, 100400. doi: 10.1016/j.scog.2025.100400

This is the published version of the paper.

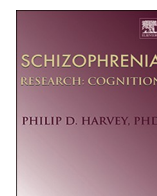
This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/36445/>

Link to published version: <https://doi.org/10.1016/j.scog.2025.100400>

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.



Short Communication

Call to action on social cognition measures in clinical research

T.B. Ziermans^{a,*}, M. Hajdúk^{b,c,d}, A.E. Pinkham^e, SIRS Social Cognition Research Harmonization Group^{a,b,c,d,e,1}

^a Department of Psychology, University of Amsterdam, Amsterdam, the Netherlands

^b Department of Psychology, Faculty of Arts, Comenius University in Bratislava, Bratislava, Slovak Republic

^c The Centre for Psychiatric Disorders Research, Science Park, Comenius University in Bratislava, Bratislava, Slovak Republic

^d Department of Psychiatry, Faculty of Medicine, Comenius University in Bratislava, Bratislava, Slovak Republic

^e Department of Psychology, School of Behavioral and Brain Sciences, The University of Texas at Dallas, Richardson, TX, USA



ARTICLE INFO

Keywords:

Social cognition
Assessment
Neuropsychology
Validation
Harmonisation
Crosscultural

ABSTRACT

Objective: To describe current practices and key barriers in social cognition (SC) assessment, given its central role in psychiatric and neurological disorders and the limitations of existing measures.

Methods: Fifty-two SC experts from 20 countries completed an online survey regarding SC tests and questions about their usage frequency and perceived obstacles.

Results: Only facial emotion recognition tasks were used frequently, while the Hinting task and Reading the Mind in the Eyes Test (RMET) were used by over half of participants. However, 10 experts also urged discontinuation of RMET, mostly due to validity concerns. Major obstacles included lack of culture-appropriate norms and poor psychometric properties.

Conclusions: SC assessment is limited by cultural bias and weak psychometrics. Developing and validating culturally sensitive tools, harmonizing protocols, and securing funding are essential to advance research, enable international trials, and improve clinical outcomes.

1. Introduction

Social cognition (SC), broadly referring to the mental capacity to perceive, interpret and respond to social information, is impaired across a broad spectrum of disorders (Cotter et al., 2018; Henry et al., 2016). These deficits are generally associated with poorer real-world functioning and reduced quality of life (Halverson et al., 2019). Therefore, SC measures can serve as an important clinical and research tool to improve our understanding of social difficulties across clinical conditions (Van den Stock et al., 2023).

In recent decades the number and transdiagnostic utilization of SC tests has rapidly expanded to a vast array of clinical and non-clinical specializations, resulting in a rich variety of available instruments and paradigms. Unfortunately, many of these measures suffer from shortcomings such as poor psychometric properties and cultural bias and continue to be used, despite criticism (Bourdage et al., 2024; Pinkham et al., 2018). Consequently, there is an urgent need for harmonized, large-scale, international efforts to develop a universal best practice

guideline for SC testing in research and clinical settings.

To address this, we previously conducted an international survey and Delphi consensus study among multidisciplinary SC experts with an international Research Harmonization Group (RHG). The goal was to identify the best transdiagnostic and cross-cultural SC measures to stimulate future harmonization among researchers and clinicians. The outcome showed that relatively few existing measures were considered adequate (Pinkham et al., 2025). Here, we report additional results from the global survey to provide further insights into the status quo of SC research practices and identify key obstacles to improve SC assessment.

2. Methods

Online survey content and expert definitions were drafted by authors and further refined with input from RHG-members. The term “expert” for academic researchers was defined as follows:

DOI of original article: <https://doi.org/10.1016/j.scog.2025.100405>.

* Corresponding author.

E-mail address: t.b.ziermans@uva.nl (T.B. Ziermans).

¹ A list of all contributors and their affiliations appears at the end of the paper.

<https://doi.org/10.1016/j.scog.2025.100400>

Received 18 September 2025; Accepted 10 October 2025

Available online 18 November 2025

2215-0013/© 2025 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

- 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 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 encouraged to complete the survey as well. For detail on additional survey procedures, we refer to [Pinkham et al. \(2025\)](#). The study was approved by the Institutional Review Board (IRB) of The University of Texas at Dallas (IRB-23-177).

The RHG pre-defined a set of 18 well-known SC tests from the literature for which respondents could indicate how often they used them on a 4-point scale. In addition, participants could insert additional tests they use and indicate which three tests they use the most. They were also asked if certain tasks should no longer be used, including their rationale. To evaluate obstacles in SC assessments, participants were asked to rate the extent to which each of 10 predefined limitations posed a problem on a 7-point Likert scale. In addition, they could suggest alternative problems and indicate what they considered to be the biggest obstacles limiting progress in international SC research.

3. Results

Fifty-two experts filled out >50 % of items for this final part of the survey. They resided in 20 countries, predominantly in North America (35 %) and Europe (50 %). Most experts identified as male (58 %) and 73 % indicated professor/lecturer (any level) as their profession. Schizophrenia/psychosis was checked by 37 % of experts as their main study population of interest, followed by high-risk for psychosis and general population (both 17 %) and 7 additional subdomains in psychiatry or neurology. For additional detail on respondent characteristics, see Supplementary Table 1.

3.1. Frequency of use

[Table 1](#) gives an overview of the user frequency of common SC measures. Only facial emotion recognition tasks (such as ER-40) are used often by experts, and more than half of participants reported using the Hinting task and Reading the Mind in the Eyes Test (RMET) at least occasionally. For the 15 other included tasks, most experts indicated that they never use them. Finally, 16 experts indicated some SC tasks should no longer be used of whom 10 mentioned the RMET. This was motivated by issues with construct validity and confounders (e.g., the influence of vocabulary).

3.2. Obstacles

[Table 2](#) shows how experts evaluated the 10 pre-defined problems within SC assessment. While all obstacles were considered moderately problematic on average, the highest mean scores were reserved for two related issues concerning cross-cultural applicability, i.e., the availability of appropriate norms and the availability of paradigms that can

Table 1

Central tendency measures for use of common social cognition measures.

	N	Mean	Med	SD	#Top3
Emotion Recognition - 40 (ER40), or equivalent ER task	52	2.52	3	1.00	25
Hinting Task	52	2.17	2	1.04	8
Reading the Mind in the Eyes Test (RMET/Eyes Task)	51	2.16	2	0.88	10
The Awareness of Social Inferences Test (TASIT)	52	1.79	1	0.96	8
Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT)	52	1.77	1	0.98	8
Faux Pas Task	51	1.63	1	0.75	5
Cartoon Theory of Mind (CToM)	52	1.62	1	0.89	7
Ambiguous Intentions and Hostility Questionnaire (AIHQ)	52	1.54	1	0.92	4
Bell Lysaker Emotion Recognition Test (BLERT)	50	1.46	1	0.71	3
Empathic Accuracy Task (EAT)	50	1.42	1	0.79	5
Trustworthiness Task	52	1.37	1	0.77	1
Mini Profile of Nonverbal Sensitivity (MiniPONS)	52	1.35	1	0.79	1
Emotion In Biological Motion Task (EBM)	52	1.25	1	0.65	3
Point-light walkers	52	1.19	1	0.56	1
Relationships Across Domains Task (RAD)	52	1.17	1	0.51	1
Intentionality Bias Task (IBT)	50	1.16	1	0.42	0
Social Attribution Task, Multiple Choice (SAT-MC)	51	1.16	1	0.46	1
Observable Social Cognition Rating Scale (OSCARS)	52	1.15	1	0.46	1

Note. This table shows frequency scores (mean, median) and their standard deviation (SD) for use of common social cognition tests by experts. Scores were based on a 4-point scale (1 = Never, 2 = Rarely, 3 = Often, 4 = Always). The final column shows how often a test was included in the top 3 of SC tests participants used most frequently.

Table 2

Central tendency measures for perceived obstacles in using social cognition measures.

	N	Mean	Med	SD
Lack of normative data and/or culture-specific norms	51	4.94	5	1.50
Problems for cross-cultural comparisons (e.g., stimuli/language)	50	4.74	5	1.60
Poor ecological validity	51	4.65	5	1.41
Weak psychometric properties	51	4.49	5	1.65
Floor/ceiling effects/limited sensitivity to capture differences	51	4.35	5	1.41
Low clinical utility (e.g., for differential diagnosis, treatment)	51	4.12	4	1.74
Utility for repeated administrations (e.g., in RCTs)	51	4.08	4	1.50
Lack of support for associations with functional outcomes	51	3.90	4	1.76
Too lengthy or cumbersome	51	3.78	4	1.40
Limited availability of test in my country	50	3.30	2.5	2.22

Note. This table shows frequency scores (Mean, Median) and their standard deviation (SD) for perceived obstacles in using social cognition tests by experts. Scores were based on a 7-point Likert-scale (1 = Not at all; 7 = Very) and indicated to what extent an obstacle was considered problematic.

be used internationally. Regarding the biggest obstacle for progression in international SC research, the diverse answers were categorized based on author consensus (see Supplemental Table 2). Most comments and themes aligned with obstacles mentioned in [Table 2](#), with insufficient test norms and validation, as well as cultural/language adaptations standing out. However, two novel themes that were reported were lack of funding and collaboration opportunities as well as theoretical/conceptual ambiguity.

4. Discussion

The findings from our global survey show that SC assessment is currently narrow and likely constrained by several perceived barriers. Straight-forward emotion perception tasks such as the ER-40 are the only type of measurement that is commonly used among experts, followed by tests measuring mentalization and empathy. Notably, the RMET is both one of the most widely used and most frequently criticized social cognition measures. In our survey, 19 % of experts recommended discontinuing its use, citing challenges in interpretation and confounding influences, such as verbal ability. These findings echo methodological criticisms from the literature (Higgins et al., 2025). Yet they do not appear to deter even narrowly defined SC experts from using RMET or other poorly evaluated SC tasks. This ambivalent attitude in the field highlights a lack of consensus, perhaps driven by insufficient theoretical/conceptual clarity on what SC tests measure (Eikelboom et al., 2025; Quesque et al., 2024).

The most prominently endorsed barriers were the lack of culture-appropriate norms and paradigms as well as poor psychometric properties. There is a strong need for funding to develop psychometrically sound and ecologically valid SC tools. Cultural considerations for such tests can be advanced by systematically developing culture-neutral, culture-sensitive, and culture-adapted tests, each accompanied by robust normative data both within and across diverse ethnocultural groups to ensure validity, fairness, and comparability. Meanwhile, it is important to remain mindful that social stimuli are interpreted within diverse sociocultural contexts and to consider modular assessment batteries to better balance cross-cultural and culture-specific elements. Together, these are essential foundations for large-scale (international) clinical trials using harmonized protocols. Even in well-powered trials with effective interventions, inadequate measurement tools may obscure treatment effects and hinder detection of meaningful change, e. g., in functioning or quality of life (Eddy, 2019). We therefore recommend prioritizing the validation of core measures and systematically adapting them across cultures before expanding to new tools.

Dedicated funding and sustained international collaboration are required to improve norm and test development. Commercial publishers and online platforms may contribute, but academic leadership, focused on both international and non-academic partnerships, is essential. We also advocate for targeted funding schemes to support international consortia.

While modern technologies enable large-scale measurement initiatives, these efforts cannot succeed without proper financial support. Clinicians require precise and reliable norms and valid tools for accurate decision-making and monitoring. Basic research should guide assessment choices, while clinical feedback should inform further refinement. Strengthening collaboration between academia and clinicians is key to bridging the research-practice gap.

In conclusion, our survey of global SC experts emphasizes the need for, and further amplifies recent calls for intensified international collaborative efforts to develop more applicable and psychometrically sound gold standard measures for SC (Corbera et al., 2025). Without these steps, it will be difficult to test the efficacy and effectiveness of SC interventions in large scale international clinical trials or routinely use SC assessment in clinical practice.

Consortia

The SIRS Social Cognition Research Harmonization Group:

Co-authors:

Maxime Bertoux⁶, Kelsey A. Bonfils⁷, Anne M. Buunk⁸, Clare M. Eddy⁹, Anne-Kathrin Fett^{10,11}, Michal Hajdúk^{2,3,4}, Monica Mazza¹², Urvakhsh Meherwan Mehta¹³, Lindsay D. Oliver¹⁴, David L. Penn¹⁵, Amy Pinkham⁵, Tamsyn Van Rheenen^{16,17} and Tim Ziermans¹

⁶ Lille Neuroscience and Cognition, Inserm, University Lille, CHU Lille, Lille, France.

⁷ Department of Psychology and Neuroscience, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

⁸ Department of Neurology, Unit Neuropsychology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands.

⁹ Research and Development, Birmingham and Solihull Mental Health NHS Foundation Trust, and College of Medicine and Health, University of Birmingham, Birmingham, UK.

¹⁰ Department of Psychology, City St George's, University of London, School of Health & Medical Sciences, London, UK.

¹¹ Department of Psychosis Studies, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK.

¹² Department of Biotechnological and Applied Clinical Sciences, University of L'Aquila, L'Aquila, Italy.

¹³ Department of Psychiatry, National Institute of Mental Health and Neuro-Sciences (NIMHANS), Bangalore, India.

¹⁴ Campbell Family Mental Health Research Institute, Centre for Addiction and Mental Health, Toronto, ON, Canada; Department of Psychiatry, University of Toronto, Toronto, ON, Canada.

¹⁵ Department of Psychiatry, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

¹⁶ Department of Psychiatry, University of Melbourne, Faculty of Medicine, Dentistry and Health Sciences, Melbourne, VIC, Australia.

¹⁷ Centre for Mental Health and Brain Sciences, Swinburne University of Technology, Sydney, Melbourne, VIC, Australia.

Non-author collaborators:

Minji Bang¹⁸, Bodi Bodenhamer¹⁹, Raymond C. C. Chan^{20,21}, Beshawn Davis²², Ana Flores²³, Taeyoung Lee²⁴, Lucy Livingston^{25,26}, Skye McDonald²⁷, Arundati Nagendra²⁸, Bram-Sieben Rosema²⁹, Julia Sheffield³⁰, Andrew Spink³¹, Tamara Tavares³²

¹⁸ Department of Psychiatry, CHA Bundang Medical Center, CHA University School of Medicine, Seongnam, Republic of Korea.

¹⁹ Department of Psychiatry, School of Medicine, The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA.

²⁰ Neuropsychology and Applied Cognitive Neuroscience Laboratory, CAS Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing, China.

²¹ Department of Psychology, The University of Chinese Academy of Sciences, Beijing, China.

²² Department of Psychiatry, University of Maryland School of Medicine, Baltimore, MD, USA.

²³ Silver School of Social Work, New York University, New York, NY, USA.

²⁴ Department of Psychiatry, Kyungpook National University School of Medicine, Daegu, Republic of Korea.

²⁵ Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, Birmingham, UK.

²⁶ Neuroscience and Mental Health Innovation Institute, Cardiff University, Birmingham, UK.

²⁷ School of Psychology, University of New South Wales, Sydney, NSW, Australia.

²⁸ Schizophrenia & Psychosis Action Alliance, Alexandria, VA, USA.

²⁹ HAMLETT study and "Kenniserwerkplaats Onbegrepen Gedrag", NHL Stenden Leeuwarden, and Expert by Experience for Anoisks, Patient Association for Susceptibility for Psychosis, Leeuwarden, The Netherlands.

³⁰ Department of Psychiatry and Behavioral Sciences, Vanderbilt University Medical Center, Nashville, TN, USA.

³¹ Noldus Information Technology BV, Wageningen, The Netherlands.

³² Department of Psychology, The Hospital for Sick Children, York University, Toronto, ON, Canada.

CRedit authorship contribution statement

T.B. Ziermans: Conceptualization, Data curation, Formal analysis,

Funding acquisition, Investigation, Methodology, Visualization, Writing – original draft. **M. Hajdúk:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Writing – review & editing. **A.E. Pinkham:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

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 (SIRS)–Research Harmonization Award. TVR (group member) was supported by an AI and Val Rosenstrauss Fellowship from the Rebecca L Cooper Medical Research Foundation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scog.2025.100400>.

References

- Bourdage, R., Narme, P., Neeskens, R., Papma, J., Franzen, S., 2024. An evaluation of cross-cultural adaptations of social cognition testing: a systematic review. *Neuropsychol. Rev.* 34 (4), 1048–1094. <https://doi.org/10.1007/s11065-023-09616-0>.
- Corbera, S., Kurtz, M.M., Achim, A.M., et al., 2025. International perspective on social cognition in schizophrenia: current stage and the next steps. *Eur. Psychiatry* 68 (1), e9. <https://doi.org/10.1192/j.eurpsy.2024.1776>.
- Cotter, J., Granger, K., Backx, R., Hobbs, M., Looi, C.Y., Barnett, J.H., 2018. Social cognitive dysfunction as a clinical marker: a systematic review of meta-analyses across 30 clinical conditions. *Neurosci. Biobehav. Rev.* 84, 92–99. <https://doi.org/10.1016/j.neubiorev.2017.11.014>.
- Eddy, C.M., 2019. What do you have in mind? Measures to assess mental state reasoning in neuropsychiatric populations. *Front. Psychiatry*. 10, 425. <https://doi.org/10.3389/fpsy.2019.00425>.
- Eikelboom, W.S., van den Berg, E., Beauchamp, M.H., et al., 2025. Providing a taxonomy for social cognition: how to bridge the gap between expert opinion, empirical data, and theoretical models. *J. Psychiatry Neurosci.* 50 (3), E157–E161. <https://doi.org/10.1503/jpn.240156>.
- 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, 212–219. <https://doi.org/10.1016/j.neubiorev.2019.07.020>.
- Henry, J.D., von Hippel, W., Molenberghs, P., Lee, T., Sachdev, P.S., 2016. Clinical assessment of social cognitive function in neurological disorders. *Nat. Rev. Neurol.* 12 (1), 28–39. <https://doi.org/10.1038/nrneurol.2015.229>.
- Higgins, W.C., Kaplan, D.M., Deschrijver, E., Ross, R.M., 2025. Why most research based on the Reading the Mind in the Eyes Test is unsubstantiated and uninterpretable: a response to Murphy and Hall (2024). *Clin. Psychol. Rev.* 115, 102530. <https://doi.org/10.1016/j.cpr.2024.102530>.
- 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., Hajdúk, M., Ziermans, T., SIRS Social Cognition Research Harmonization Group, 2025. Harmonizing cross-cultural and transdiagnostic assessment of social cognition by expert panel consensus. *Schizophrenia* 11 (1), 25. <https://doi.org/10.1038/s41537-024-00540-7>.
- Quesque, F., Apperly, I., Baillargeon, R., et al., 2024. Defining key concepts for mental state attribution. *Commun. Psychol.* 2 (1), 29. <https://doi.org/10.1038/s44271-024-00077-6>.
- Van den Stock, J., Bertoux, M., Diehl-Schmid, J., et al., 2023. Current potential for clinical optimization of social cognition assessment for frontotemporal dementia and primary psychiatric disorders. *Neuropsychol. Rev.* 33 (2), 544–550. <https://doi.org/10.1007/s11065-022-09554-3>.