brain sciences

MDPI

Comment

Comment on Panuccio et al. Quality of Assessment Tools for Aphasia: A Systematic Review. *Brain Sci.* 2025, *15*, 271

Sarah J. Wallace [1,2,*], Katerina Hilari [3], Klaus Willmes [4], Marcus Meinzer [5], Claudia Peñaloza [6], Lizet van Ewijk [7], Rebecca Palmer [8], Sally Zingelman [1,2], William D. Hula [9,10,11] and Caterina Breitenstein [12] on behalf of The Collaboration of Aphasia Trialists (CATs)

1 Queensland Aphasia Research Centre, The University of Queensland, Brisbane 4029, Australia; s.zingelman@uq.edu.au
2 Surgical Treatment and Rehabilitation Service (STARS) Education and Research Alliance, The University of Queensland and Metro North, Brisbane 4029, Australia
3 School of Health and Medical Sciences, Centre for Language and Communication Science Research, City St George's, University of London, London EC1V 0HB, UK; k.hilari@citystgeorges.ac.uk
4 Department of Neurology, RWTH Aachen University Hospital, 52074 Aachen, Germany; willmes@neuropsych.rwth-aachen.de
5 Department of Neurology, University Medicine Greifswald, 17489 Greifswald, Germany; marcus.meinzer@med.uni-greifswald.de
6 Department of Cognition, Development and Educational Psychology, Institute of Neurosciences, University of Barcelona, 08035 Barcelona, Spain; claudia_penaloza@ub.edu
7 Research Group Speech and Language Therapy, Research Center Healthy and Sustainable Living, HU University of Applied Sciences, 3584CS Utrecht, The Netherlands; lizet.vanewijk@hu.nl
8 Sheffield Centre for Health and Related Research, University of Sheffield, Sheffield S10 2TN, UK; r.l.palmer@sheffield.ac.uk
9 Geriatric Research, Education, and Clinical Center, VA Pittsburgh Healthcare System, Pittsburgh, PA 15240, USA; william.hula@va.gov
10 Audiology and Speech Pathology Service, VA Pittsburgh Healthcare System, Pittsburgh, PA 15240, USA
11 Department of Communication Science and Disorders, University of Pittsburgh, Pittsburgh, PA 15260, USA
12 Department of Neurology, University of Muenster, 48149 Muenster, Germany; breitens@uni-muenster.de
* Correspondence: s.wallace3@uq.edu.au

Academic Editors: Tracy E. Love and Stephen D. Meriney

Received: 11 April 2025
Revised: 24 September 2025
Accepted: 29 October 2025
Published: 17 November 2025

Citation: Wallace, S.J.; Hilari, K.; Willmes, K.; Meinzer, M.; Peñaloza, C.; van Ewijk, L.; Palmer, R.; Zingelman, S.; Hula, W.D.; Breitenstein, C., on behalf of The Collaboration of Aphasia Trialists (CATs). Comment on Panuccio et al. Quality of Assessment Tools for Aphasia: A Systematic Review. *Brain Sci.* 2025, *15*, 271. *Brain Sci.* **2025**, *15*, 1233. https://doi.org/10.3390/brainsci15111233

We write on behalf of the Collaboration of Aphasia Trialists and in response to an article recently published in the journal *Brain Sciences*: Quality of Assessment Tools for Aphasia: A Systematic Review by Panuccio and colleagues [1]. While we applaud the authors' efforts to provide a comprehensive review of aphasia measurement instruments, we have identified numerous significant methodological concerns and factual errors that undermine the quality, validity, and utility of this review. We have outlined some of these concerns below. In the interest of providing a timely response, these constitute only selected examples.

First, an outdated version of the COSMIN quality rating criteria [2] is used instead of the more recent checklist [3], and there are substantial errors in its application, including that, in many cases, citations do not support the ratings made. For example:

- The study evaluating the Turkish version of the Aphasia Rapid Test (ART) [4] is awarded the highest quality rating of all measurement instruments in the systematic review, with uniformly positive ratings across 9 of 10 COSMIN criteria, despite the supporting paper only evaluating one aspect of one of the COSMIN quality criteria (inter-rater agreement as one aspect of reliability).
- The paper reporting on the adaptation of the Stroke Specific Quality of Life scale (SS-QOL, Williams et al., 1999) [5] to develop an aphasia-adapted version, the English language Stroke and Aphasia Quality of Life Scale (SAQOL-39) and test its content validity [6] is listed as a Dutch publication in Table 2 and not considered in Table 3

for the development of the SAQOL-39. The structural validity of the original English-language SAQOL-39 [7] and SAQOL-39g [8] is rated as insufficient despite both studies reporting results of Exploratory Factor Analysis, while the Japanese SAQOL-39 [9] received a positive rating despite no reported factor analysis at all in the cited article.

- The Aphasia Communication Outcome Measure (ACOM) was rated negatively for internal consistency, even though the cited paper [10] reports an IRT- based marginal reliability coefficient, an internal consistency measure.
- For psychometric evaluation of the original German-language Communicative Activity Log (CAL), the authors refer to an evaluation study for the Korean version of the CAL [[11], Table 2], which does not include any data for the German CAL, and merely cites a review article for the German CAL standardization. This review article includes the CAL questions in an appendix, without reporting any psychometric data.
- Ratings for the Communication Participation Item Bank (CPIB) [12] do not accurately reflect available information on Patient Reported Outcome Measure (PROM) development or psychometric information related to item response theory (IRT) analyses [12].
- In many cited articles that include general stroke samples, the proportion of people with aphasia is not specified, for example, for the German-language screening (LAST) [13]. It therefore remains unclear whether the corresponding measurement instrument has even been evaluated in the target population (people with aphasia) at all.

These examples raise significant questions about the rigour of the quality assessment process and erode the confidence we can have in the findings. Furthermore, Table 3 lacks documentation supporting the authors' quality ratings, making it impossible to verify their judgments.

Second, the review is limited by selection bias, having excluded published test manuals containing robust standardization data for well-established aphasia measures. Previous reviews on aphasia assessment instruments have underscored that search strategies restricted to just research databases (i.e., peer-reviewed articles) are likely to miss available psychometric data [14–16]. Similarly, other systematic review authors have noted that "comprehensive language assessments often report their psychometric properties within their purchased test manuals or through online sources and not within peer-reviewed journals" [[17], p. 3] leading them to refine their search strategy to include test manuals and other sources to access their psychometric data. In this view, the review by Panuccio and colleagues [1] shows notable omissions, including the test manuals for:

- The German-language Aachen Aphasia Test (AAT) [18] and Scenario Test [19], and the Dutch Amsterdam-Nijmegen Everyday Language Test (ANELT) [20] all of which contain substantial psychometric data, with the AAT being one of the most extensively psychometrically evaluated measures in the field.
- The original English version of the Comprehensive Aphasia Test (CAT) [21], a significant oversight given the CAT's increasing importance in aphasia research over the past two decades and the multiple adaptations it has inspired [22].

Third, the structure used to categorize measurement instruments (Table 2) lacks coherence, sometimes referring to constructs, sometimes to target populations, and other times to combinations of populations and recovery phases. The fundamental distinction between language and communication is not recognized in this structure, a significant oversight when it comes to aphasia measurement instruments. Accordingly, measures of communication, e.g., the Scenario Test [19], Aphasia Communication Outcome Measure ACOM (Hula et al., 2015) [10], and American Speech-Language and Hearing Association Functional Assessment of Communication Skills for Adults [23], are miscategorized as measures of language. Within the category structure, there are numerous further exam-

ples of measurement instruments being incorrectly categorized. Such examples included the following:

- The Apraxia of Speech Rating Scale [24] is categorized as a language measure when it assesses apraxia of speech, a *motor speech disorder*.
- The Abbey Pain Scale [25] is categorized as a measure of language, when it measures *pain*.
- The Auditory-Perceptual Rating of Connected Speech in Aphasia [26] is categorized as an "Auditory-perceptive" measure, rather than a multidimensional measure of *connected speech performance*.
- The ACOM [10] is categorized as a quality-of-life measure, when its authors specifically identify it as a patient-reported measure of *communicative function*.
- The CPIB [12] is categorized as a quality-of-life scale, rather than a measure of *communicative experience.*

Finally, throughout the paper, there are numerous referencing errors, e.g., Kavakci et al., (2022) [4], which is reference #65 in the reference list in Panuccio et al. [1], is cited as reference #66 in the text, or measures are attributed to the wrong author team, e.g., the SS-QOL [developed by Williams et al. (1999) [5]] is attributed to Hilari (2001) ref #299 in Table 2 and to Northcott (2013) ref 298 in Table 3. These errors impede the reader's ability to link ratings with supporting evidence, making verification difficult and reducing confidence in this review's ratings.

The issues in Panuccio et al. (2025) [1], at best, undermine the review's utility as a guide to measurement instrument selection, and, at worst, provide information which could compromise future aphasia research design and clinical outcomes. Low-quality systematic reviews of assessment and outcome measurement instruments can have many negative consequences, including the following:

1. **Producing misleading conclusions** about the psychometric quality of measurement instruments, which may misinform decision-making in healthcare and research.
2. **Hindering the development of effective interventions or treatments** if unreliable and invalid measurement instruments are selected as outcome measures.
3. **Negatively affecting patient care** by impacting aphasia assessment guidelines, which could lead to incorrect diagnoses, poor treatment choices, and worse health outcomes.

Over more than ten years, the Collaboration of Aphasia Trialists' members (>300 across >40 countries) have made painstaking efforts to improve the quality of aphasia research. These include multiple initiatives to improve the quality, efficiency, and global relevance of measurement instruments and practices [22,27–36]. As a collaboration focused on enhancing the quality and reporting of aphasia research, we are compelled to draw attention to the issues in this paper. The authors' endeavour to critically evaluate the quality of available aphasia tests within the framework of a systematic review is highly commendable. However, given the potential impacts outlined above, we recommend that the authors review and revise the manuscript.

Author Contributions: Conceptualization, S.J.W., K.H. and C.B.; writing—original draft preparation, S.J.W., K.H. and C.B.; writing—review and editing, S.J.W., K.H., K.W., M.M., C.P., L.v.E., R.P., S.Z., W.D.H. and C.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding. S.J.W receives fellowship funding from the National Health and Medical Research Council (NHMRC) Australia (2032983).

Disclosure Statement: The authors are members of the Collaboration of Aphasia Trialists and have professional interests in advancing evidence-based aphasia assessment and outcome measurement practices.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Panuccio, F.; Rossi, G.; Di Nuzzo, A.; Ruotolo, I.; Cianfriglia, G.; Simeon, R.; Sellitto, G.; Berardi, A.; Galeoto, G. Quality of Assessment Tools for Aphasia: A Systematic Review. *Brain Sci.* **2025**, *15*, 271. [CrossRef]
2. Mokkink, L.; Terwee, C.; Patrick, D.; Alonso, J.; Stratford, P.; Knol, D.; Bouter, L.; Vet, H.W. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Qual. Life Res.* **2010**, *19*, 539–549. [CrossRef]
3. Mokkink, L.B.; Boers, M.; van der Vleuten, C.P.M.; Bouter, L.M.; Alonso, J.; Patrick, D.L.; de Vet, H.C.W.; Terwee, C.B. COSMIN Risk of Bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: A Delphi study. *BMC Med. Res. Methodol.* **2020**, *20*, 293. [CrossRef] [PubMed]
4. Kavakci, M.; Engin, K.; Melike, T.; Emre, A.; Yasar, E. The inter-rater reliability of the Turkish version of Aphasia Rapid Test for stroke. *Top. Stroke Rehabil.* **2022**, *29*, 272–279. [CrossRef] [PubMed]
5. Williams, L.S.; Weinberger, M.; Harris, L.E.; Clark, D.O.; Biller, J. Development of a stroke-specific quality of life scale. *Stroke* **1999**, *30*, 1362–1369. [CrossRef]
6. Hilari, K.; Byng, S. Measuring quality of life in people with aphasia: The Stroke Specific Quality of Life Scale. *Int. J. Lang. Commun. Disord.* **2001**, *36* (Suppl. S1), 86–91. [CrossRef] [PubMed]
7. Hilari, K.; Byng, S.; Lamping, D.L.; Smith, S.C. Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39): Evaluation of acceptability, reliability, and validity. *Stroke* **2003**, *34*, 1944–1950. [CrossRef]
8. Hilari, K.; Lamping, D.L.; Smith, S.C.; Northcott, S.; Lamb, A.; Marshall, J. Psychometric properties of the Stroke and Aphasia Quality of Life Scale (SAQOL-39) in a generic stroke population. *Clin. Rehabil.* **2009**, *23*, 544–557. [CrossRef]
9. Kamiya, A.; Kamiya, K.; Tatsumi, H.; Suzuki, M.; Horiguchi, S. Japanese adaptation of the Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39): Comparative study among different types of aphasia. *J. Stroke Cerebrovasc. Dis.* **2015**, *24*, 2561–2564. [CrossRef]
10. Hula, W.D.; Doyle, P.J.; Stone, C.A.; Austermann Hula, S.N.; Kellough, S.; Wambaugh, J.L.; Ross, K.B.; Schumacher, J.G.; St Jacque, A. The Aphasia Communication Outcome Measure (ACOM): Dimensionality, item bank calibration, and initial validation. *J. Speech Lang. Hear. Res.* **2015**, *58*, 906–919. [CrossRef]
11. Kim, D.Y.; Pyun, S.-B.; Kim, E.J.; Ryu, B.J.; Choi, T.W.; Pulvermüller, F. Reliability and validity of the Korean version of the Communicative Activity Log (CAL). *Aphasiology* **2016**, *30*, 96–105. [CrossRef]
12. Baylor, C.; Yorkston, K.; Eadie, T.; Kim, J.; Chung, H.; Amtmann, D. The Communicative Participation Item Bank (CPIB): Item bank calibration and development of a disorder-generic short form. *J. Speech Lang. Hear. Res.* **2013**, *56*, 1190–1208. [CrossRef]
13. Koenig-Bruhin, M.; Vanbellingen, T.; Schumacher, R.; Pflugshaupt, T.; Annoni, J.M.; Muri, R.M.; Bohlhalter, S.; Nyffeler, T. Screening for language disorders in stroke: German validation of the Language Screening Test (LAST). *Cerebrovasc. Dis. Extra* **2016**, *6*, 27–31. [CrossRef]
14. El Hachioui, H.; Visch-Brink, E.G.; de Lau, L.M.L.; van de Sandt-Koenderman, M.W.M.E.; Nouwens, F.; Koudstaal, P.J.; Dippel, D.W.J. Screening tests for aphasia in patients with stroke: A systematic review. *J. Neurol.* **2017**, *264*, 211–220. [CrossRef]
15. Salter, K.; Jutai, J.; Foley, N.; Hellings, C.; Teasell, R. Identification of aphasia post stroke: A review of screening assessment tools. *Brain Inj.* **2006**, *20*, 559–568. [CrossRef]
16. Vogel, A.P.; Maruff, P.; Morgan, A.T. Evaluation of communication assessment practices during the acute stages post stroke. *J. Eval. Clin. Pract.* **2010**, *16*, 1183–1188. [CrossRef]
17. Rohde, A.; Worrall, L.; Godecke, E.; O'Halloran, R.; Farrell, A.; Massey, M. Diagnosis of aphasia in stroke populations: A systematic review of language tests. *PLoS ONE* **2018**, *13*, e0194143. [CrossRef] [PubMed]
18. Huber, W.; Poeck, K.; Weniger, D.; Willmes, K. *Der Aachener Aphasie Test (AAT)*; Hogrefe: Göttingen, Germany, 1983.
19. van der Meulen, I.; van de Sandt-Koenderman, W.M.; Duivenvoorden, H.J.; Ribbers, G.M. Measuring verbal and non-verbal communication in aphasia: Reliability, validity, and sensitivity to change of the Scenario Test. *Int. J. Lang. Commun. Disord.* **2010**, *45*, 424–435. [CrossRef] [PubMed]
20. Blomert, L.; Kean, M.L.; Koster, C.; Schokker, J. Amsterdam-Nijmegen Everyday Language Test: Construction, reliability and validity. *Aphasiology* **1994**, *8*, 381–407. [CrossRef]
21. Swinburn, K.; Porter, G.; Howard, D. *Comprehensive Aphasia Test*; The Psychology Press, Taylor and Francis: Abingdon, UK, 2005.
22. Martinez-Ferreiro, S.; Arslan, S.; Fyndanis, V.; Howard, D.; Kraljevic, J.K.; Skoric, A.M.; Munarriz-Ibarrola, A.; Norvik, M.; Penaloza, C.; Pourquie, M.; et al. Guidelines and recommendations for cross-linguistic aphasia assessment: A review of 10 years of comprehensive aphasia test adaptations. *Aphasiology* **2024**, 1–25. [CrossRef]
23. Frattali, C.M.; Thompson, C.M.; Holland, A.L.; Wohl, C.B.; Ferketic, M.M. The FACS of life ASHA facs—A functional outcome measure for adults. *ASHA* **1995**, *37*, 40.

24. Strand, E.A.; Duffy, J.R.; Clark, H.M.; Josephs, K. The Apraxia of Speech Rating Scale: A tool for diagnosis and description of apraxia of speech. *J. Commun. Disord.* **2014**, *51*, 43–50. [CrossRef]
25. Abbey, J.; Piller, N.; De Bellis, A.; Esterman, A.; Parker, D.; Giles, L.; Lowcay, B. The Abbey pain scale: A 1-minute numerical indicator for people with end-stage dementia. *Int. J. Palliat. Nurs.* **2004**, *10*, 6–13. [CrossRef]
26. Casilio, M.; Rising, K.; Beeson, P.M.; Bunton, K.; Wilson, S.M. Auditory-Perceptual Rating of Connected Speech in Aphasia. *Am. J. Speech-Lang. Pathol.* **2019**, *28*, 550–568. [CrossRef]
27. Ali, M.; Basat, A.; Berthier, M.; Blom Johansson, M.; Breitenstein, C.; Cadilhac, D.; Constantinidou, F.; Cruice, M.; Dávila, G.; Gandolfi, M.; et al. Protocol for the development of the international population registry for aphasia after stroke (I-PRAISE). *Aphasiology* **2022**, *36*, 534–554. [CrossRef]
28. Ali, M.; Soroli, E.; Jesus, L.M.T.; Cruice, M.; Isaksen, J.; Visch-Brink, E.; Grohmann, K.K.; Jagoe, C.; Kukkonen, T.; Varlokosta, S.; et al. An aphasia research agenda—A consensus statement from the collaboration of aphasia trialists. *Aphasiology* **2022**, *36*, 555–574. [CrossRef]
29. Arslan, S.; Peñaloza, C. Across countries and cultures: The assessment of aphasia in linguistically diverse clinical populations. *Aphasiology* **2025**, 1–6. [CrossRef]
30. Behn, N.; Harrison, M.; Brady, M.; Breitenstein, C.; Carragher, M.; Fridriksson, J.; Godecke, E.; Hillis, A.; Kelly, H.; Palmer, R.; et al. Developing, monitoring, and reporting of fidelity in aphasia trials: Core recommendations from the collaboration of aphasia trialists (CATs) trials for aphasia panel. *Aphasiology* **2022**, *37*, 1733–1755. [CrossRef]
31. Brady, M.C.; Ali, M.; Fyndanis, C.; Kambanaros, M.; Grohmann, K.K.; Laska, A.-C.; Hernández-Sacristán, C.; Varlokosta, S. Time for a step change? Improving the efficiency, relevance, reliability, validity and transparency of aphasia rehabilitation research through core outcome measures, a common data set and improved reporting criteria. *Aphasiology* **2014**, *28*, 1385–1392. [CrossRef]
32. Breitenstein, C.; Wallace, S.J.; Gilmore, N.; Finch, E.; Pettigrove, K.; Brady, M.C.; Brady, M.C.; Breitenstein, C.; Hilari, K.; Wallace, S.J.; et al. Invaluable benefits of 10 years of the international Collaboration of Aphasia Trialists (CATs). *Stroke* **2024**, *55*, 1129–1135. [CrossRef]
33. Fyndanis, V.; Lind, M.; Varlokosta, S.; Kambanaros, M.; Soroli, E.; Ceder, K.; Grohmann, K.K.; Rofes, A.; Simonsen, H.G.; Bjekić, J.; et al. Cross-linguistic adaptations of The Comprehensive Aphasia Test: Challenges and solutions. *Clin. Linguist. Phon.* **2017**, *31*, 697–710. [CrossRef] [PubMed]
34. Wallace, S.J.; Isaacs, M.; Ali, M.; Brady, M.C. Establishing reporting standards for participant characteristics in post-stroke aphasia research: An international e-Delphi exercise and consensus meeting. *Clin. Rehabil.* **2023**, *37*, 199–214. [CrossRef]
35. Wallace, S.J.; Worrall, L.; Rose, T.; Le Dorze, G.; Breitenstein, C.; Hilari, K.; Babbitt, E.; Bose, A.; Brady, M.; Cherney, L.R.; et al. A core outcome set for aphasia treatment research: The ROMA consensus statement. *Int. J. Stroke* **2019**, *14*, 180–185. [CrossRef] [PubMed]
36. Wallace, S.J.; Worrall, L.; Rose, T.A.; Alyahya, R.S.W.; Babbitt, E.; Beeke, S.; de Beer, C.; Bose, A.; Bowen, A.; Brady, M.C.; et al. Measuring communication as a core outcome in aphasia trials: Results of the ROMA-2 international core outcome set development meeting. *Int. J. Lang. Commun. Disord.* **2023**, *58*, 1017–1028. [CrossRef] [PubMed]