



Validation of the reading the mind in the eyes test requires an interpretable factor model

Wendy C. Higgins^{a,1} , Victoria Savalei^b, Vince Polito^a , and Robert M. Ross^c

In a recent study of the Reading the Mind in the Eyes Test (Eyes Test), Greenberg et al. (1) claim to have provided “robust validation” for the test’s psychometric properties, supported by “acceptable-to-good reliability of the Eyes Test across datasets and countries” (p. 8). However, we have two critical concerns.

First, Greenberg et al. (p. 3) state that the reliability of the Eyes Test was “acceptable-to-good” based on both coefficient omega total (ω_t ; range: 0.74 to 0.80) and coefficient omega hierarchical (ω_h ; range: 0.32 to 0.50). However, this is incorrect. The ω_h values are substantially lower than acceptable (i.e., $\omega_h \ll 0.80$; 2, 3). Both ω_h and ω_t are based on bifactor modelling. Bifactor models comprise a general factor that influences all items (which should capture the target construct) and several group factors, which are independent of the general factor and typically represent construct-irrelevant variance. While ω_h captures the proportion of variance in the test scores due to the general factor only (i.e., variance related to the target construct), ω_t is based on variance due to both the general factor and the (usually construct-irrelevant) group factors. Critically, when these two estimates differ substantially, ω_h better represents the extent to which the sum scores reflect the target construct (4–6). In addition, only ω_t and not ω_h values were reported for individual countries, making it impossible to identify which, if any, countries had acceptable ω_h values.

Second, Greenberg et al. provide no details of the bifactor model used to calculate ω_t and ω_h . Details of the fitted model, including the estimated solution, need to be reported because acceptable omega estimates can hide poor solutions, such as uninterpretable loading patterns, common in these models (6). However, even basic properties of the model, such as the number and composition of group factors (and whether they were the same across countries), were not reported. Critically, given the low values of ω_h relative to ω_t , group factors may

have contributed substantially to the reported effects. Details of the bifactor model are necessary to evaluate the extent to which reported effects (e.g., sex differences) are due to the target construct of interest (theory of mind) rather than group factors. Further, when comparing performance across samples (e.g., by sex), the same bifactor model should be fit to each sample to establish that the number of group factors and the indicators they comprise are the same (i.e., configural invariance) (6). Higher levels of measurement invariance (e.g., metric and scalar) would need to be established to claim that differences in performance are due to the target construct rather than, for example, differing strengths of group factors across populations (7).

Given serious concerns about the Eyes Test as a measure of theory of mind expressed here and elsewhere (8–10), we strongly caution against interpreting Greenberg et al.’s study as providing “robust validation for the psychometric properties of the Eyes Test” (p. 8) or evidence for sex differences in theory of mind ability.

ACKNOWLEDGMENTS. W.C.H. is funded by an Australian Government Research Training Program Scholarship and a Macquarie University Research Excellence Scholarship, R.M.R. is funded by a Macquarie University Research Fellowship and the John Templeton Foundation (Grant ID: 62631), and V.P. is funded by a Macquarie University Research Fellowship.

Author affiliations: ^aSchool of Psychological Sciences, Macquarie University, North Ryde, NSW 2109, Australia; ^bDepartment of Psychology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada; and ^cDepartment of Philosophy, Macquarie University, North Ryde, NSW 2109, Australia

Author contributions: W.C.H., V.S., V.P., and R.M.R. wrote the paper.

The authors declare no competing interest.

Copyright © 2023 the Author(s). Published by PNAS. This article is distributed under [Creative Commons Attribution License 4.0 \(CC BY\)](https://creativecommons.org/licenses/by/4.0/).

¹To whom correspondence may be addressed. Email: wendy.higgins@mq.edu.au.

Published December 18, 2023.

1. D. M. Greenberg et al., Sex and age differences in “theory of mind” across 57 countries using the English version of the “Reading the Mind in the Eyes” Test. *Proc. Natl. Acad. Sci. U.S.A.* **120**, e2022385119 (2023).
2. M. K. Forbes et al., Three recommendations based on a comparison of the reliability and validity of the predominant models used in research on the empirical structure of psychopathology. *J. Abnorm. Psychol.* **130**, 297–317 (2021).
3. A. Rodriguez, S. P. Reise, M. G. Haviland, Applying bifactor statistical indices in the evaluation of psychological measures. *J. Pers. Assess.* **98**, 223–237 (2016).
4. D. B. Flora, Your coefficient alpha is probably wrong, but which coefficient omega is right? A tutorial on using R to obtain better reliability estimates. *Adv. Meth. Pract. Psychol. Sci.* **3**, 484–501 (2020).
5. V. Savalei, S. P. Reise, Don’t forget the model in your model-based reliability coefficients: A reply to Mcneish (2018). *Collabra Psychol.* **5**, 36 (2019).
6. M. A. Bornovalova, A. M. Choate, H. Fatimah, K. J. Petersen, B. M. Wiernik, Appropriate use of bifactor analysis in psychopathology research: Appreciating benefits and limitations. *Biol. Psychiatry*, **88**, 18–27 (2020).
7. R. B. Kline, *Principles and Practice of Structural Equation Modelling* (Guilford Press, ed. 5, 2023).
8. W. C. Higgins, R. M. Ross, R. Langdon, V. Polito, The “Reading the Mind in the Eyes” Test shows poor psychometric properties in a large, demographically representative U.S. sample. *Assessment* **30**, 1777–1789 (2023).
9. W. C. Higgins, R. M. Ross, V. Polito, D. M. Kaplan, Three threats to the validity of the Reading the Mind in the Eyes Test: A commentary on Pavolova and Sokolov (2022). *Neurosci. Biobehav. Rev.* **147**, 105088 (2023).
10. A. F. D. Kittel, S. Olderbak, O. Wilhelm, Sty in the mind’s eye: A meta-analytic investigation of the nomological network and internal consistency of the “Reading the Mind in the Eyes” test. *Assessment* **29**, 872–895 (2022).