



Editorial

Stephen C. Bowden¹ · David W. Loring²

Received: 12 July 2023 / Accepted: 12 July 2023 / Published online: 18 August 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

As graduate students, we were taught that unusual test score variability was a sign of malingering. While we were quite comfortable making that diagnostic inference, since the advice came from respected mentors, when we re-evaluated these patients at a later time, we found that some had neuropsychological profiles reflecting clear decline associated with disease progression. Since those early faltering steps in our personal careers, the study of performance validity tests (PVTs) has provided multiple objective approaches to characterize patient task engagement during neuropsychological evaluation, although these too have been associated with methodological limitations.

Assessing performance validity as a routine component of a neuropsychological evaluation is now a recommended practice standard (e.g., Heilbronner et al., 2009; Sherman et al., 2020; Sweet et al., 2021). Since the initial descriptions of performance validity tests (for review see Frederick, 2003), there has been remarkable maturation of research methods used to establish the validity of validity tests themselves. Unfortunately, this field has also been the subject of controversy and debate, not only with regard to specific tests and assessment techniques, but more importantly with the definitional challenges of “validity” in the absence of any clearly established and well-performing external criteria or clinical biomarkers to serve as a “gold standard.” The absence of a gold-standard poses special challenges for performance validity validation.

In this issue of *NEUROPSYCHOLOGY REVIEW*, Dr. Christoph Leonard addresses a range of issues regarding the validation and clinical interpretation of performance

validity assessments. In Part I, Dr. Leonhard discusses (a) the importance of reporting full statistics on classification accuracy rather than point estimates of sensitivity and specificity alone, (b) the challenges of interpretation of two or more PVTs administered to a patient, and (c) conceptual and empirical issues in interpretation of assessment algorithms that include different sources of information. In Part II, Dr. Leonard addresses study design limitations in the validation of PVTs, highlighting (a) the risk-of-bias issues in some published studies, (b) the distinction between convergent construct validation versus criterion-referenced validation, and (c) sample selection and the impact of exclusion of too-close-to-call cases in reporting diagnostic accuracy. He concludes with a discussion of risk-of-bias evaluation according to the QUADAS criteria.

Because of the importance of the methodological limitations identified by Dr. Leonhard, we have sought commentary from three senior clinicians or researchers who have addressed the area of PVT characterization and use or test validation, namely, Drs. Shane Bush, David Faust, and Paul Jewsbury.

Dr. Bush highlights the history of professional, legal, and ethical issues associated with the PVT application. He describes the long-standing contentions associated with the use of some PVT testing practices and the transition from more subjective techniques to more empirically based approaches, and he highlights the continuing need to address important professional issues.

Dr. Faust notes the serious methodological challenges underlying the effective estimation of accuracy rates derived from PVTs, and in particular, the potential bias associated with excluding cases judged “too-close-to-call,” redundancy of measures in validation studies, and the impact of clinician factors. Dr Faust discusses priorities for further research in validity evaluations, and concludes with discussion of criterion group composition and the impact that selection procedures may have on validity evaluation.

Dr. Jewsbury addresses critical statistical issues related to the optimal interpretation of multiple PVTs derived from a single assessment, and notes continuing discussion

✉ Stephen C. Bowden
sbowden@unimelb.edu.au

David W. Loring
dloring@emory.edu

¹ Melbourne School of Psychological Sciences, The University of Melbourne, 3010 Melbourne, VIC, Australia

² Department of Neurology, Emory University School of Medicine, Atlanta, GA, USA

in the literature regarding the best methods of interpretation of multiple results. This issue is particularly relevant since consensus guidelines recommend administration and algorithmic interpretation of multiple PVTs for all patients. There are important confounds that arise from PVT interpretation when statistical independence of PVTs cannot be assumed (the usual circumstance), and the critical distinction between conditional independence versus unconditional independence for multiple PVTs is emphasized. Conditional independence refers to multiple test results being independent from each other (uncorrelated) and established separately for both the patient and control groups. Unconditional independence refers to the correlation between multiple test results in the total population (combined patient and control groups).

We expect that Dr. Leonhard's critiques and our invited commentaries will foster thoughtful debate and discussion about the current state of PVT use in neuropsychology, which we regard as an important vehicle for improving evidence-based PVT practices. As Dr. Faust observed, we have come a long way in refining PVT research methods, but we have not reached the end of the road.

References

- Frederick, R. I. (2003). A review of Rey's strategies for detecting malingered neuropsychological impairment. *Journal of Forensic Neuropsychology*, 2, 3–4. https://doi.org/10.1300/J151v02n03_01
- Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., Millis, S. R., & Participants, C. 1 (2009). American academy of clinical neuropsychology consensus conference statement on the neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist*, 23(7), 1093–1129. <https://doi.org/10.1080/13854040903155063>
- Sherman, E. M. S., Slick, D. J., & Iverson, G. L. (2020). Multidimensional malingering criteria for neuropsychological assessment: A 20-year update of the malingered neuropsychological dysfunction criteria. *Archives of Clinical Neuropsychology*, acaa019. <https://doi.org/10.1093/arclin/acaa019>
- Sweet, J. J., Heilbronner, R. L., Morgan, J. E., Larrabee, G. J., Rohling, M. L., Boone, K. B., Kirkwood, M. W., Schroeder, R. W., & Suhr, J. A. (2021). American Academy of Clinical Neuropsychology (AACN) 2021 consensus statement on validity assessment: Update of the 2009 AACN consensus conference statement on neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist*, 35(6), 1053–1106. <https://doi.org/10.1080/13854046.2021.1896036>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.