Psychological Injury and Law (2021) 14:227–231
https://doi.org/10.1007/s12207-021-09425-x

Future Directions in Performance Validity Assessment to Optimize Detection of Invalid Neuropsychological Test Performance: Special Issue Introduction

Jason R. Soble[1]

Received: 23 August 2021 / Accepted: 13 September 2021 / Published online: 22 September 2021

This special issue on performance validity assessment in neuropsychological testing stems from a long tradition in the field, although the research has accelerated only recently. The term performance validity tests (PVTs) was introduced by Larrabee (2012) to distinguish testing in neuropsychological assessment aimed at determining extent of cognitive test validity from symptom validity testing (SVTs), such as through scales in self-report instruments (e.g., the F tests in the MMPI-2-RF; Ben-Porath & Tellegen, 2008; Ben-Porath, 2012), aimed at determining extent of symptom plausibility and exaggeration.

History of PVTs

In these respects, the concept of performance validity assessment is not entirely novel considering that the earliest known performance validity test equivalents were developed 60–80 years ago by Andre Rey (see Frederick, 2003; Greiffenstein et al., 1996), including the Dot Counting Test, Word Recognition Test, and 15-Item Test (Rey, 1941, 1964). Notwithstanding, research examining PVTs in the context of neuropsychological evaluation was largely dormant over the ensuing 40–50 years such that a virtually non-existent literature base existed as recently as the 1980s (Boone, 2007). By contrast, the 1990s ushered in a greatly renewed interest in performance validity assessment within the field of clinical neuropsychology along with burgeoning research on validating/cross-validating various PVTs (e.g., more than 300 publications on PVTs from 1990 to 2007; Boone, 2007). Indeed, it was during this time that some of the most commonly-administered and well-known PVTs were published, such as the Word Memory Test (Green et al., 1996), Test of Memory Malingering (Tombaugh, 1996), Victoria Symptom Validity Test (Slick et al., 1997), and Reliable Digit Span (Greiffenstein et al., 1994). In addition, standardized criteria for identifying malingered neurocognitive dysfunction (Slick et al., 1999) were introduced and included central roles for performance validity testing equivalents (the term had yet to be created at that time).

Despite these seminal advances, at the onset of the 2000s, several key limitations persisted in the PVT literature base. Chief among these were the near-exclusive emphasis on forced choice measures as PVTs, such as with the TOMM, overreliance of forensic/medicolegal cross-validation samples, and the largely synonymous linkage of PVT failure and malingering (Boone, 2007).

Current State of PVT Research

Building on the research of the 1990s, the 2000s saw both further rapid growth of the PVT literature base. Indeed, from 2007 to 2015, more than 1400 publications on the topic of performance validity assessment have been introduced to the literature (Boone, 2021; Martin et al., 2015). Several factors have contributed to this burgeoning research on PVTs, including codification of formal practice standards for validity assessment published by the major professional organizations in the field, including the National Academy Neuropsychology (Bush et al., 2005) and the American Academy of Clinical Neuropsychology (Heilbronner et al., 2009; Sweet et al., 2021) as well as revised structured criteria for identifying non-credible neuropsychological test performance (Sherman et al., 2020). Greater appreciation of base rates of performance invalidity in non-forensic clinical samples (e.g., Martin & Schroeder, 2020) further resulted in a more nuanced understanding of the importance of integrating

✉ Jason R. Soble
soble@uic.edu

[1] Department of Psychiatry and Department of Neurology, University of Illinois At Chicago College of Medicine, Chicago, USA

objective performance validity assessment in *all* neuropsychological evaluations, not just forensic/medicolegal exams. Lastly, an increasing number of graduate training programs offering dual degrees in psychology and law have increased the number of early career professionals advocating for high-quality PVT research that translates to and informs evidence-based forensic case work.

Several key empirical findings also have emerged from the rapidly expanding PVT literature base over the past 25 years and currently allow for more precise, refined, and evidence-based assessment of performance validity in neuropsychological evaluations. Among these are establishment of clear benchmarks for classification of invalid neuropsychological test performance (i.e., failure on ≥ 2 independent PVTs; Boone, 2013; Critchfield et al., 2019; Jennette et al., 2021; Larrabee, 2008; Meyers et al., 2014; Rhoads et al., 2021b; Sherman et al., 2020; Webber et al., 2020), elucidation of best practices for validity assessment that includes continuous sampling of validity via administration of multiple freestanding and embedded PVTs throughout neuropsychological evaluations (Boone, 2009; Sweet et al., 2021), and greater empirical support to inform critical clinical decisions related to validity assessment, such as the number and type(s) of PVTs administered (Soble et al., 2020). Perhaps most importantly, the extant PVT literature has continued to firmly establish the psychometric properties and effectiveness of many freestanding and embedded PVTs for detecting performance invalidity across diverse clinical populations with and without cognitive impairment (see Soble et al., 2021b).

Special Issue Focus: Future Directions for PVT Research

While significant empirical strides pertaining to performance validity assessment have been made over the past 2.5 decades, the PVT science must continually evolve to meet the changing needs of the field and larger sociopolitical factors. Accordingly, future directions for performance validity research is the common theme underlying many of the articles featured in this special issue, with each article highlighting one or more aspects of PVT research that should continue to progress. For instance, Ovsiew et al. (2021) (featured in *Psychological Injury and Law* 14(2) demonstrated that abbreviated versions of the TOMM, particularly Trial 1, evidence classification accuracy and psychometric properties that mirror the traditional two-trial administration, but with the advantage of half the administration time. Studies such as this allow for advancement of PVT science in a manner consistent with current healthcare trends emphasizing cost-containment and shorter, more focused evaluations that minimize patient burden and associated costs. The research in this regard continues to develop more effective embedded PVTs, such as indices derived from common cognitive tests, including the Stroop (White et al., 2020a), Rey Auditory Verbal Learning Test (Pliskin et al., 2020; Soble et al., 2021a), Hopkins Verbal Learning Test-Revised (Bailey et al., 2018); Brief Visuospatial Memory Test-Revised (Bailey et al., 2018; Resch et al., 2020), Digit Span (Schroeder et al., 2012; Webber & Soble, 2018), California Verbal Learning Test (Schwartz et al., 2016), and Repeatable Battery for the Assessment of Neuropsychological Status (Shura et al., 2018).

Moreover, the applicability and utility of various validity measures must continue to be cross-validated in diverse medical and neuropsychiatric populations. Two of the articles featured in this issue, Modiano et al. (2021) and Tierney et al. (2021), highlight this research principle well. Notably, Modiano et al. (2021) demonstrated that the Amnestic Disorders Scale of the Structured Inventory of Malingered Symptomatology (Widows & Smith, 2005) had excellent classification accuracy for detecting invalid cognitive symptom reporting irrespective of the presence of actual cognitive impairment, whereas Tierney et al. (2021) provided preliminary evidence that the Miller Forensic Assessment of Symptoms Test (M-FAST; Miller, 2001) accurately identifies invalid symptom reporting among neurological patients admitted for inpatient epilepsy monitoring/workup.

In a related vein, the relationship between performance validity and symptom validity must continue to be clarified across diverse clinical populations. Notably, it is established that symptom validity and performance validity are separate constructs with varying degrees of interrelatedness depending on the clinical population (Gervais et al., 2007; Larrabee, 2012; Leib et al., 2021; White et al., 2020a, 2020b). In this issue, Shura et al. (2021) further expanded the current understanding of how symptom and performance validity are dissociable in veteran populations with mild traumatic brain injury and posttraumatic stress disorder.

PVT research must evolve to meet the changing demographics of the USA as well as increasing applicability among international samples by establishing the accuracy and cross-validating PVTs in diverse racial/ethnic groups. In an earlier article in this journal, Bailey and colleagues (2021) published novel cross validation findings of the TOMM in a large Colombian sample and identified several relevant demographic factors (e.g., age, education) that may affect performance on this test. More recently, Rhoads et al. (2021a) further highlighted some potential limitations of using PVT cut-scores derived from English-speaking populations among Spanish-speaking patients residing in the USA and emphasized the need for more extensive cross-validation of various PVTs in non-English-speaking populations. Undoubtedly, PVT research in diverse and/or non-English-speaking

populations is currently its early stages and remains a fertile area for future empirical investigation.

Finally, although beyond the focus of the specific articles included in this special issue, some additional emerging areas of future PVT research are noteworthy. Future PVT research should continue to capitalize on meta-analytic and systematic review methodologies (e.g., Bernstein et al., 2021; Martin et al., 2020; Resch et al., 2021) to enhance findings from single cross-validation studies and make use of more advanced methodological (e.g., machine learning) and/or statistical approaches (e.g., measurement invariance) to enhance their utility and applicability across a wider range of populations. Additional research on validity testing via computer-based and telehealth modalities (e.g., O'Rourke et al., under review) also will be critical considering how the COVID-19 pandemic has resulted in opportunities for change in psychological/neuropsychological assessment practices. Future research examining the relationship and concordance of PVT performance with neuroimaging or other techniques assessing neural activation also may yield fruitful results.

The past 25 years have resulted in a robust literature base supporting the effectiveness and accuracy of PVTs for detecting invalid neuropsychological test performance across medicolegal, clinical, and research settings and have provided clinical neuropsychologists with a wealth of free-standing and embedded measures at their disposal. However, the practice and science of performance validity assessment must continue to develop in order to meet the demands of changing demographics and healthcare factors. To this end, it is the hope that many of the articles contained in this special issue provide steps and ideas for future PVT research.

References

Bailey, K. C., Goatte, W., Ramos-Usuga, D., Rivera, D., & Arango-Lasprilla, J. C. (2021). Cross-validation of the utility of Test of Memory Malingering (TOMM) cut-offs in a large Colombian sample. *Psychological Injury and Law,* 14, 114–126.

Bailey, K. C., Soble, J. R., Bain, K. M., & Fullen, C. (2018). Embedded performance validity tests in the Hopkins Verbal Learning Test-Revised and the Brief Visuospatial Memory Test-Revised: A replication study. *Archives of Clinical Neuropsychology,* 33(7), 895–900. https://doi.org/10.1093/arclin/acx111

Ben-Porath, Y. S. (2012). *Interpreting the MMPI-2-RF*. University of Minnesota Press.

Ben-Porath, Y. S., & Tellegen, A. (2008). *MMPI-2-RF: Manual for administration, scoring, and interpretation*. University of Minnesota Press.

Bernstein, M. T., Resch, Z. J., Ovsiew, G. P., & Soble, J. R. (2021). A systematic review and meta-analysis of the diagnostic accuracy of the Advanced Clinical Solutions Word Choice Test as a performance validity test. *Neuropsychology Review,* 31(2), 349–359.

Boone, K. B. (2007). *Assessment of feigned cognitive impairment: A neuropsychological perspective*. Guilford Press.

Boone, K. B. (2009). The need for continuous and comprehensive sampling of effort/response bias during neuropsychological examinations. *The Clinical Neuropsychologist,* 23(4), 729–741. https://doi.org/10.1080/13854040802427803

Boone, K. B. (2013). *Clinical practice of forensic neuropsychology: An evidence-based approach*. Guilford Press.

Boone, K. B. (2021). *Assessment of feigned cognitive impairment: A neuropsychological perspective* (2nd ed.). Guilford Press.

Bush, S. S., Ruff, R. M., Troster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H., & Silver, C. H. (2005). Symptom validity assessment: Practice issues and medical necessity: NAN Policy & Planning Committee. *Archives of Clinical Neuropsychology,* 20(4), 419–426.

Critchfield, E., Soble, J. R., Marceaux, J. C., Bain, K. M., Bailey, K. C., Webber, T. A., Alverson, W. A., Messerly, J., González, D. A., & O'Rourke, J. J. F. (2019). Cognitive impairment does not cause performance validity failure: Analyzing performance patterns among unimpaired, impaired, and noncredible participants across six tests. *The Clinical Neuropsychologist,* 33(6), 1083–1101. https://doi.org/10.1080/13854046.2018.1508615

Frederick, R. I. (2003). A review of Rey's strategies for detecting malingered neuropsychological impairment. *Journal of Forensic Neuropsychology,* 3–4, 1–25.

Gervais, R. O., Ben-Porath, Y. S., Wygant, D. B., & Green, P. (2007). Development and validation of a Response Bias Scale (RBS) for the MMPI-2. *Assessment,* 14(2), 196–208. https://doi.org/10.1177/1073191106295861

Green, P., Allen, L., & Astner, K. (1996). *The Word Memory Test: A user's guide to the oral and computer administered forms, US version 1.1*. Durham, NC: Cognisyst.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures in a large clinical sample. *Psychological Assessment,* 6, 218–224.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1996). Comparison of multiple scoring methods for Rey's malingered amnesia measures. *Archives of Clinical Neuropsychology,* 11, 283–293.

Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., & Millis, S. R. (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.

Jennette, K. J., Williams, C. P., Resch, Z. J., Ovsiew, G. P., Durkin, N. M., O'Rourke, J. J. F., Marceaux, J. C., Critchfield, E. C., & Soble, J. R. (2021). Assessment of differential neurocognitive performance based on the number of performance validity tests failures: A cross-validation study across multiple mixed clinical samples. *The Clinical Neuropsychologist*. Advance online publication. https://doi.org/10.1080/13854046.2021.1900398

Larrabee, G. J. (2008). Aggregation across multiple indicators improves the detection of malingering: Relationship to likelihood ratios. *The Clinical Neuropsychologist,* 22(4), 666–679. https://doi.org/10.1080/13854040701494987

Larrabee, G. J. (2012). Performance validity and symptom validity in neuropsychological assessment. *Journal of the International Neuropsychological Society,* 18(4), 625–630. https://doi.org/10.1017/s1355617712000240

Leib, S. I., Schieszler-Ockrassa, C. White, D. J., Gallagher, V. T., Carter, D. A., Basurto, K. S., Ovsiew, G. P., Resch, Z. J., Jennette, K. J., & Soble, J. R. (2021). Concordance between the Minnesota Multiphasic Personality Inventory-2-Restructured Form (MMPI-2-RF) and Clinical Assessment of Attention Deficit-Adult (CAT-A) over-reporting validity scales for detecting invalid ADHD symptom reporting. *Applied Neuropsychology: Adult*. Advance online publication. https://doi.org/10.1080/23279095.2021.1894150

Martin, P. K., & Schroeder, R. W. (2020). Base rates of invalid test performance across clinical non-forensic contexts and settings.

Archives of Clinical Neuropsychology, 35(6), 717–725. https://doi.org/10.1093/arclin/acaa017

Martin, P. K., Schroeder, R. W., & Odland, A. P. (2015). Neuropsychologists' validity testing beliefs and practices: A survey of North American professionals. *The Clinical Neuropsychologist,* 29(6), 741–776.

Martin, P. K., Schroeder, R. W., Olsen, D. H., Maloy, H., Boettcher, A., Ernst, N., & Okut, H. (2020). A systematic review and meta-analysis of the Test of Memory Malingering in adults: Two decades of deception detection. *The Clinical Neuropsychologist,* 34(1), 88–119.

Meyers, J. E., Miller, R. M., Thompson, L. M., Scalese, A. M., Allred, B. C., Rupp, Z. W., Dupaix, Z. P., & Junghyun Lee, A. (2014). Using likelihood ratios to detect invalid performance with performance validity measures. *Archives of Clinical Neuropsychology,* 29(3), 224–235. https://doi.org/10.1093/arclin/acu001

Miller, H. A. (2001). *Miller Forensic Assessment of Symptoms Test (M-FAST): Professional manual*. Psychological Assessment Resources.

Modiano, Y. A., Taiwo, Z., Pastorek, N. J., & Webber, T. A. (2021). *The Structured Inventory of Malingered Symptomatology Amnestic Disorders Scale (SIMS-AM) is insensitive to cognitive impairment while accurately identifying invalid cognitive symptom reporting*. Advance online publication. https://doi.org/10.1007/s12207-021-09420-2

Ovsiew, G. P., Carter, D. A., Rhoads, T., Resch, Z. J., Jennette, K. J., & Soble, J. R. (2021). Concordance between standard and abbreviated administrations of the Test of Memory Malingering: Implications for streamlining performance validity assessment. *Psychological Injury and Law,* 14(2), 134–143.

Pliskin, J. I., DeDios Stern, S., Resch, Z. J., Saladino, K. F., Ovsiew, G. P., Carter, D. A., Soble, J. R. (2020). Comparing the psychometric properties of 8 embedded performance validity tests in the Rey Auditory Verbal Learning Test, Wechsler Memory Scale Logical Memory, and Brief Visuospatial Memory Test–Revised recognition trials for detecting invalid neuropsychological test performance. *Assessment*. Advance online publication. https://doi.org/10.1177/1073191120929093

Resch, Z. J., Pham, A. T., Abramson, D. A., White, D. J., DeDios-Stern, S., Ovsiew, G. P., Castillo, L., & Soble, J. R. (2020a). Examining independent and combined accuracy of embedded performance validity tests in the California Verbal Learning Test-II and Brief Visuospatial Memory Test-Revised for detecting invalid performance. *Applied Neuropsychology: Adult*. https://doi.org/10.1080/23279095.2020.1742718

Resch, Z. J., Webber, T. A., Bernstein, M. T., Rhoads, T., Ovsiew, G. P., & Soble, J. R. (2021). Victoria Symptom Validity Test: A systematic review and cross-validation study. *Neuropsychology Review,* 31(2), 331–348.

Rey, A. (1941). L'examen psychologique dans les cas d'encéphalopathie traumatique [The psychological examination in cases of traumatic encephalopathy]. *Archives de Psychologie,* 28, 286–340.

Rey, A. (1964). *L'examen Clinique en psychologie [The clinical examination in psychology]*. Presses Universitaires de France.

Rhoads, T. Leib, S. I., Resch, Z. J., Basurto, K., Castillo, L. R., Jennette, K. J., & Soble, J. R. (2021a). Relative base rates of invalidity for the Test of Memory Malingering and the Dot Counting Test among Spanish-speaking patients residing in the United States. *Psychology Injury and Law*.

Rhoads, T., Neale, A. C., Resch, Z. J., Cohen, C. D., Keezer, R. D., Cerny, B. M., Jennette, K. J., Ovsiew, G. P., & Soble, J. R. (2021b). Psychometric implications of failure on one performance validity test: A cross-validation study to inform criterion group definition. *Journal of Clinical and Experimental Neuropsychology,* 43(5), 437–448. https://doi.org/10.1080/13803395.2021.1945540

Schroeder, R. W., Twumasi-Ankrah, P., Baade, L. E., & Marshall, P. S. (2012). Reliable Digit Span: A systematic review and cross-validation study. *Assessment,* 19(1), 21–30.

Schwartz, E. S., Erdodi, L., Rodriguez, N., Ghosh, J. J., Curtain, J. R., Flashman, L. A., & Roth, R. M. (2016). CVLT-II Forced Choice Recognition Trial as an embedded validity indicator: A systematic review of the evidence. *Journal of the International Neuropsychological Society, 22*(8), 851–858. https://doi.org/10.1017/S1355617716000746

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,* 35(6), 735–764. https://doi.org/10.1093/arclin/acaa019

Shura, R. D., Brearly, T. W., Rowland, J. A., Martindale, S. L., Miskey, H. M., & Duff, K. (2018). RBANS validity indices: A systematic review and meta-analysis. *Neuropsychology Review,* 28(3), 269–284.

Shura, R. D., Yoash-Gantz, R. E., Pickett, T. C., McDonald, S. D., & Tupler, L. A. (2021). *Relations among performance and symptom validity, mild traumatic brain injury, and posttraumatic stress disorder symptom burden in postdeployment veterans*. Advance online publication. https://doi.org/10.1007/s12207-021-09415-z

Slick, D. J., Hopp, G., Strauss, E., & Thompson, G. B. (1997). *Victoria Symptom Validity Test: Professional manual*. Psychological Assessment Resources.

Slick, D. J., Sherman, E. M., & Iverson, G. L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist,* 13(4), 545–561.

Soble, J. R., Alverson, W. A., Phillips, J. I., Critchfield, E. A., Fullen, C., O'Rourke, J. J. F., Messerly, J., Highsmith, J. M., Bailey, K. C., Webber, T. A., & Marceaux, J. M. (2020). Strength in numbers or quality over quantity? Examining the importance of criterion measure selection to define validity groups in performance validity test (PVT) research. *Psychological Injury and Law,* 13, 44–56. https://doi.org/10.1007/s12207-019-09370-w

Soble, J. R., Sharp, D. W., Carter, D. A., Jennette, K. J., Resch, Z. J., Ovsiew, G. P., & Critchfield, E. A. (2021a). Cross-validation of a forced-choice validity indicator to enhance the clinical utility of the Rey Auditory Verbal Learning Test. *Psychological Assessment,* 33(6), 568–573.

Soble, J. R., Webber, T. A., & Bailey, K. C. (2021b). An overview of common stand-alone and embedded PVTs for the practicing clinician: Cutoffs, classification accuracy, and administration times. In R. W. Schroeder & P. K. Martin (Eds.), *Validity assessment in clinical neuropsychological practice: Evaluating and managing noncredible performance* (pp. 126–149). Guilford.

Sweet, J. J., Heilbronner, R. L., Morgan, J. E., Larrabee, G. J., Rohling, M. L., Boone, K. B., et al. (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*. Advance online publication. https://doi.org/10.1080/13854046.2021.1896036

Tierney, S. M., Webber, T. A., Collins, R. L., Pacheco, V. H., & Grayban, J. M. (2021). *Validity and utility of the Miller Forensic Assessment of Symptoms Test (M-FAST) on an inpatient epilepsy monitoring unit*. Advance online publication. https://doi.org/10.1007/s12207-021-09418-w

Tombaugh, T. N. (1996). *Test of Memory Malingering (TOMM)*. North Tonawanda: Multi Health Systems.

Webber, T. A., & Soble, J. R. (2018). Utility of various WAIS-IV Digit Span indices for identifying noncredible performance validity among cognitively impaired and unimpaired examinees. *The*

Clinical Neuropsychologist, 32(4), 657–670. https://doi.org/10.1080/13854046.2017.1415374

Webber, T. A., Critchfield, E. A., & Soble, J. R. (2020). Convergent, discriminant, and concurrent validity of non-memory-based performance validity tests. *Assessment,* 27(7), 1399–1415. https://doi.org/10.1177/1073191118804874

White, D. J., Korinek, D., Bernstein, M. T., Ovsiew, G. P., Resch, Z. J., & Soble, J. R. (2020a). Cross-validation of non-memory-based embedded performance validity tests for detecting invalid performance among patients with and without neurocognitive impairment. *Journal of Clinical and Experimental Neuropsychology,* 42, 459–472.

White, D. J., Ovsiew, G. P., Rhoads, T., Resch, Z. J., Lee, M., Oh, A. J., & Soble, J. R. (2020b). The divergent roles of symptom and performance validity in the assessment of ADHD. *Journal of Attention Disorders*. Advance online publication. https://doi.org/10.1177/1087054720964575

Widows, M. R., & Smith, G. P. (2005). *Structured Inventory of Malingered Symptomatology: Professional manual*. Psychological Assessment Resources.