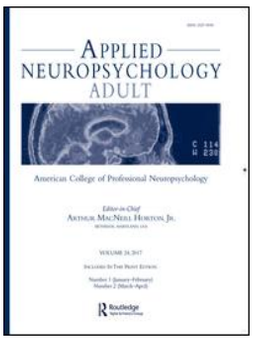

Applied Neuropsychology: Adult

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hapn21

Valid or not: A critique of Graver and Green

David W. Loring , Kimford J. Meador & Felicia C. Goldstein

To cite this article: David W. Loring , Kimford J. Meador & Felicia C. Goldstein (2020): Valid or not: A critique of Graver and Green, Applied Neuropsychology: Adult

To link to this article: https://doi.org/10.1080/23279095.2020.1798961

Published online: 31 Jul 2020.

APPLIED NEUROPSYCHOLOGY: ADULT
https://doi.org/10.1080/23279095.2020.1798961

Valid or not: A critique of Graver and Green

David W. Loring[a,b], Kimford J. Meador[c], and Felicia C. Goldstein[a]

[a]Department of Neurology, Emory University School of Medicine, Atlanta, GA, USA; [b]Department of Pediatrics, Emory University School of Medicine, Atlanta, GA, USA; [c]Department of Neurology & Neurological Sciences, Stanford University School of Medicine, Stanford, CA, USA

ABSTRACT

Disagreements in science and medicine are not uncommon, and formal exchanges of disagreements serve a variety of valuable roles. As identified by a *Nature Methods* editorial entitled "The Power of Disagreement" (2016), disagreements bring attention to best practices so that differences in interpretation do not result from inferior data sets or confirmation bias, "prompting researchers to take a second look at evidence that is not in agreement with their hypothesis, rather than dismiss it as artifacts." Graver and Green published reasons why they disagree with a recent clinical case report and a decades old randomized control trial characterizing the effect of an acute 2 mg dosing of lorazepam on the Word Memory Test. In this article, we formally responded to their commentary to further clarify the reasons for our data interpretations. These two opposing views provide an excellent learning opportunity, particularly for students, demonstrating the importance of careful articulation of the rationale behind certain conclusions from different perspectives. We encourage careful review of the original articles being discussed so the neuropsychologists can read both positions and decide which interpretation of the findings they consider most sound.

KEYWORDS

Lorazepam; multiple sclerosis; performance validity tests; word memory test

Graver and Green (2020) express disagreement with our interpretation of a clinical case study and a randomized clinical trial, both of which included the Word Memory Test (WMT), a performance validity test (PVT) developed and commercialized by one of the commentary authors (i.e., Paul Green). While their commentary did not appear in either of the two journals in which our reports appeared (*Archives of Clinical Neuropsychology*, *The Clinical Neuropsychologist*), we appreciate the willingness of the Editors of *Applied Neuropsychology: Adult* to allow our response to Graver and Green to appear in the same journal. Our goal is not to change their position, but rather to respond formally to inaccuracies and to encourage readers, particularly students, to carefully consider both positions to better understand clinical case formulation from different perspectives.

The two primary issues described by Graver and Green are that (1) our clinical patient with an established diagnosis of multiple sclerosis (MS) who failed two computerized performance validity tests displayed a pattern of neuropsychological performance that was inconsistent with that diagnosis, thus supporting the validity test interpretation of poor effort, and (2) the results of our randomized clinical trial examining the acute effects of 2 mg of lorazepam on the WMT reflect invalid performance rather than drug effects.

Multiple sclerosis case report

Our case report described a woman with an established diagnosis of MS who was referred for neuropsychological evaluation of subjective memory concerns (Loring & Goldstein, 2019). Her findings included a neuropsychological profile consistent with an MS syndrome occurring in the context of failed stand-alone computerized PVTs. In an attempt to reconcile these two superficially inconsistent findings, we "speculate(d) that poor computerized (performance validity test) scores resulted from the disease-related features of MS, although we also discuss approaches to reconcile apparently contradictory performance validity testing (PVT) vs. neuropsychological results if the contributions of disease-related cognitive factors on PVTs scores are discounted. This case demonstrates the value of completing the assessment protocol despite obtaining PVT scores below publisher recommended cutoffs in clinical evaluations. If sub-threshold PVT scores are considered evidence of performance invalidity, it is still necessary to have an approach for interpreting seemingly credible neuropsychological test results rather than simply dismissing them as invalid" (p. 1192).

One of the diagnostic dilemmas using PVT measures outside of medico-legal evaluations of mild traumatic brain injury (TBI) is the scarcity of appropriate validation studies to guide clinical interpretation. We were misquoted by Graver and Green as saying "there are no available data on the WMT in MS patients," and actually state that the "WMT has not been systematically studied in MS," (p. 1200) to point out the lack of appropriate validation in our specific clinical context (McWhirter et al., 2020). Unfortunately, most stand-alone and embedded performance validity measures have not been systematically characterized

CONTACT David W. Loring dloring@emory.edu Emory Brain Health Center, 12 Executive Park, Atlanta, GA 30329, USA.

in homogenous neurological diagnostic groups across a spectrum of disease severity, and generalization to these populations needs to be made cautiously while trying to provide the highest quality patient care.

After deciding that the general neuropsychological results must be invalid based upon their interpretation of subthreshold WMT scores, Graver and Green seek support for that interpretation without consideration of the consistency of the neuropsychological profile with the MS diagnosis. Given the well-established white matter pathophysiology of MS affecting frontal system efficiency (Grzegorski & Losy, 2017), Graver and Green's citation of a prolonged Trail Making Part B is an illogical choice to infer evidence of poor effort. Graver and Green also characterize infrequency of poor scores compared to their expectations in support of the inference of poor effort, although the MS studies that they cite are actually studies of Clinically Isolated Syndrome (CIS) rather than MS (Uher et al., 2014; Viterbo et al., 2013). While suggestive of an eventual clinical diagnosis of MS, CIS commonly presents with optic neuritis, or an isolated brainstem or partial spinal-cord syndrome (Loring, 2015), and likely provides an under estimation of cognitive deficits in MS when based upon patterns obtained in CIS patients. While a review of neuropsychological aspects of MS is beyond the scope of this critique, several good reviews are available (Amato et al., 2008; Chiaravalloti & DeLuca, 2008; Oreja-Guevara et al., 2019; Sumowski et al., 2018).

All of the embedded PVT measures in our patient's neuropsychological assessment are based on peer-reviewed studies that characterize her performance levels as in the valid range. These embedded measures included Reliable Digit Span (Greiffenstein et al., 1994), Wisconsin Card Sorting Test (WCST) failure to maintain set errors (Greve et al., 2009), recognition scores for both Logical Memory and Visual Reproduction from the Wechsler Memory Scale-IV (Pearson, 2009), logistic regression estimation using the Rey Auditory Verbal Learning Test (Davis et al., 2012), and Complex Figure recognition (Lu et al., 2003). Although Graver and Green cite references suggesting the possibility of inadequate effort on several of these embedded measures, they are for measures that are often affected in patients with executive function and working memory impairment in MS. We again emphasize that embedded PVT measures have not been systematically examined in MS.

Graver and Green present a lengthy discussion of their own unpublished MS data, but it is difficult to evaluate these data since their methodology, findings, and conclusions have not been subjected to a full formal peer review—even basic characteristics of their MS patients are unknown (e.g., disease duration, MRI lesion burden, MS subtype, or whether McDonald criteria (Thompson et al., 2018) were employed). Further, we note that despite information in *Applied Neuropsychology: Adult's* Instructions to Authors to include a data availability statement, Graver and Green failed to include any declaration about how the data supporting their assertions could be obtained for independent review.

What is lost in the Graver and Green discussion of our case report is the patient's normal learning and memory scores, which enabled us to explicitly answer the referral question addressing her memory concerns, a fact that provided reassurance during her feedback session. Her normal memory scores on neuropsychological testing are inconsistent with a poor effort inference from WMT. *Primum non nocere.*

Lorazepam effects on WMT

Graver and Green spend considerable effort addressing what they perceive as incorrect conclusions in a healthy volunteer study of lorazepam effects on WMT. That study examined the effects of an acute 2 mg dose on WMT in a randomized, double-blind, placebo-controlled, crossover investigation in healthy volunteers published in 2011 in *The Clinical Neuropsychologist*. We concluded that "these data indicate that multiple WMT measures may be affected by acute (lorazepam) dosing, and provide additional evidence that potential latent variables and their effects on both (symptom validity test) performance and cognitive function should be part of the clinical decision-making process" (p. 799).

Regarding the lorazepam study design, we have successfully employed a similar experimental design across multiple studies examining the cognitive side effects of anti-seizure medications (ASMs). Our neuropsychological studies in ASMs have involved a broad range of medications including brivaracetam, carbamazepine, gabapentin, lamotrigine, levetiracetam, lorazepam, phenobarbital, phenytoin, topiramate, valproate, with all studies producing results supporting our *a priori* experimental hypotheses of differential ASM cognitive side-effects (Meador et al., 1995, 1999, 2005, 2007, 2011, 2019). We also note that this experimental design and study findings have been accepted in well-established peer-review journals.

We have conducted two independent studies demonstrating the negative cognitive effects of acute 2 mg administration of lorazepam (Loring et al., 2011; Meador et al., 2011); one of these studies also included neurophysiological measures demonstrating that acute 2 mg lorazepam decreases EEG alpha band peak power and also decreases beta band power reflecting decreased neuronal excitability (Meador et al., 2011). Neurocognitive findings in methodologically rigorous randomized double-blind, placebo-controlled trials cannot be simply dismissed because they do not conform to preconceived biases of the WMT test publisher.

As proponents of data sharing, these data were gladly provided to a colleague for *post-hoc* exploratory analyses, and findings were presented in a poster presentation at the 2013 meeting of the *American Society for Clinical Psychopharmacology* (Rohling, 2013). While it is beyond the scope of this report to provide a scientific critique of a conference abstract, In brief, the project was an exploratory analysis in which 16 embedded "validity" indices were derived internally from the dataset, with no available validity of the derived "validity" indices themselves. Thus, the meaningfulness of these "validity" measures is unknown. The scatter plot of these findings (see Figure 1) displays Overall Test Battery Mean (OTBM) scores across experimental

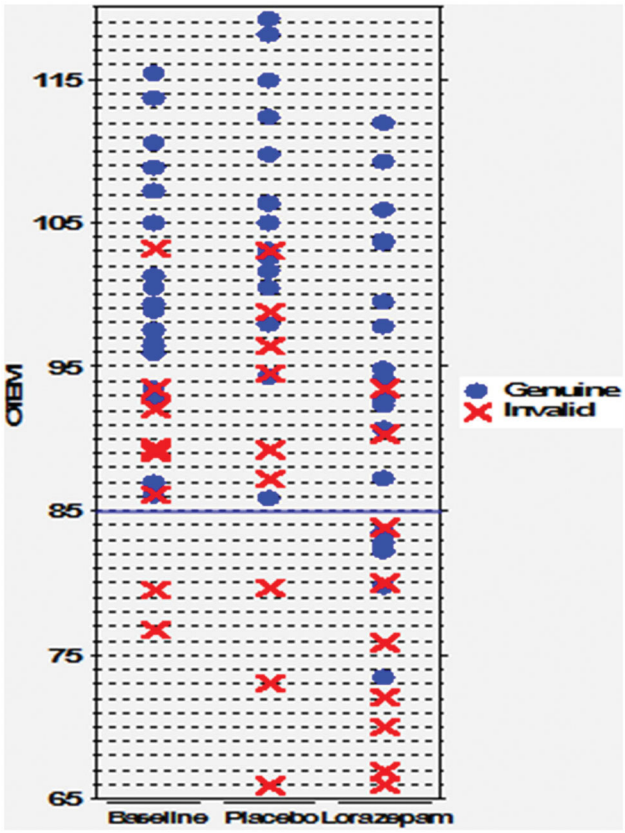

Figure 1. Scattergram of subjects characterized as having "invalid" peromances, represented by red X's of Overall Test Battery Mean across trial conditions (Rohling, 2013).

conditions, with subjects identified as having "invalid" profiles identified with a red X. It is difficult to reconcile the large number of "invalid" scores with OTBM scores in the normal range, broadly defined as OTBM scores greater than *SS* = 85. Unfortunately, internally derived measures from this dataset suffer from contamination artifacts that cannot be disentangled. Most scientifically rigorous neuropsychologists will not rely on results from a *post-hoc* exploratory abstract to refute a peer-reviewed double-blind, placebo-controlled randomized trial, an experimental design considered to reflect the highest level of evidence by multiple systems characterizing evidence quality (e.g., Burns et al., 2011; French & Gronseth, 2008; Straus et al., 2018). Interested readers can download a detailed critique that identifies multiple specific flaws in the Rohling *post-hoc* analysis (Loring, 2013).

Summary

We acknowledge that there will be clinicians who disagree with our interpretation, and the MS case report was intended to highlight approaches when attempting to reconcile two superficially discrepant findings. Consequently, we devoted an entire section of our case report to "'Ontological-Epistemological One-Worldness' and Choosing Between Apparent Contradictory Results." Faust's (2003) provocative article introduces the concept of "ontological-epistemological one-worldness," that is, the belief that "careful analysis and synthesis (of seemingly discrepant data) allows one to integrate them into meaningful or orderly results and patterns" (p. 430). Thus, it is easy to dismiss any unexpected variability simply as reflecting poor effort. As we note, however, neuropsychological evaluations often contain cognitive scores that are internally inconsistent (Brooks et al., 2009), and cognitively healthy individuals frequently have scores in the impaired range (Binder et al., 2009).

We concluded our case report by stating that "this case suggests that computerized PVTs *may* be affected in some MS patients by disease-related decreased information processing speed and working memory impairment (emphasis added). Even for those who disagree, there remains a need for viewing the apparently valid neuropsychological test profile within a consistent interpretative framework that addresses superficially incompatible findings within some logical decision-making framework. Different recommendations regarding best practice after obtaining subthreshold scores on PVT measures will vary based upon the clinical context in which the evaluations are being performed (i.e., forensic vs. medical). However, this case demonstrates that credible patterns of neuropsychological performance addressing the primary referral question may be obtained in at least some neurologic patients with subthreshold PVTs scores. Clinical neuropsychological evaluation requires judicious use of clinical judgment, and a 'one-size fits all' approach to interpretation and assessment protocol when subthreshold PVT scores are obtained is insufficient to replace good clinical practice" (p. 1200).

We appreciate this opportunity to provide lessons in clinical case formulation to our colleagues and peers, and to encourage all to read both positions critically and decide which interpretation of the findings they consider most sound. Articles to which Graver and Green refer, in addition to the 2013 *American Society for Clinical Psychopharmacology* PowerPoint and poster critique, may be downloaded for educational purposes at http://neurology.emory.edu/faculty/neuropsychology/loring_david.html.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Amato, M. P., Zipoli, V., & Portaccio, E. (2008). Cognitive changes in multiple sclerosis. *Expert Review of Neurotherapeutics*, *8*(10), 1585–1596. https://doi.org/10.1586/14737175.8.10.1585

Binder, L. M., Iverson, G. L., & Brooks, B. L. (2009). To err is human: "Abnormal" neuropsychological scores and variability are common in healthy adults. *Archives of Clinical Neuropsychology : The Official Journal of the National Academy of Neuropsychologists*, *24*(1), 31–46. https://doi.org/10.1093/arclin/acn001

Brooks, B. L., Iverson, G. L., Feldman, H. H., & Holdnack, J. A. (2009). Minimizing misdiagnosis: Psychometric criteria for possible or probable memory impairment. *Dementia and Geriatric Cognitive Disorders*, *27*(5), 439–450. https://doi.org/10.1159/000215390

Burns, P. B., Rohrich, R. J., & Chung, K. C. (2011). The levels of evidence and their role in evidence-based medicine. *Plastic and*

Reconstructive Surgery, *128*(1), 305–310. https://doi.org/10.1097/PRS.0b013e318219c171

Chiaravalloti, N. D., & DeLuca, J. (2008). Cognitive impairment in multiple sclerosis. *The Lancet Neurology*, *7*(12), 1139–1151. https://doi.org/10.1016/S1474-4422(08)70259-X

Davis, J. J., Millis, S. R., & Axelrod, B. N. (2012). Derivation of an embedded Rey Auditory Verbal Learning Test performance validity indicator. *The Clinical Neuropsychologist*, *26*(8), 1397–1408. https://doi.org/10.1080/13854046.2012.728627

Faust, D. (2003). Holistic thinking is not the whole story: Alternative or adjunct approaches for increasing the accuracy of legal evaluations. *Assessment*, *10*(4), 428–441. https://doi.org/10.1177/1073191103259534

French, J., & Gronseth, G. (2008). Lost in a jungle of evidence: We need a compass. *Neurology*, *71*(20), 1634–1638. https://doi.org/10.1212/01.wnl.0000336533.19610.1b

Graver, C., & Green, P. (2020). *Misleading conclusions about word memory test results in multiple sclerosis (MS) by Loring and Goldstein (2019)*. Adult.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment*, *6*(3), 218–224. https://doi.org/10.1037/1040-3590.6.3.218

Greve, K. W., Heinly, M. T., Bianchini, K. J., & Love, J. M. (2009). Malingering detection with the Wisconsin Card Sorting Test in mild traumatic brain injury. *The Clinical Neuropsychologist*, *23*(2), 343–362. https://doi.org/10.1080/13854040802054169

Grzegorski, T., & Losy, J. (2017). Cognitive impairment in multiple sclerosis – a review of current knowledge and recent research. *Reviews in the Neurosciences*, *28*(8), 845–860. https://doi.org/10.1515/revneuro-2017-0011

Loring, D. W. (2015). *INS dictionary of neuropsychology and clinical neurosciences* (2nd ed.). Oxford University Press.

Loring, D. W., & Goldstein, F. C. (2019). If invalid PVT scores are obtained, can valid neuropsychological profiles be believed? *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *34*(7), 1192–1202. https://doi.org/10.1093/arclin/acz028

Loring, D. W., Marino, S. E., Drane, D. L., Parfitt, D., Finney, G. R., & Meador, K. J. (2011). Lorazepam effects on Word Memory Test performance: A randomized, double-blind, placebo-controlled, crossover trial. *The Clinical Neuropsychologist*, *25*(5), 799–811. https://doi.org/10.1080/13854046.2011.583279

Loring, D. W. (2013). *ASCP post-hoc analysis comments*. http://neurology.emory.edu/faculty/neuropsychology/loring_david.html.

Lu, P. H., Boone, K. B., Cozolino, L., & Mitchell, C. (2003). Effectiveness of the Rey-Osterrieth Complex Figure Test and the Meyers and Meyers recognition trial in the detection of suspect effort. *The Clinical Neuropsychologist*, *17*(3), 426–440. https://doi.org/10.1076/clin.17.3.426.18083

McWhirter, L., Ritchie, C. W., Stone, J., & Carson, A. (2020). Performance validity test failure in clinical populations: A systematic review. *Journal of Neurology, Neurosurgery, and Psychiatry*. https://doi.org/10.1136/jnnp-2020-323776

Meador, K. J., Gevins, A., Leese, P. T., Otoul, C., & Loring, D. W. (2011). Neurocognitive effects of brivaracetam, levetiracetam, and lorazepam. *Epilepsia*, *52*(2), 264–272. https://doi.org/10.1111/j.1528-1167.2010.02746.x

Meador, K. J., Gevins, A., Loring, D. W., McEvoy, L. K., Ray, P. G., Smith, M. E., Motamedi, G. K., Evans, B. M., & Baum, C. (2007). Neuropsychological and neurophysiologic effects of carbamazepine and levetiracetam. *Neurology*, *69*(22), 2076–2084. https://doi.org/10.1212/01.wnl.0000281104.55418.60

Meador, K. J., Loring, D. W., Moore, E. E., Thompson, W. O., Nichols, M. E., Oberzan, R. E., Durkin, M. W., Gallagher, B. B., & King, D. W. (1995). Comparative cognitive effects of phenobarbital, phenytoin, and valproate in healthy adults. *Neurology*, *45*(8), 1494–1499. https://doi.org/10.1212/wnl.45.8.1494

Meador, K. J., Loring, D. W., Ray, P. G., Murro, A. M., King, D. W., Nichols, M. E., Deer, E. M., & Goff, W. T. (1999). Differential cognitive effects of carbamazepine and gabapentin. *Epilepsia*, *40*(9), 1279–1285. https://doi.org/10.1111/j.1528-1157.1999.tb00858.x

Meador, K. J., Loring, D. W., Vahle, V. J., Ray, P. G., Werz, M. A., Fessler, A. J., Ogrocki, P., Schoenberg, M. R., Miller, J. M., & Kustra, R. P. (2005). Cognitive and behavioral effects of lamotrigine and topiramate in healthy volunteers. *Neurology*, *64*(12), 2108–2114. https://doi.org/10.1212/01.WNL.0000165994.46777.BE

Meador, K. J., Seliger, J., Boyd, A., Razavi, B., Falco-Walter, J., Le, S., & Loring, D. W. (2019). Comparative neuropsychological effects of carbamazepine and eslicarbazepine acetate. *Epilepsy & Behavior: E&B*, *94*, 151–157. https://doi.org/10.1016/j.yebeh.2019.02.034

Oreja-Guevara, C., Ayuso Blanco, T., Brieva Ruiz, L., Hernández Pérez, M., Meca-Lallana, V., & Ramió-Torrentà, L. (2019). Cognitive dysfunctions and assessments in multiple sclerosis. *Frontiers in Neurology*, *10*, 581. https://doi.org/10.3389/fneur.2019.00581

Pearson. (2009). *Advanced clinical solutions for WAIS-IV and WMS-IV: Clinical and interpretive manual*. Pearson.

Rohling, M. L. (2013). When are your trial data real? [Poster presented]. American Society for Clinical Psychopharmacology, Boca Raton, FL.

Straus, S., Glasziou, P., Richardson, W. S., & Haynes, R. B. (2018). *Evidence-based medicine*. Elsivier. https://doi.org/10.1136/ebm.12.1.2-a

Sumowski, J. F., Benedict, R., Enzinger, C., Filippi, M., Geurts, J. J., Hamalainen, P., Hulst, H., Inglese, M., Leavitt, V. M., Rocca, M. A., Rosti-Otajarvi, E. M., & Rao, S. (2018). Cognition in multiple sclerosis: State of the field and priorities for the future. *Neurology*, *90*(6), 278–288. https://doi.org/10.1212/wnl.0000000000004977

The power of disagreement. (2016). *Nature Methods*, *13*(3), 185–185. https://doi.org/10.1038/nmeth.3798

Thompson, A. J., Banwell, B. L., Barkhof, F., Carroll, W. M., Coetzee, T., Comi, G., Correale, J., Fazekas, F., Filippi, M., Freedman, M. S., Fujihara, K., Galetta, S. L., Hartung, H. P., Kappos, L., Lublin, F. D., Marrie, R. A., Miller, A. E., Miller, D. H., Montalban, X., … Cohen, J. A. (2018). Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *The Lancet Neurology*, *17*(2), 162–173. https://doi.org/10.1016/S1474-4422(17)30470-2

Uher, T., Blahova-Dusankova, J., Horakova, D., Bergsland, N., Tyblova, M., Benedict, R. H. B., Kalincik, T., Ramasamy, D. P., Seidl, Z., Hagermeier, J., Vaneckova, M., Krasensky, J., Havrdova, E., & Zivadinov, R. (2014). Longitudinal MRI and neuropsychological assessment of patients with clinically isolated syndrome. *Journal of Neurology*, *261*(9), 1735–1744. https://doi.org/10.1007/s00415-014-7413-9

Viterbo, R. G., Iaffaldano, P., & Trojano, M. (2013). Verbal fluency deficits in clinically isolated syndrome suggestive of multiple sclerosis. *Journal of the Neurological Sciences*, *330*(1–2), 56–60. https://doi.org/10.1016/j.jns.2013.04.004