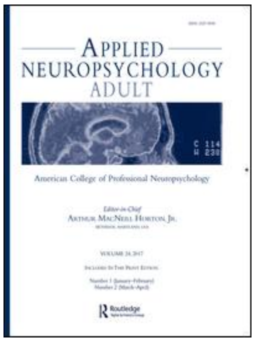

Applied Neuropsychology: Adult

ISSN: 2327-9095 (Print) 2327-9109 (Online) Journal homepage: https://www.tandfonline.com/loi/hapn21

Comparison of the Word Memory Test and the Test of Memory Malingering in detecting invalid performance in neuropsychological testing

Chloe Bhowmick, Rayna Hirst & Paul Green

To cite this article: Chloe Bhowmick, Rayna Hirst & Paul Green (2019): Comparison of the Word Memory Test and the Test of Memory Malingering in detecting invalid performance in neuropsychological testing, Applied Neuropsychology: Adult, DOI: 10.1080/23279095.2019.1658585

To link to this article: https://doi.org/10.1080/23279095.2019.1658585

Published online: 13 Sep 2019.

https://doi.org/10.1080/23279095.2019.1658585

Comparison of the Word Memory Test and the Test of Memory Malingering in detecting invalid performance in neuropsychological testing

Chloe Bhowmick[a], Rayna Hirst[a], and Paul Green[b]

[a]Department of Psychology, Palo Alto University, Palo Alto, CA, USA; [b]William Green, Greens Publishing, Kelowna, British Columbia, Canada

ABSTRACT

Given the prevalence of compensation seeking patients who exaggerate or fabricate their symptoms, the assessment of performance and symptom validity throughout testing is vital in neuropsychological evaluations. Two of the most commonly utilized performance validity tests (PVTs) are the Word Memory Test (WMT) and the Test of Memory Malingering (TOMM). While both have proven successful in detecting invalid performance, some studies suggest greater sensitivity in the WMT relative to the TOMM. To improve upon previous research, this study compared performance in individuals who completed both the WMT and TOMM during a neuropsychological evaluation. Participants included 268 cases from a clinical private practice consisting of primarily disability claimants. One-way multivariate analysis of variance (MANOVA) compared neuropsychological performance of participants who passed both PVTs ($n = 198$) versus those who failed the WMT but passed the TOMM ($n = 70$). Global suppression of neuropsychological scores was found for participants who failed the WMT but passed the TOMM, as well as more psychiatric symptoms reported on questionnaires, relative to those who passed both PVTs. These findings suggest that those passing the TOMM but failing the WMT demonstrated performance invalidity, which illustrates the WMT's enhanced sensitivity.

KEYWORDS
Forced choice; symptom validity testing; tests; performance validity

Estimates for the prevalence of compensation seeking neurological and chronic pain patients who exaggerate or fabricate symptoms vary (25–50%), but even conservative estimates indicate that this is a serious public health problem (Greve, Ord, Curtis, Bianchini, & Brennan, 2008). Given its prevalence, the assessment of performance and symptom validity is a vital part of the neuropsychological evaluation (Bush et al., 2005). Performance validity testing involves determining whether neuropsychological test scores are truly reflective of current cognitive functioning, while symptom validity assessment is the process of determining the accuracy of reported clinical symptoms. Patients have numerous internal and external motivations to either exaggerate or suppress symptoms (Bush et al., 2005; Green et al., 2001). Additionally, many patients have motivations to deny or suppress psychiatric distress or neurological impairment, such as in evaluations regarding decision-making capacity (Bush et al., 2005). Although patients can also fail validity testing unintentionally due to severe cognitive impairment or psychiatric symptoms (e.g., Delis & Wetter, 2007), the identification of invalid performance is still critical to ensure accurate results. Given these factors, experts agree that an evaluation without a formal assessment of validity is incomplete or negligent, particularly in litigious or forensic contexts (Bush et al., 2005; Green et al., 2001; Iverson, 2003).

Many stand-alone performance validity tests (PVTs) have been developed for administration during neuropsychological evaluations. The two most widely used PVTs are the Word Memory Test (WMT; Green, 1996) and the Test of Memory Malingering (TOMM; Tombaugh, 1996). The WMT is a stand-alone PVT assessing both verbal memory and performance validity on testing, while the TOMM is a stand-alone PVT assessing performance validity under the guise of visual memory. Both the WMT and TOMM have been well validated as measuring performance validity (Allen & Green, 1999; Green & Allen, 1999; Heyanka et al., 2015; Rees, Tombaugh, Gansler, & Moczynski, 1998; Tombaugh, 1996). Factor analysis has shown that performance validity is

CONTACT Chloe Bhowmick cbhowmick@paloaltou.edu Palo Alto University, 1791 Arastradero Road, Palo Alto, CA 94304-1337, USA.

distinct from actual memory abilities (Heyanka et al., 2015).

The WMT proves useful for assessing performance validity even in the presence of extensive neurological impairment or traumatic brain injury (TBI; e.g. Carone, 2014; Green & Flaro, 2016). Green and Allen (Allen & Green, 1999; Green & Allen, 1999) found that almost all patients with severe TBI or neurological diseases exceeded the cutoff scores of the primary effort indicators of the WMT. Goodrich-Hunsaker and Hopkins (2009) found that even three amnestic patients with bilateral hippocampal damage were able to score above established cutoffs on the WMT's primary validity indicators. The utility of the WMT has additionally been established in children with severe brain volume loss and extensive neurological and functional impairments (Carone, 2014), as well as those with intellectual disability (Green & Flaro, 2016). Green and Flaro (2019) also reported that WMT recognition scores in developmentally disabled children were unrelated to age and full-scale intelligence and unaffected by brain disease. Even children with a mean full scale intelligence quotient (FSIQ) of 59 obtained near perfect mean scores (over 95% correct).

The TOMM has also been extensively examined as a PVT. Rees et al. (1998) conducted validation studies that sought to explore the clinical utility of the TOMM under various simulation conditions (e.g., participants coached to malinger v.s. told to "try your best") and with a variety of participants (e.g., hospital outpatients, participants with TBI, and university students). This series of validation studies showed that the TOMM had 100% specificity, with sensitivity ranging from 84 to 100% (Rees et al., 1998). Similar to the WMT, the TOMM was also found to have a lack of sensitivity to cognitive deficits from TBI, supporting its use as a performance validity indicator which is relatively unaffected by cognitive impairment.

Given that both the WMT and TOMM are commonly utilized PVTs, researchers have compared the two, particularly in forensic contexts. Both the WMT and TOMM produced good discriminatory performance in a criminally forensic population; the TOMM had greater sensitivity and negative predictive power than the WMT, but the WMT had greater specificity and positive predictive power (Fazio, Sanders, & Denney, 2015). Bauer, O'Bryant, Lynch, McCaffrey, and Fisher (2007) found that the WMT had greater predictive power and greater sensitivity than the TOMM in a sample of mild head injury litigants from a forensic private practice. Greiffenstein and colleagues (Greiffenstein, Greve, Bianchini, & Baker, 2008) showed that when using balanced comparison cutoff criteria (i.e., comparing WMT's Immediate Recognition, Delayed Recognition, and Internal Consistency primary validity indicators to TOMM Trial 1, Trial 2, and Retention, as opposed to traditional cutoffs), the two tests exhibited comparable failure rates, though the WMT still had higher sensitivity. While both the WMT and TOMM are successful at detecting simulated mild TBI malingerers, the WMT has been found to be more sensitive than the TOMM (Lau et al., 2017; Tan, Slick, Strauss, & Hultsch, 2002). Greater sensitivity has also been shown in non-head injury disability claimants, where about half of participants who scored between 45 and 49 on TOMM Trial 2 or Retention (scores that qualify as passing, per the TOMM manual) fell below the recommended cutoffs for the WMT (Gervais, Rohling, Green, & Ford, 2004). Furthermore, Rienstra et al. (2013) found that the WMT, and not the TOMM, was able to accurately classify patients with mild cognitive impairment (MCI) into groups of credible and non-credible responders.

One criticism of the WMT is that it produces high false positive rates. Multiple studies found that interference or distraction tasks impact WMT performance more than TOMM (Batt, Shores, & Chekaluk, 2008; Eglit, Lynch, & McCaffrey, 2016; Greve et al., 2008), suggesting WMT could be vulnerable to false positives. Merten, Bossink, and Schmand (2007) found that the TOMM was less impacted than the WMT by severe cognitive impairment in two separate studies of neurological patients, again suggesting possible false positives.

Yet, numerous research suggests that the WMT does not demonstrate disproportionate false positive rates. Gervais et al. (2004) did not find excessive false positive rates when comparing WMT performance to the Computerized Assessment of Response Bias (CARB). Fazio et al. (2015) found that when the Genuine Memory Impairment Profile (GMIP; an algorithm assessing whether failed WMT performance is due to genuine cognitive impairment by comparing WMT to other memory test performance) was implemented, false positive rates fell below that of the TOMM. Additionally, Green, Flaro, and Courtney (2009) assessed false positives in a shorter version of the WMT, the Medical Symptom Validity Test (MSVT), and found that adults with mild TBI failed both the WMT and MSVT nearly 10 times as often as children with extensive cognitive impairment, suggesting a negligible false positive rate. Notably, Erdodi, Green, Sirianni, and Abeare (2019) recently reported minimally higher failure rates on the WMT (44.7% of patients) compared to other PVTs (39.1–41.8%), again suggesting minimal or no false positive errors. Additionally, the vast majority of patients who failed the WMT (94.7%) had independent evidence of performance invalidity on other PVTs.

In light of the somewhat discrepant findings comparing WMT and TOMM performance, Mossman, Wygant, and Gervais (2012) searched for the "gold standard" PVT by examining the CARB, WMT, and TOMM. They controlled for methodological bias by using latent class modeling of data from forensic cases, instead of using simulated malingering, to increase ecological validity. When they held false positives at a rate of 0.02, the WMT was able to detect true positives at a rate of approximately 65%, compared to a rate of approximately 49% for the TOMM and 35% for the CARB. These results suggest that WMT is closer to the gold standard than TOMM or CARB.

This review of the literature shows that there remains some disagreement as to which PVT is superior in detecting invalid performance. This disagreement indicates that further research is vital, as both the WMT and TOMM are widely utilized PVTs (Martin, Schroeder, & Odland, 2015). Therefore, the present study compared the ability of the WMT and the TOMM to detect performance validity on neuropsychological testing in a large predominantly disability claimant sample. It was hypothesized that the WMT would be more sensitive than the TOMM in detecting performance validity across a battery of neuropsychological testing. This hypothesis was tested by examining neuropsychological performance in a group of predominantly disability claimants to determine whether those who failed the WMT but passed the TOMM exhibited lower neuropsychological test performance across multiple cognitive domains than participants who passed both the WMT and TOMM. Lower performance across all neuropsychological tests would imply that failing only the WMT (while passing the TOMM) suggests performance invalidity and, thus, that the WMT is more sensitive to invalid performance than the TOMM. The main alternative outcome would be no difference in neuropsychological test scores between those passing both PVTs versus those who failed only the WMT. Such an outcome would support lower specificity for the WMT than the TOMM.

Method

Participants

Participants were garnered from a total data set consisting of 2,174 cases from a clinical private practice containing evaluations for predominantly disability claimants seen by one of the authors [P.G.] for psychological assessment. From the initial 2,174 cases, participants who completed both the WMT and TOMM were assigned to one of two groups: (a) individuals who passed both the WMT and TOMM, and (b) those who passed the TOMM but failed the WMT. Final analysis of data consisted of 268 participants who were either in the group of participants who passed both the WMT and TOMM (Pass Both) or the group that passed the TOMM but failed the WMT (Pass TOMM, Fail WMT). None of the participants who passed the WMT failed the TOMM; thus, this was not a group that was included in the final analysis. Only three participants reported not speaking English fluently, and 92.2% of the sample reported that English was their first language. Please see Table 1 for demographic information for the experimental groups.

Measures

Participants were assessed using a flexible battery approach. As a result, not every participant completed the same tests; however, the assessments that were used are common to psychologists. Assessment batteries were typically completed over a two-day period and involved the assessment of domains such as general intellectual ability, academic achievement, verbal and visual memory, gross and fine motor abilities, and executive functions. Additionally, participants completed self-report symptom inventories. Of note, the *n* per test will be listed in the Results section.

The Wechsler Adult Intelligence Scale – Revised (WAIS-R; Wechsler, 1981) was administered to measure general cognitive ability. Academic achievement was measured using the Wide Range Achievement Test – Third Edition (WRAT-3; Wilkinson, 1993). The California Verbal Learning Test – Second Edition (CVLT-II; Delis, Kramer, Kaplan, & Ober, 2000), the Wechsler Memory Scale – Third Edition Visual Memory Span subtest (WMS-III VMS; Wechsler, 1997), the Rey Complex Figure Test (RCFT; Meyers & Meyers, 1995), and the Story Recall Test (Green, 1996) were all administered to assess for various components of memory functioning. The Story Recall Test may not be as widely known as other measures. It involves being presented with five tape-recorded

Table 1. Demographic information per experimental group.

	Pass Both ($n = 198$)	Pass TOMM/Fail WMT ($n = 70$)	Total ($N = 268$)
Age	$M = 45.49$ ($SD = 11.96$)	$M = 47.34$ ($SD = 11.10$)	$\mu = 45.97$ ($SD = 11.75$)
Sex			
Female, *n* (%)	87 (44)	38 (54)	125 (47)
Male, *n* (%)	111 (56)	32 (46)	143 (53)
Years of Education	$M = 12.90$ ($SD = 2.58$)	$M = 12.53$ ($SD = 2.61$)	$\mu = 12.80$ ($SD = 2.59$)
Referral Diagnosis			
Chronic Pain, *n* (%)	12 (6)	3 (4)	15 (6)
Dementia Evaluation, *n* (%)	8 (5)	3 (4)	11 (4)
Mild Traumatic Brain Injury, *n* (%)	27 (13)	21 (30)	48 (18)
Moderate or Severe Traumatic Brain injury, *n* (%)	19 (10)	5 (7)	24 (9)
Neurological Conditions, *n* (%)	38 (19)	7 (10)	45 (17)
Orthopedic Concerns, *n* (%)	4 (2)	0	4 (1)
Primary Diagnosis of Anxiety, *n* (%)	4 (2)	1 (1)	5 (2)
Primary Diagnosis of Bipolar Disorder, *n* (%)	3 (1)	3 (4)	6 (2)
Primary Diagnosis of Depression, *n* (%)	30 (15)	7 (10)	37 (14)
Primary Diagnosis of Psychotic Disorder, *n* (%)	2 (1)	0	2 (1)
Police Recruitment Process, *n* (%)	8 (5)	0	8 (3)
Other, *n* (%)	39 (19)	16 (23)	55 (21)
Data Not Available, *n* (%)	4 (2)	4 (6)	8 (3)

Note. TOMM = Test of Memory Malingering; WMT = Word Memory Test. There were no significant differences between groups in any demographic variables.

stories via headphones and being tested for both immediate recall and recall after a 30-minute delay (Green, 1996). The Category Test (Halstead, 1947; Reitan & Wolfson, 1993), the Trail Making Test (TMT; Reitan, 1979), Gorham's Proverbs Test (Gorham, 1956), and the Thurstone Word Fluency Test (Thurstone, 1938) were administered to assess for executive functioning. Executive functioning on a visual task was measured using Ruff's Figural Fluency Test (Ruff, 1996). The Grip Strength Test (GST; Reitan & Wolfson, 1993) and the Grooved Pegboard Test (GPT; Kløve, 1963) were administered to measure motor functioning. The Alberta Smell Test (Green & Iverson, 2001) assessed for impairment of olfactory identification. The Emotional Perception Test (Green, 1997) assessed judgment of emotion in tone of voice. Lastly, patients completed the Beck Depression Inventory – Second Edition (BDI-II; Beck, Steer, & Brown, 1996) and the Memory Complaints Inventory (Green, 2004) as self-reports of symptom complaints regarding depression and memory, respectively.

All participants completed the Word Memory Test (WMT; Green, 1996) and the Test of Memory Malingering (TOMM; Tombaugh, 1996). The WMT consists of multiple verbal memory subtests ranging from extremely easy recognition memory tasks to a quite difficult free recall task. The primary effort indicators are the Immediate Recognition (IR) and Delayed Recognition (DR) subtests, which assess for the recognition of an earlier presented word list using a forced-choice paradigm. The computer program calculates the consistency of recognition performance across the IR and DR trials. Given that only a few of the participants (4%) were referred for a potential dementia diagnosis, the GMIP was not applied to this sample.

The TOMM assesses for effort under the guise of testing visual memory (Tombaugh, 1996). The TOMM consists of two main trials, with a retention trial administered if suboptimal effort is suspected. In Trial 1, test takers are shown a series of 50 pictures in a flipbook format and then are presented with a forced-choice recognition task regarding earlier stimuli. Trial 2 consists of the same format as Trial 1, but the same stimuli are presented in a different order. The TOMM manual indicates that if a test taker scores less than 45 out of 50 on Trial 2, then the examiner should administer the Retention trial 10 minutes later, which consists of just the forced-choice recognition of the same stimuli.

Procedure

Neuropsychological testing performance was analyzed in relation to performance on both the WMT and TOMM. PVT data were available for all participants and completion of both PVTs was a prerequisite for inclusion into the study. To ensure an adequate sample size for reliable results, neuropsychological tests were examined only if they had a sample of at least 70 participants who had completed them.

The relationship between the WMT and the TOMM was explored with regards to overall neuropsychological testing scores. Analyses examined whether neuropsychological performance differed in those who passed both PVTs (demonstrating valid performance) relative to those who passed the TOMM but failed the WMT (suspect performance). Raw scores were used for analyses to maintain adequate distribution of scores and prevent the restricted range that can occur with analyzing standardized scores. For tests that included multiple outcome measures, such as the CVLT-II, a Multivariate Analysis of Variance (MANOVA) was utilized to compare neuropsychological performance between the two groups to account for intercorrelations between the multiple outcome measures of a given test. On tests where there was only one outcome measure, an Analysis of Variance (ANOVA) was utilized to compare the groups on their neuropsychological test performance. Separate MANOVAs and ANOVAs were conducted due to the different *n* between tests.

Results

For this study we focused on two out of four possible groups. No participant passed the WMT but failed the TOMM, so that was not a choice for group comparison. The group who failed both the WMT and TOMM exhibited a global suppression of neuropsychological test scores and reported more cognitive and psychiatric difficulties compared to participants who only failed the WMT; these findings are beyond the scope of the present study, and are not presented to preserve clarity and conciseness. It is commonly known that the more PVTs failed, the greater the suppression of neuropsychological test scores (e.g., Boone, 2007).

Comparison statistics were used to examine whether there were significant demographic differences between the remaining two groups: those who passed the TOMM but failed the WMT ("Pass TOMM, Fail WMT" $n = 70$), and those who passed both the TOMM and WMT ("Pass Both," $n = 198$). ANOVA was used to examine continuous variables (e.g., age, years of education), while a chi-square analysis was used for categorical variables (e.g., gender). There were no significant differences between the two groups with regard to age ($F(1) = 1.288$, $p = .237$), years of education ($F(1) = 1.061$, $p = .304$), or gender ($\chi^2 = 2.224$, $p = .136$).

A one-way MANOVA or ANOVA compared neuropsychological test performance between the two groups. Overall, results in Table 2 consistently showed that those who passed the TOMM but failed the WMT exhibited significantly lower neuropsychological test performance across nearly all cognitive tests relative to participants who passed both the TOMM and WMT. Results are described in more detail in the following section.

General intellectual and academic functioning

A one-way MANOVA examined between-group differences regarding WAIS-R performance. The Pass Both group ($n = 178$) performed significantly better than the Pass TOMM, Fail WMT group ($n = 63$; Wilks' $\Lambda(3,237) = 6.447$, $p < .001$, multivariate $\eta^2 = .075$). Follow-up univariate ANOVAs found that Pass TOMM, Fail WMT performed worse on the WAIS-R Full Scale IQ ($F(1,239) = 12.368$, $p = .001$, $\eta^2 = .049$), WAIS-R Verbal IQ ($F(1,240) = 5.383$, $p = .021$, $\eta^2 = .022$), and WAIS-R Performance IQ ($F(1,240) = 18.939$, $p < .001$, $\eta^2 = .073$).

On the WRAT-3, the Pass Both group ($n = 57$) performed significantly better than the Pass TOMM, Fail WMT group ($n = 15$; Wilks' $\Lambda(3,68) = 3.254$, $p = .021$, multivariate $\eta^2 = .126$). Univariate ANOVAs revealed that the Pass TOMM, Fail WMT group performed worse on the math subtest ($F(1,70) = 6.342$, $p = .014$, $\eta^2 = .083$). There was no group difference found for scores on the reading ($F(1,70) = .147$, $p = .700$) or spelling ($F(1,70) = .056$, $p = .813$) subtests.

Learning and memory testing

A one-way MANOVA examined between-group differences in verbal learning and memory performance. On the CVLT-II, the Pass Both group ($n = 191$) performed significantly better than the Pass TOMM, Fail WMT group ($n = 68$; Wilks' $\Lambda(9,249) = 8.174$, $p < .001$, multivariate $\eta^2 = .228$). Subsequent univariate ANOVAs revealed that Pass TOMM, Fail WMT performed worse on all subtests examined: List A Trials 1–5 Total Correct ($F(1,257) = 43.478$, $p < .001$, $\eta^2 = .145$), List A Trial 1 Total Correct ($F(1,257) = 10.235$, $p = .002$, $\eta^2 = .038$), List A Trial 5 Total Correct ($F(1,257) = 48.093$, $p < .001$, $\eta^2 = .158$), Short-Delay Free Recall ($F(1,257) = 41.019$, $p < .001$, $\eta^2 = .113$), Long-Delay Free Recall ($F(1,257) = 45.116$, $p < .001$, $\eta^2 = .149$), and Recognition Hits ($F(1,257) = 42.193$, $p < .001$, $\eta^2 = .141$). On a story recall task, the Pass Both group ($n = 162$) performed significantly better than the Pass TOMM, Fail WMT group ($n = 61$; Wilks' $\Lambda(2,220) = 15.320$, $p < .001$, multivariate $\eta^2 = .122$). Follow-up univariate ANOVAs revealed that Pass TOMM, Fail WMT performed worse on both immediate ($F(1,221) = 17.447$, $p < .001$, $\eta^2 = .073$) and delayed recall ($F(1,221) = 30.779$, $p < .001$, $\eta^2 = .122$).

Table 2. Mean and standard deviation for MANOVAs/ANOVAs of neuropsychological test performance.

Test	Mean (*SD*) for Pass Both	Mean (*SD*) for Pass T/Fail W	Overall Result
Alberta Smell Test			
Right Nostril	4.80 (2.66)	4.19 (2.44)	No Significant Difference
Left Nostril	4.88 (2.48)	4.56 (2.56)	
California Verbal Learning Test – 2nd Edition			
List A Trials Total Correct (Raw)	51.91 (10.18)	42.28 (10.78)	Pass Both > Pass T/Fail W
List A Trial 1 Number Correct (Raw)	6.61 (2.14)	5.65 (2.09)	
List A Trial 5 Number Correct (Raw)	12.42 (2.37)	10.01 (2.70)	
List B Number Correct (Raw)	6.24 (2.10)	5.09 (2.23)	
List A Short-Delay Free Recall (Raw)	10.56 (2.96)	7.88 (2.94)	
List A Short-Delay Cued Recall (Raw)	11.60 (2.69)	9.38 (2.89)	
List A Long-Delay Free Recall (Raw)	10.92 (2.98)	8.03 (3.21)	
List A Long-Delay Cued Recall (Raw)	11.64 (2.87)	9.16 (3.06)	
Recognition Hits (Raw)	14.63 (1.55)	13.06 (2.09)	
Category Test			
Raw Score (number of errors)	58.05 (30.23)	68.52 (28.89)	Pass Both > Pass T/Fail W
Emotional Perception Test			
Raw Score out of 30	10.65 (4.24)	10.79 (3.81)	No Significant Difference
Raw Score out of 45	14.78 (5.87)	15.48 (5.63)	
Gorham's Proverbs Test	12.09 (5.23)	9.96 (5.00)	Pass Both > Pass T/Fail W
Grip Strength			
Right Hand (Raw)	38.55 (15.04)	31.24 (13.20)	Pass Both > Pass T/Fail W
Left Hand (Raw)	35.90 (15.24)	30.29 (14.37)	
Grooved Pegboard			
Right Hand (Raw)	71.62 (19.60)	83.69 (24.54)	Pass Both > Pass T/Fail W
Left Hand (Raw)	81.40 (22.88)	92.39 (35.84)	
Rey Complex Figure Test			
Copy	31.15 (3.76)	29.14 (4.89)	Pass Both > Pass T/Fail W
Immediate Recall (Percentile)	32.33 (29.75)	21.49 (25.42)	
Delayed Recall (Percentile)	30.57 (29.25)	18.96 (23.23)	
Recognition (Percentile)	44.26 (30.93)	31.58 (28.27)	
Ruff Figural Fluency Test			
Unique Designs	75.16 (22.32)	68.11 (20.98)	No Significant Difference
Perseverations	9.24 (11.61)	9.26 (15.42)	
Story Recall Test			
Immediate Recall	45.98 (9.99)	39.79 (9.50)	Pass Both > Pass T/Fail W
Delayed Recall	35.33 (12.29)	25.10 (12.26)	
Thurstone Word Fluency Test			
Raw Score	49.89 (17.21)	42.23 (17.57)	Pass Both > Pass T/Fail W
Trail Making Test			
Part A (in Seconds)	33.91 (12.79)	38.40 (14.00)	Pass Both > Pass T/Fail W
Part B (in Seconds)	77.50 (35.07)	102.29 (49.00)	
Wechsler Adult Intelligence Scale - Revised			
Full Scale IQ	106.13 (12.20)	99.83 (12.34)	Pass Both > Pass T/Fail W
Verbal IQ	104.57 (13.02)	100.09 (13.56)	
Performance IQ	108.76 (12.29)	100.86 (12.66)	
Wechsler Memory Scale – Visual Memory Span			
Forward	43.09 (30.08)	34.47 (25.91)	Pass Both > Pass T/Fail W
Backward	55.60 (25.66)	41.93 (25.43)	
Wide Range Achievement Test – 3rd Edition			
Reading (Raw)	47.40 (4.42)	46.93 (3.31)	Pass Both > Pass T/Fail W
Spelling (Raw)	41.88 (5.36)	41.53 (3.14)	
Math (Raw)	39.04 (4.75)	35.67 (4.01)	

A one-way MANOVA also examined between-group differences in visual learning and memory performance. On WMS-III Visual Memory Span, the Pass Both group ($n = 189$) performed significantly better than the Pass TOMM, Fail WMT group ($n = 68$; Wilks' $\Lambda(2,254) = 7.159$, $p = .001$, multivariate $\eta^2 = .053$). Follow-up univariate ANOVAs revealed that Pass TOMM, Fail WMT performed worse on both the forward ($F(1,255) = 4.405$ $p = .037$, $\eta^2 = .017$) and backward items ($F(1,255) = 14.273$, $p < .001$; $\eta^2 = .053$). Additionally, the Pass Both group ($n = 187$) performed significantly better than the Pass TOMM, Fail WMT group on the Rey Complex Figure Test ($n = 65$; Wilks' $\Lambda(4,247) = 4.789$, $p = .001$, multivariate $\eta^2 = .072$). Further univariate ANOVAs found that Pass TOMM, Fail WMT performed worse on the copy trial ($F(1,250) = 11.744$, $p = .001$, $\eta^2 = .045$), immediate recall trial ($F(1,250) = 6.882$, $p = .009$), $\eta^2 = .027$), delayed recall trial ($F(1,250) = 8.384$, $p = .004$, $\eta^2 = .032$), and recognition trial ($F(1,250) = 8.458$, $p = .004$, $\eta^2 = .033$).

Executive functioning

A one-way MANOVA examined between-group differences in executive functioning performance. The

Pass Both group ($n = 193$) performed significantly better than the Pass TOMM, Fail WMT group ($n = 70$; Wilks' $\Lambda(2,260) = 7.121$, $p = .001$, multivariate $\eta^2 = .052$) on the TMT. The Pass TOMM, Fail WMT group obtained lower scores on both Part A ($F(1,261) = 5.995$, $p = .015$, $\eta^2 = .022$) and Part B ($F(1,261) = 13.780$, $p < .001$, $\eta^2 = .050$). A one-way ANOVA on the Gorham's Proverbs Test revealed that Pass TOMM, Fail WMT ($n = 68$) scored significantly lower compared to Pass Both ($n = 192$; $F(1,258) = 8.532$, $p = .004$, $\eta^2 = .032$). A one-way ANOVA also found that Pass TOMM, Fail WMT ($n = 55$) scored lower than Pass Both on the Thurstone Word Fluency Test ($n = 168$; $F(1,221) = 8.117$, $p = .005$, $\eta^2 = .35$). Finally, a one-way MANOVA of the Category Test performance revealed that the Pass Both group ($n = 86$) performed significantly better than the Pass TOMM, Fail WMT group ($n = 65$; Wilks' $\Lambda(2,248) = 3.153$, $p = .044$, multivariate $\eta^2 = .025$). The Pass TOMM, Fail WMT group committed more errors ($m = 68.52$) on the Category Test than the Pass Both group ($m = 58.05$). A one-way MANOVA revealed that the Pass Both ($n = 165$) and Pass TOMM, Fail WMT ($n = 53$) groups did not differ regarding performance on the Ruff Figural Fluency Test (Wilks' $\Lambda(2,225) = .2.065$, $p = .129$, multivariate $\eta^2 = .019$).

Sensory-motor testing

A one-way MANOVA examined between-group differences in sensory and motor functioning. The Pass Both group ($n = 187$) performed significantly better on a Grip Strength task than the Pass TOMM, Fail WMT group ($n = 67$; Wilks' $\Lambda(2,252) = 6.293$, $p = .002$, multivariate $\eta^2 = .048$). Pass TOMM, Fail WMT obtained worse scores using both the right ($F(1,253) = 12.256$, $p = .001$, $\eta^2 = .047$) and left hands ($F(1,253) = 6.902$, $p = .009$, $\eta^2 = .027$). On Grooved Pegboard, the Pass Both group ($n = 191$) performed significantly better than the Pass TOMM, Fail WMT group ($n = 67$; Wilks' $\Lambda(2,255) = 8.283$, $p < .001$, multivariate $\eta^2 = .061$). Pass TOMM, Fail WMT performed significantly worse on both the right ($F(1,256) = 16.411$, $p < .001$, $\eta^2 = .060$) and left hands ($F(1,256) = 8.311$, $p = .004$, $\eta^2 = .031$). There were no between-group differences on the Alberta Smell Test between the Pass Both ($n = 193$) and Pass TOMM, Fail WMT groups ($n = 68$; Wilks' $\Lambda(2,258) = 1.384$, $p = .252$, multivariate $\eta^2 = .011$).

Emotional perception

A one-way MANOVA examined for between group differences in emotional perception. No differences were found between the Pass Both ($n = 124$) and Pass TOMM, Fail WMT ($n = 48$; Wilks' $\Lambda(2,169) = 1.372$, $p = .256$, multivariate $\eta^2 = .016$) groups in performance on the Emotional Perception Test.

Self-report inventories

A one-way ANOVA examined between-group differences concerning BDI-II scores and found that those in the Pass TOMM, Fail WMT group ($n = 67$) obtained higher scores than the Pass Both group ($n = 186$; $F(1,251) = 4.346$, $p = .038$, $\eta^2 = .017$). A one-way MANOVA examined Memory Complaints Inventory scores, and found that the Pass Both group ($n = 197$) reported significantly fewer memory complaints than the Pass TOMM, Fail WMT group ($n = 69$; Wilks' $\Lambda(9,256) = 6.734$, $p < .001$, multivariate $\eta^2 = .191$). A follow-up univariate ANOVA found that Pass TOMM, Fail WMT had a higher total mean of subscale scores ($F(1,264) = 43.949$, $p < .001$, $\eta^2 = .143$). Furthermore, Pass TOMM, Fail WMT obtained higher scores for the following subscales: General Memory Problems ($F(1,264) = 44.925$, $p < .001$, $\eta^2 = .145$), Numeric Information Problems ($F(1,264) = 36.604$, $p < .001$, $\eta^2 = .122$), Visuospatial Memory Problems ($F(1,264) = 39.354$, $p < .001$, $\eta^2 = .130$), Verbal Memory Problems ($F(1,264) = 33.603$, $p < .001$, $\eta^2 = .113$), Memory Complaints due to Pain ($F(1,264) = 4.718$, $p = .031$, $\eta^2 = .018$), Memory Problems Interfering at Work ($F(1,264) = 21.921$, $p < .001$, $\eta^2 = .077$), Impairment of Remote Memory ($F(1,264) = 23.862$, $p < .001$, $\eta^2 = .083$), Amnesia for Complex Behavior ($F(1,264) = 40.057$, $p < .001$, $\eta^2 = .132$), and Memory for Antisocial Behavior ($F(1,264) = 18.089$, $p < .001$, $\eta^2 = .064$). See Table 3 for means and standard deviations regarding self-report data between experimental groups.

Other PVT performance

Given the suggestion from prior researchers that the WMT may be vulnerable to false positive errors, we examined performance on other PVTs (MSVT, Nonverbal-MSVT, and Reliable Digit Span; RDS) within the Pass TOMM, Fail WMT group ($n = 70$) to determine whether their other PVT performance was consistent with invalid performance. One participant did not complete any PVTs besides the TOMM and WMT. Twenty-three participants failed one other

Table 3. Group comparison of means and standard deviations for self-report measures.

Test	Mean (SD) for Pass Both	Mean (SD) for Pass T/Fail W	Overall Result
Beck Depression Inventory – 2nd Edition			
Raw Score	15.96 (11.37)	19.36 (11.61)	Pass T/Fail W Reported More Depression Symptoms
Memory Complaints Inventory			
Total Score	22.03 (15.92)	37.45 (18.53)	Pass T/Fail W Reported More Memory Complaints
General Memory Problems	24.07 (19.13)	42.91 (22.65)	
Numeric Information	29.76 (21.09)	54.00 (43.61)	
Visuospatial Memory	20.61 (18.74)	38.26 (23.61)	
Verbal Memory	33.90 (23.77)	53.41 (24.86)	
Pain Interferes with Memory	25.07 (29.30)	34.17 (31.77)	
Memory Interferes with Work	29.09 (26.80)	47.03 (29.03)	
Impairment of Remote Memory	15.11 (14.05)	25.04 (15.85)	
Amnesia for Complex Behavior	14.68 (15.29)	29.46 (20.30)	
Amnesia for Antisocial Behavior	5.96 (9.57)	12.80 (15.78)	

PVT along with the WMT, 17 participants failed two PVTs along with the WMT, and four participants failed three PVTs along with the WMT. Twenty-five participants passed all PVTs except for the WMT.

Discussion

When a patient fails a PVT, we expect to find that their scores across a neuropsychological test battery will be significantly lower than those who pass that PVT. This is one of the primary acid tests of whether that PVT is truly measuring the validity of testing results. If a failed PVT does not predict lowered scores in at least some components of a neuropsychological battery, indicating performance that is not truly indicative of an individual's cognitive abilities, it is not really acting as a PVT (Larrabee, 2012).

The current study examined the scores from ability-based tests and self-report measures of psychiatric symptoms from participants who completed both the WMT and TOMM. Groups were categorized as those who passed the TOMM but failed the WMT and those who passed both. The Pass Both group was presumed to have put forth valid performance on testing, given that they passed two objective measures of performance validity. Importantly, it should be noted that no participant passed the WMT but failed the TOMM. In nearly every cognitive domain assessed, neuropsychological test scores were lower in participants who passed the TOMM but failed the WMT compared to participants who passed both PVTs. Thus, those who passed the TOMM but failed the WMT showed a global suppression of test scores. While failure of the TOMM has been shown to predict lower neuropsychological test scores (Constantinou, Bauer, Ashendorf, Fisher, & McCaffrey, 2005), the present data suggest that the WMT is superior in predicting lower scores when the TOMM is passed.

A common criticism of the WMT is that it functions more as an assessment of memory than performance validity (e.g., Greve et al., 2008). If this were true, then in the present study, lower neuropsychological test scores would be expected for only tests of verbal memory. Instead, our study found that participants who passed the TOMM but failed the WMT performed significantly worse on visuospatial tasks, tests of emotional judgment, problem-solving tasks, academic/achievement testing, and global intellectual functioning. This finding suggests that the WMT is more sensitive to performance invalidity than the TOMM, which replicates previous findings in the literature (e.g., Gervais et al., 2004). Erdodi et al. (2019) found a similar global suppression of neuropsychological test scores on participants who failed the WMT, indicating that this is not an isolated finding. Similarly, Green and Flaro (2019) observed a widespread suppression of neuropsychological test scores in developmentally disabled children who failed the WMT, even if other PVTs were passed.

In the current study, the group who passed the TOMM but failed the WMT reported significantly more depressive symptoms (per BDI-II) and significantly more memory complaints (per Memory Complaints Inventory) when compared with the Pass Both group. Irrespective of the TOMM being passed, the presence of a WMT failure indicates a higher likelihood of reporting both more cognitive and psychiatric difficulties.

The limitations of the current study include the fact that it involved a mixed diagnostic group. The effects were not examined in relation to diagnosis. Not all participants were administered a uniform battery, so the *n* of each test administered differs. Furthermore, while these participants were from the same private practice, different examiners could have administered batteries across participants, and different examiner characteristics such as administration style or preference for a particular test could influence outcome measures. Future studies examining PVT utility could administer a uniform battery to better control for examiner characteristics or test order. Additionally, given that the data were presented as raw scores, the number of patients with impaired performances on neuropsychological tests cannot be calculated. While raw scores are preferable for maintaining the range of data, given that standardized scores can truncate the distribution (e.g., McSweeney, Naugle, Chelune, & Lüders, 1993; Strauss, Sherman, & Spreen, 2006), future studies may wish to incorporate standardized scores into analyses.

An inherent limitation with any PVT research study is determining whether a participant was truly giving an invalid performance or if he/she did not pass a PVT due to genuine cognitive impairment (e.g., Bush et al., 2005). We chose to test the clinical utility of two commonly administered PVTs within the real-world context utilizing actual patients/litigants to avoid common limitations associated with using a simulated-malingerers design (e.g., inadequate knowledge about the disorder and lack of incentive for accurate simulating). However, this methodology does limit the findings in that knowledge about whether performance was truly valid or invalid is unknown. We included failure rates of other PVTs to support our suggestion that WMT failure was due to invalid performance rather than false positive errors. The majority of participants who passed TOMM but failed WMT (65%) failed at least one other PVT, suggesting failure was due to invalid performance. Notably, other research conducted with the present sample found that nearly all participants (94.7%) who failed the WMT failed at least one embedded validity indicator as well (Erdodi et al., 2019). However, without definitively knowing which participants were giving invalid performance, the possibility of false positive errors cannot be ruled out for certain. It is possible that the 25 participants who only failed the WMT consisted of a mixed-validity group with both true and false positives.

In future studies, the method used for this study could be extended to analyzing discrepancies between other PVTs. Some PVTs may turn out not to be good predictors of performance validity. Failure on these PVTs might not be linked with a global suppression of neuropsychological test scores. Only by examining the actual suppression of neuropsychological test scores can we apply this aspect of the acid test to other PVTs. Future studies could examine WMT versus TOMM performance in further detail to discover, for example, whether there are differences in the suppression of cognitive test scores in those with specific diagnoses, such as TBI or chronic pain. Further, it is only by examining other research data in further studies that we can show that the WMT is, in fact, uncorrelated with age, intelligence, and with impairment due to brain disease (see Green & Allen, 1999; Green & Flaro, 2019). Despite these properties, WMT does explain a large portion of the variance in the overall test battery (Green et al., 2001).

The global suppression in test scores across multiple cognitive domains seen in participants who failed the WMT but passed the TOMM indicates that the WMT is more sensitive than the TOMM in detecting performance validity. This study provides a useful model for examining the clinical utility of other PVTs.

ORCID

Rayna Hirst http://orcid.org/0000-0001-5586-9641

References

Allen, L. M., & Green, P. (1999). *Severe traumatic brain injury performance on CARB and the WMT. Supplement to the CARB and WMT manuals*. Durham, NC: CogniSyst, Inc.

Batt, K., Shores, E. A., & Chekaluk, E. (2008). The effect of distraction on the Word Memory Test and Test of Memory Malingering performance in patients with a severe brain injury. *Journal of the International Neuropsychological Society*, *14*(6), 1074–1080. doi:10.1017/S135561770808137X

Bauer, L., O'Bryant, S. E., Lynch, J. K., McCaffrey, R. J., & Fisher, J. M. (2007). Examining the Test of Memory Malingering Trial 1 and Word Memory Test Immediate Recognition as screening tools for insufficient effort. *Assessment*, *14*(3), 215–222. doi:10.1177/1073191106297617

Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Beck depression inventory* (2nd ed.). San Antonio, TX: The Psychological Corporation.

Boone, K. B. (2007). A reconsideration of the Slick et al., (1999) criteria for malingered neurocognitive dysfunction. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective*. New York, NY: Guilford Publications, Inc.

Bush et al. (2005). Symptom validity assessment: Practice issues and medical necessity. *The Clinical Neuropsychologist*, *20*(4), 419–426.

Carone, D. A. (2014). Young child with severe brain volume loss easily passes the Word Memory Test and Medical Symptom Validity Test: Implications for mild TBI. *The Clinical Neuropsychologist*, *28*(1), 146–162. doi:10.1080/13854046.2013.861019

Constantinou, M., Bauer, L., Ashendorf, L., Fisher, J. M., & McCaffrey, R. J. (2005). Is poor performance on recognition memory effort measures indicative of generalized poor performance on neuropsychological tests? *Archives* *of Clinical Neuropsychology*, *20*(2), 191–198. doi:10.1016/j.acn.2004.06.002

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *California verbal learning test – second edition: Adult version manual.* San Antonio, TX: The Psychological Corporation.

Delis, D. C., & Wetter, S. R. (2007). Cogniform Disorder and Cogniform Condition: Proposed diagnoses for excessive cognitive symptoms. *Archives of Clinical Neuropsychology*, *22*(5), 589–604. doi:10.1016/j.acn.2007.04.001

Eglit, G., Lynch, J. L., & McCaffrey, R. J. (2016). Not all performance validity tests are created equal: The role of recollection and familiarity in the Test of Memory Malingering and Word Memory Test. *Journal of Clinical and Experimental Neuropsychology*, *39*(2), 173–189. doi:10.1080/13803395.2016.1210573

Erdodi, L. A., Green, P., Sirianni, C. D., & Abeare, C. A. (2019). The myth of high false-positives on the Word Memory Test in mild TBI. *Psychological Injury and Law*, *12*(2), 155–169. doi:10.1007/s12207-019-09356-8

Fazio, R. L., Sanders, J. F., & Denney, R. L. (2015). Comparison of performance of the Test of Memory Malingering and Word Memory Test in a criminal forensic sample. *Archives of Clinical Neuropsychology*, *30*(4), 293–301. doi:10.1093/arclin/acv024

Gervais, R. O., Rohling, M. L., Green, P., & Ford, W. (2004). A comparison of WMT, CARB, and TOMM failure rates in non-head injury disability claimants. *Archives of Clinical Neuropsychology*, *19*(4), 475–487. doi:10.1016/j.acn.2003.05.001

Goodrich-Hunsaker, N. J., & Hopkins, R. O. (2009). Word memory test performance in amnesic patients with hippocampal damage. *Neuropsychology*, *23*(4), 529–534. doi:10.1037/a0015444

Gorham, D. R. (1956). *Gorham's proverbs test.* Missoula, MT: Psychological Tests Specialists.

Green, P. (1996). *Manual for the Story Recall Test (ACT).* Edmonton, AB: Green's Publishing.

Green, P. (1997). *Manual for the emotional perception test.* Edmonton, AB: Green's Publishing.

Green, P. (2004). *Memory complaints inventory*. Edmonton, AB: Green's Publishing.

Green, P., & Allen, L. M. (1999). *Performance of neurological patients on CARB and the WMT. Supplement to the CARB and WMT manuals*. Durham, NC: CogniSyst, Inc.

Green, P., & Iverson, G. L. (2001). Effects of injury severity and cognitive exaggeration on olfactory deficits in head injury compensation claims. *NeuroRehabilitation*, *16*(4), 237–243.

Green, P., & Flaro, L. (2016). Results from three performance validity tests in children with intellectual disability. *Applied Neuropsychology: Child*, *5*(1), 25–34.

Green, P., & Flaro, L. (2019). Performance validity test failure predicts suppression of neuropsychological test results in developmentally disabled children. *Applied Neuropsychology – Child*, 1–17. doi:10.1080/21622965.2019.1604342

Green, P., Flaro, L., & Courtney, J. (2009). Examining false positives on the Word Memory Test in adults with mild traumatic brain injury. *Brain Injury*, *23*(9), 741–750. doi:10.1080/02699050903133962

Green, P., Rohling, M. L., Lees-Haley, P. R., & Allen, L. I. (2001). Effort has a greater effect on test scores than severe brain injury in compensation claimants. *Brain Injury*, *15*(12), 1045–1060. doi:10.1080/02699050110088254

Greiffenstein, M. F., Greve, K. W., Bianchini, K. J., & Baker, W. J. (2008). Test of Memory Malingering and Word Memory Test: A new comparison of failure concordance rates. *Archives of Clinical Neuropsychology*, *23*(7–8), 801–807. doi:10.1016/j.acn.2008.07.005

Greve, K. W., Ord, J., Curtis, K. L., Bianchini, K. J., & Brennan, A. (2008). Detecting malingering in traumatic brain injury and chronic pain: A comparison of three forced- choice symptom validity tests. *The Clinical Neuropsychologist*, *22*(5), 896–918. doi:10.1080/13854040701565208

Halstead, W. C. (1947). *Brain and intelligence: A quantitative study of the frontal lobes*. Chicago, IL: University of Chicago Press.

Heyanka, D. J., Thaler, N. S., Linck, J. F., Pastorek, N. J., Miller, B., Romesser, J., & Sim, A. H. (2015). A factor analytic approach to the validation of the Word Memory Test and Test of Memory Malingering as measures of effort and not memory. *Archives of Clinical Neuropsychology*, *30*(5), 369–376. doi:10.1093/arclin/acv025

Iverson, G. L. (2003). Detecting malingering in civil forensic evaluations. In A. M. Horton & L. C. Hartlage (Eds.), *Handbook of forensic neuropsychology* (pp. 137–177). New York, NY: Springer Publishing Company.

Kløve, H. (1963). Clinical neuropsychology. In F. M. Forster (Ed.), *The medical clinics of North America*. New York, NY: WB Saunders.

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

Lau, L., Basso, M. R., Estevis, E., Miller, A., Whiteside, D. M., Combs, D., & Arentsen, T. J. (2017). Detecting coached neuropsychological dysfunction: A simulation experiment regarding mild traumatic brain injury. *The Clinical Neuropsychologist*, *31*(8), 1412–1431. doi:10.1080/13854046.2017.1318954

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. doi:10.1080/13854046.2015.1087597

McSweeney, A. J., Naugle, R. I., Chelune, G. J., & Lüders, H. (1993). T scores for change:" An illustration of a regression approach to depicting change in clinical neuropsychology. *Clinical Neuropsychologist*, *7*(3), 300–312. doi:10.1080/13854049308401901

Merten, T., Bossink, L., & Schmand, B. (2007). On the limits of effort testing: Symptom validity tests and severity of neurocognitive symptoms in nonlitigant patients. *Journal of Clinical and Experimental Neuropsychology*, *29*(3), 308–318. doi:10.1080/13803390600693607

Meyers, J. E., & Meyers, K. R. (1995). *Rey Complex Figure Test and Recognition Trial: Professional manual*. Lutz, FL: Psychological Assessment Resources.

Mossman, D., Wygant, D. B., & Gervais, R. O. (2012). Estimating the accuracy of neurocognitive effort measures in the absence of a "gold standard". *Psychological Assessment*, *24*(4), 815–822. doi:10.1037/a0028195

Rees, L. M., Tombaugh, T. N., Gansler, D. A., & Moczynski, N. P. (1998). Five validation experiments of the Test of Memory Malingering (TOMM). *Psychological Assessment*, *10*(1), 10–20. doi:10.1037/1040-3590.10.1.10

Reitan, R. M. (1979). *Manual for administration and scoring: Trail Making Test*. Dallas, TX: Neuropsychology Center.

Reitan, R. M., & Wolfson, D. (1993). *The Halstead-Reitan Neuropsychological Test Battery: Theory and critical interpretation*. Tucson, AZ: Neuropsychology Press.

Rienstra, A., Groot, P. F. C., Spaan, P. E. J., Majoie, C. B. L. M., Nederveen, A. J., Walstra, G. J. M., … Schmand, B. (2013). Symptom validity testing in memory clinics: Hippocampal-memory associations and relevance for diagnosing mild cognitive impairment. *Journal of Clinical and Experimental Neuropsychology*, *35*(1), 59–70. doi:10.1080/13803395.2012.751361

Ruff, R. M. (1996). *RFFT: Ruff Figural Fluency Test: Professional manual*. Lutz, FL: Psychological Assessment Resources.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York, NY: Oxford University Press.

Tan, J. E., Slick, D. J., Strauss, E., & Hultsch, D. F. (2002). How'd they do it? Malingering strategies on symptom validity tests. *The Clinical Neuropsychologist*, *16*(4), 495–505. doi:10.1076/clin.16.4.495.13909

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago, IL: Chicago University Press.

Tombaugh, T. N. (1996). *Test of memory malingering*. Toronto, Canada: MultiHealth Systems.

Wechsler, D. (1981). *Wechsler adult intelligence scale – revised*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997). *Wechsler memory scale – third edition*. San Antonio, TX: The Psychological Corporation.

Wilkinson, G. S. (1993). *Wide range achievement test* (3rd ed.). Wilmington, DE: Wide Range, Inc.