Archives of Clinical Neuropsychology
19 (2004) 693–702

Archives of CLINICAL NEUROPSYCHOLOGY

Detecting Symptom- and Test-Coached Simulators with the Test of Memory Malingering☆

Matthew R. Powell[a,b,*], Jeffrey D. Gfeller[b], Bryan L. Hendricks[c], Michael Sharland[b]

[a] *Department of Psychiatry and Psychology (W-11-B), Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA*
[b] *Saint Louis University, St. Louis, MO, USA*
[c] *University of Wisconsin–Marathon County, Wausau, WI, USA*

Accepted 10 March 2004

Abstract

The ability of the Test of Memory Malingering (TOMM; Tombaugh, 1996) to detect feigned-memory impairment was explored. The TOMM was administered to three groups: (a) a control group instructed to perform optimally, (b) a symptom-coached group instructed to feign memory problems after being educated about traumatic brain injury symptomatology, and (c) a test-coached group instructed to feign memory problems after being educated about test-taking strategies to avoid detection. The recommended cutoff scores (Tombaugh, 1996) on Trial 2 and the Retention Trial produced overall classification accuracy rates of 96%, with high levels of sensitivity and specificity. Although the symptom-coached group performed more poorly on the TOMM relative to the test-coached group, the test was equally sensitive in detecting suboptimal effort across the different coaching paradigms.

Keywords: Symptom validity; Effortful; Malingering; Coaching; Assessment; Forensics

☆ Portions of this study were presented at the 22nd Annual Conference of the National Academy of Neuropsychology, Tampa Bay, FL, October 2002.

* Corresponding author. Fax: +44-507-284-4158.
E-mail address: powell.matthew@mayo.edu (M.R. Powell).

0887-6177/$ – see front matter © 2004 National Academy of Neuropsychology.
doi:10.1016/j.acn.2004.04.001

1. Introduction

There has been an emergent body of research over the past decade focusing on the assessment and detection of poor effort within neuropsychological evaluations (Reynolds, 1998). Mild head injured patients involved in litigation have historically served as the primary target for empirical evaluation of poor effort (Binder & Rohling, 1996; Binder & Willis, 1991; Green, Iverson, & Allen, 1999; Langeluddecke & Lucas, 2003). More recent evidence indicates that a variety of other medical, litigious or psychiatric patient populations (e.g., disability claimants or chronic pain syndromes) may also exhibit compromised effort during neuropsychological evaluation (Gervais, Green, Allen, & Iverson, 2001; Gervais et al., 2001; Mittenberg, Patton, Canyock, & Condit, 2002; Rohling, Green, Allen, & Iverson, 2002). Although base rates of symptom exaggeration for any group will vary widely due to methodological and sample characteristics (Reynolds, 1998), a recent survey of American Board of Clinical Neuropsychology members provides benchmark prevalence estimates of suboptimal effort for a variety of patient populations (Mittenberg et al., 2002).

Given the existence of compromised effort and symptom magnification during neuropsychological evaluation, energy has been devoted to the development and validation of symptom-validity tests (SVT). SVTs are measures designed to be insensitive to bonefide neurocognitive deficits and contribute valuable information to the diagnostic process by objectively assessing patient effort. SVTs often use a forced-choice paradigm and function on the premise that symptom exaggerators (a) perceive the task as more difficult than is actually the case, (b) perform worse than severely impaired clinical groups (e.g., severe traumatic brain injured traumatic brain injury patients), or (c) perform at a level worse than chance (Bianchini, Mathias, & Greve, 2001; Williams, 1998). Bianchini et al. (2001) provide a more comprehensive discussion of symptom-validity testing.

The TOMM (Tombaugh, 1996) is a 50-item, two-alternative forced-choice SVT of visual recognition memory investigated in the current study. In a validation study, Tombaugh (1997) administered three trials of the TOMM to community volunteers and various clinical groups (e.g., cognitively impaired, aphasic, traumatic brain injury, and demented patients). Results support the TOMMs insensitivity to neurological dysfunction; 95% of all nondemented patients (91% of all patients), and 100% of normal controls, performed above 45/50 correct (the suggested cutoff) on Trial 2 of the TOMM. The TOMM was also highly successful at differentiating between nonpatient and patient simulators and controls, and was sensitive to poor effort in litigious samples (Rees, Tombaugh, Gansler, & Moczynski, 1998; Tombaugh, 1997). The TOMM has also been shown to be insensitive to clinically significant depression (Rees, Tombaugh, & Boulay, 2001).

As the efficacy of SVTs is dependent on persons being naïve regarding their purpose, the utility of effort measures could be compromised by education or coaching being provided to clients by third parties prior to the evaluation process (Lees-Haley, 1997; Youngjohn, 1995). Such counseling may encourage symptom exaggeration, resistance and defensiveness during the assessment, or increase the sophistication of feigning symptoms. Given the possibility of clients altering their symptom presentation and exhibiting sophisticated test-taking behaviors on instruments following coaching, investigators have attempted to explore how coaching may influence performance on effort measures.

Frederick and Foster (1991) compared the performance of test-coached and naïve (uncoached) simulators against effortful responders on a forced-choice measure of nonverbal cognitive ability. They instructed the coached group to (a) get at least half of the answers correct, (b) answer the easy ones correctly, and (c) miss only the more difficult items. Results indicated that although most of the test-coached participants did not avoid detection (approximately 73% were classified as simulating malingering), more were undetected when compared to naïve simulators. Several other investigators provide evidence that coaching participants can lead to more sophisticated simulation behavior (e.g., Feldstein, Durham, Keller, Klebe, & Davis, 2000; Rose, Hall, Szalda-Petree, & Bach, 1998).

DiCarlo, Gfeller and Oliveri (2000) explored the ability of the Category Test (CT) to identify nonoptimal effort under different coaching paradigms. They administered the CT to a group who received information about common symptoms of head injury (symptom-coached group), a group who received this symptom coaching plus test coaching (instructions from Frederick & Foster, 1991), and to normal controls and traumatic brain injury patients. Results suggested that the combined symptom- and test-coached participants were more often misclassified as performing optimally on various CT indices than were the symptom-coached individuals.

Coaching (symptom-and/or test-coaching) prior to the administration of neuropsychological tests appears to lead to more sophisticated and evasive behavior by simulators. With respect to type of coaching, test-coaching (and combined test and symptom-coaching) may provide the most helpful information for those who are attempting to avoid detection (DiCarlo et al., 2000; Gfeller & Morasco, 2001). One might speculate that test coaching provides more practical and tangible strategies relative to symptom-and test-coaching for simulators to adopt during evaluations to help avoid detection.

The present study further explored the effects of symptom-and test-coaching using the TOMM, and is unique in that it examined the influence of two independent types of coaching strategies on participant performance. Previous studies exploring the TOMMs effectiveness in detecting poor effort only offered what appeared to be minimal test-coaching to their simulator participants (Rees et al., 1998; Tombaugh, 1997).

Several research hypotheses or questions were explored. First, the normal controls were predicted to obtain the highest raw scores on the three trials of the TOMM, while the symptom-coached group was expected to have the lowest raw scores. Second, it was predicted that the TOMM would be highly sensitive and specific in classifying simulators and nonsimulators. Classification accuracy statistics (i.e., sensitivity, specificity, and overall classification accuracy) were determined for Trial 2 and the Retention Trial of the TOMM. Third, based on the findings from previous coaching literature, it was predicted that significantly more test-coached simulators would go undetected compared to symptom-coached simulators.

2. Method

2.1. Measures

Demographic questionnaire: Participants reported their age, current level of education, gender, ethnicity/race, country of birth, region of the country where they were educated, and

area of residence. They were also asked to rate their degree of compliance with experimental instructions.

North American Adult Reading Test (NAART; Blair & Spreen, 1989): The NAART is a reading recognition test that consists of 61 irregular words that is often used to estimate an individual's level of intellectual functioning (Spreen & Strauss, 1998). Verbal, performance, and full scale IQ estimates were generated using equations from Blair and Spreen (1989).

Test of Memory Malingering (Tombaugh, 1996): The TOMM is a 50-item, forced-choice test of visual memory designed to assess feigned-memory impairment. The test contains line drawings (targets) of common objects, each presented for 3 s, followed by three trials of recognition memory testing. Trials 1 and 2 are tests of immediate recognition memory (learning trials), while the Retention Trial is a test of delayed recognition memory (following a 20 min delay). During each trial, the examinee chooses between each target (correct object) and a distracter object.

Three tests were used as fillers between Trial 2 and the Retention Trial of the TOMM: Rey Auditory Verbal Learning Test (RAVLT; Rey, 1964), Trail Making Test Forms A and B (Reitan & Wolfson, 1992), and the Wechsler Adult Intelligence Scale, Third Edition (WAIS-III) Digit Span Subtest (Wechsler, 1997). These filler measures were selected because they are common neuropsychological tasks and adequately filled the 20 min delay needed for the TOMM.

2.2. Participants

Ninety-five undergraduate volunteers from the Saint Louis University subject pool were recruited for the present investigation. Any individuals who reported a history of severe psychopathology (e.g., psychosis, major depression), acquired neurocognitive risk factors (e.g., TBI, stroke), learning disorders (during a brief screening interview), or who were not native English speakers, were excluded. All participants received extra credit or course participation units for their participation.

Although data were collected from 95 participants, 15 participants (1 normal control, 6 symptom-coached and 8 test-coached participants) were excluded from the analyses because they reported low compliance with experimental instructions, consistent with recommendations outlined by Langeluddecke and Lucas (2003). Normal controls, who reported not trying their best on a dichotomous yes/no item, which asked if they gave their best effort during testing, were excluded from analyses. For the symptom- and test-coached participants, those reporting low motivation to feign cognitive impairment (a score of 3 or below on a 5-point Likert scale assessing one's motivation to simulate malingering) were excluded from the analyses. Justifying their exclusion, those participants reporting low motivation to fake across both simulating conditions performed significantly better than those reporting high motivation to fake on Trial 1 ($t[64] = 3.7$, $p < .001$), Trial 2 ($t[64] = 3.8$, $p < .001$), and on the Retention Trial ($t[64] = 3.8$, $p < .001$) of the TOMM.

Of the 80 participants (normal controls = 28, symptom coached = 27, test coached = 25) included in the analyses, 65% were female, 60% were freshmen in college, 92.5% were Caucasian, 80% were raised in the North Central region of the United States, and 71.3% grew up in an urban environment. The average age of the participants was 19.2 years.

2.3. Procedures

Participants meeting the inclusion criteria were randomly assigned to one of three experimental conditions: (a) a control group, (b) a symptom-coached simulation group, or (c) a test-coached simulation group. All participants were asked to imagine that they were in a motor vehicle accident. Members of the control group were told they suffered no injuries and were asked to try their best on all tests administered. Those in the symptom-coached group were asked to believably feign cognitive impairment during the assessment in an attempt to obtain imaginal financial compensation, and were given instructions describing typical symptomatology associated with mild traumatic brain injury (e.g., memory problems). Those in the test-coached simulation group were instructed to feign cognitive impairment during the assessment in an attempt to obtain imaginal financial compensation, and were given instructions describing effective test-taking strategies that would help them avoid detection (see test-coaching used by Frederick & Foster, 1991, outlined above). The specific instructions for the three groups are available from the first author.

Following the manipulation, each participant completed the following tasks in the order listed: (a) the TOMM Trial 1 and Trial 2, (b) the RAVLT (trials 1–5, Trial B, and the immediate recall portion), (c) the Trail Making Test Forms A and B, (d) the Digit Span Subtest from the WAIS-III, (e) the TOMM delayed Retention Trial, (f) the RAVLT delayed recall and recognition tests, and (g) ratings of the perceived difficulty of the TOMM before and after testing.

Trained examiners, blind to participants' group assignment, administered the assessment instruments. Following administration of the assessment instruments, all participants completed a post-investigation questionnaire to evaluate their compliance with experimental instructions.

3. Results

Chi-square analyses and one-way analyses of variance (ANOVA) were conducted to investigate group equivalency on important demographic variables and cognitive measures. All three groups were similar regarding gender, ethnicity, level of education, and region of country where they were raised. Likewise, there were no significant differences on verbal, performance, and full scale IQ scores derived from the NAART.

One-way ANOVAs with Tukey HSD post hoc analyses were conducted to explore group differences on the three trials of the TOMM. All groups significantly differed on Trial 1 ($F[2, 77] = 112.79$, $p < .001$), Trial 2 ($F[2, 77] = 80.19$, $p < .001$), and the Retention Trial ($F[2, 77] = 81.61$, $p < .001$). As hypothesized, the symptom-coached group obtained the lowest mean score and the control group obtained the highest mean score (see Table 1). More specifically, however, the symptom-coached group performed below both the test-coached and control groups for each trial, while the control group performed better than both simulation groups for each trial. These results indicate that the experimental instructions and diverse coaching paradigms had a significant impact on participants' absolute performance on all trials of the TOMM.

The overall classification accuracy, sensitivity and specificity of the TOMM was then calculated using the recommended cutoff score of 45 (Tombaugh, 1996) on Trial 2 and the Retention

Table 1
Comparison of the normal control, symptom-coached, and test-coached groups on the TOMM

	Normal controls		Symptom coached		Test coached		Significance	
Trial	*M*	S.D.	*M*	S.D.	*M*	S.D.	*F*	*p*
Trial 1	48.82[a]	1.72	26.56[b]	8.49	32.32[c]	4.81	112.79	<.001
Trial 2	49.96[a]	0.19	26.67[b]	10.40	34.64[c]	6.10	80.19	<.001
Retention Trial	49.93[a]	0.38	26.70[b]	9.94	34.52[c]	6.63	81.61	<.001

Note. Values with different superscript letters indicate statistically significant differences between groups.

Trial of the TOMM (see Table 2). Across both trials, the TOMM demonstrated an exceptional overall classification accuracy of 96.3% (simulation and nonsimulation groups combined), with a specificity of 100% (nonsimulation group) and a sensitivity of 94.2% (with both simulation groups combined). Only three participants from the simulation group were misclassified as giving adequate effort using this method. No normal controls were misclassified. For both trials, 92.6% of the symptom-coached participants and 96% of the test-coached participants were correctly classified. Two symptom-coached participants and one test-coached participant were misclassified as performing optimally. The classification accuracy rates of the TOMM in this study are compatible with rates identified in other investigations (Rees et al., 1998; Tombaugh, 1997).

The impact of different types of coaching on simulation behavior was also explored. Consistent with previous research, it was hypothesized that significantly more symptom-coached participants would be detected than test-coached participants. Inspection of Table 2 indicates this hypothesis was not correct, as a marginally higher percentage of test-coached participants versus symptom-coached participants were detected. Chi-square analyses for Trial 2 and the Retention Trial revealed the frequency of detection across coaching groups did not significantly differ.

Table 2
Percentage of participants correctly classified by recommended cutting score of 45 on the TOMM

Trial	Normal controls (%)	Symptom coached (%)	Test coached (%)	Combined simulators (%)[a]	Total (%)[b]
Current study					
Trial 2	100.0 ($n = 28/28$)	92.6 ($n = 25/27$)	96.0 ($n = 24/25$)	94.2 ($n = 49/52$)	96.2 ($n = 77/80$)
Retention Trial	100.0 ($n = 28/28$)	92.6 ($n = 25/27$)	96.0 ($n = 24/25$)	94.2 ($n = 49/52$)	96.2 ($n = 77/80$)
Rees et al. (1998)					
Trial 2	100.0	NA	NA	84.0	92.0
Retention Trial	100.0	NA	NA	88.0	94.0
Tombaugh (1997)					
Trial 2	100.0	NA	NA	100.0	100.0

[a] The classification accuracy of both test- and symptom-coached groups combined.
[b] Reflects the overall classification accuracy for all groups combined (simulation and nonsimulation groups).

The group means outlined in Table 1, however, suggest that the different coaching strategies did lead to significantly different response styles on the TOMM. As indicated, the test-coached participants performed significantly better on all trials of the TOMM than did the symptom-coached participants. While test coaching appeared to lead to a more sophisticated simulation style when compared to the symptom-coached participants (i.e., higher mean scores), the finding that the TOMM was equally successful in detecting the test- and symptom-coached participants suggests that the TOMM is an effective SVT that remains sensitive to the different simulation strategies explored here.

4. Discussion

The primary purpose of this study was to explore the TOMM's classification accuracy statistics under different types of coaching conditions. The impact of test-and symptom-coaching on simulation behavior was also explored.

The TOMM was very sensitive in detecting suboptimal effort while maintaining very high specificity (100% accuracy). The recommended cutoff score (Tombaugh, 1997) proved to be an effective criterion for classifying the experimental groups. Approximately 92% of the symptom-coached and 96% of the test-coached participants were correctly classified using either Trial 2 or the Retention Trial. As seen in Table 2, the TOMM's overall classification accuracy statistics were largely consistent with values reported in other studies (Rees et al., 1998; Tombaugh, 1997).

As Faust and Ackley (1998) indicate, a sensitive measure must predict group membership at a level greater than the population's base rate. In the present study, the overall base rate of poor effort was 65%. The fact that the TOMM demonstrated an overall classification accuracy of 96.2% suggests that it added incrementally to our predictions of poor effort within a simulation context with known base rates. While such results will not generalize perfectly to a clinical context with a poorly defined base rate of symptom exaggeration, it provides useful data for the decision-making process.

Sensitivity is the probability that a test detects a condition when it exists in a sample (true positive), while specificity is the probability that a test does not detect a condition when it in fact does not exist in a sample (true negative). Positive predictive value, however, refers to the probability that an individual actually has the condition when a test detects the condition, whereas negative predictive power is the probability that the individual does not have the condition when a test does not detect the condition. Contrasting predictive values with sensitivity and specificity rates makes it apparent that positive and negative predictive values provide probability estimates most relevant for individual patients (Smith, Cerhan, & Ivnik, 2003).

Regarding predictive values, the present results reveal that all 49 participants performing below the cutoff criterion on the TOMM were simulators. Conversely, while a total of 31 participants scored above the cutoff, only 29 of these 31 participants were members of the control group. Therefore, the TOMM demonstrated a positive predictive power of 100% (49/49 participants) and a negative predictive power of 90% (28/31 participants). These statistics indicate that we can be 90% confident that a person gave good effort when he or she scored

above the suggested cutoff value. On the other hand, when a participant scored below the cutoff, we can have 100% confidence that he or she performed suboptimally. While these classification rates are promising, one should not overgeneralize these values to other samples. Predictive values will not generalize perfectly to groups with different characteristics (e.g., clinical samples) and will be significantly affected by the base rate of poor effort within the sample under investigation (Smith et al., 2003).

It appears that the test-coached group engaged in less dramatic exaggeration of memory impairment than did the symptom-coached group, given their significantly higher performance on all trials of the TOMM. Despite this finding, the TOMM was equally successful at detecting participants from both simulation groups. This finding stands in contrast to other studies that found that invalidity indicators derived from the Category Test (DiCarlo et al., 2000) and the Validity Indicator Profile (Gfeller & Morasco, 2001) detected significantly fewer test-coached participants compared to symptom-coached participants. The fact that the TOMM's cutting score for separating optimal and suboptimal performance is high (45/50 correct) may explain why the TOMM appears to more successfully detect poor effort for the test-savvy-simulation groups when compared to other invalidity indicators.

Several methodological limitations in this study warrant discussion. First, the use of student simulators decreases the generalizability of this study. However, Haines and Norris (2001) provide evidence in support of simulation paradigms, indicating that student simulators performed more similarly to mild traumatic brain injury patients than did non-neurological patient simulators. Additionally, the use of student populations who have no financial incentive to simulate malingering may also limit generalizability. While research indicates that financial compensation does affect patients' performance in clinical contexts (Binder & Rohling, 1996; Binder & Willis, 1991), using small financial incentives with simulators may not significantly affect performance (Bernard, 1990). Lastly, readers should be aware that this study's use of simulators and normal controls without inclusion of clinical groups (e.g., memory disordered patients or at-risk malingering patients) likely inflated the classification accuracy statistics documented above (e.g., specificity). For example, few if any bright, normal controls without motivation to perform poorly and free of complex medical histories should perform below the TOMMs cutoff for poor effort.

The overall findings support the TOMM's utility in detecting symptom magnification, even for sophisticated simulators, and provide further validation of the cutting score recommended by Tombaugh (1996) for making judgments about effort. Additional research to further explore classification accuracy rates of the TOMM with patient populations in various clinical contexts is warranted. Furthermore, additional studies focusing on the effects of coaching are needed. Different levels and types of symptom- and test-coached conditions could be formulated and evaluated using the TOMM or other SVTs.

Acknowledgments

The authors gratefully acknowledge Angie Capps and Nicole Salazar for their assistance with data collection, and Michael Oliveri and Donna LaVoie for their helpful comments throughout the design and implementation of this work.

References

Bernard, L. C. (1990). Prospects for faking believable memory deficits on neuropsychological tests and the use of incentives in simulation research. *Journal of Clinical and Experimental Neuropsychology*, *12*, 715–728.

Bianchini, K. J., Mathias, C. W., & Greve, K. W. (2001). Symptom validity testing: A critical review. *The Clinical Neuropsychologist*, *15*, 19–45.

Binder, L. M., & Rohling, M. L. (1996). Money matters: Meta-analytic review of the effects of financial incentives on recovery after closed-head injury. *American Journal of Psychiatry*, *153*, 7–10.

Binder, L. M., & Willis, S. C. (1991). Assessment of motivation after financially compensable minor head trauma. *Psychological Assessment*, *3*, 175–181.

Blair, J. R., & Spreen, O. (1989). Predicting premorbid IQ: A revision of the National Adult Reading Test. *Clinical Neurologist*, *3*, 129–136.

DiCarlo, M. A., Gfeller, J. D., & Oliveri, M. V. (2000). Effects of coaching on detecting feigned cognitive impairment with the Category Test. *Archives of Clinical Neuropsychology*, *15*, 399–413.

Faust, D., & Ackley, M. A. (1998). Did you think it was going to be easy? Some methodological suggestions for the investigation and development of malingering detection techniques. In C. R. Reynolds (Ed.), *Detection of malingering during head injury litigation: Critical issues in neuropsychology* (pp. 1–54). New York: Plenum Press.

Feldstein, S., Durham, R. L., Keller, F. R., Klebe, K., & Davis, H. P. (2000). Classification of malingerers on a nondeclarative memory task as a function of coaching and task difficulty. *American Journal of Forensic Psychology*, *18*, 57–78.

Frederick, R. I., & Foster, H. G. (1991). Multiple measures of malingering on a forced-choice test of cognitive ability. *Psychological Assessment*, *3*, 596–602.

Gervais, R. O., Green, P., Allen, L. M., & Iverson, G. L. (2001). Effects of coaching on symptom validity testing in chronic pain patients presenting for disability assessments. *Journal of Forensic Neuropsychology*, *2*, 1–19.

Gervais, R. O., Russell, A. S., Green, P., Allen, L. M., Ferrari, R., & Pieschl, S. D. (2001). Effort testing in patients with fibromyalgia and disability incentives. *The Journal of Rheumatology*, *28*, 1892–1899.

Gfeller, J. D., & Morasco, B. J. (2001). Detecting feigned cognitive impairment with the Validity Indicator Profile: The differential effects of coaching. *Archives of Clinical Neuropsychology*, *16*, 848–849.

Green, P., Iverson, G. L., & Allen, L. (1999). Detecting malingering in head injury litigation with the Word Memory Test. *Brain Injury*, *13*, 813–819.

Haines, M. E., & Norris, M. P. (2001). Comparing student and patient simulated malingerers' performance on standard neuropsychological measures to detect feigned cognitive deficits. *Clinical Neuropsychologist*, *15*, 171–182.

Langeluddecke, P. M., & Lucas, S. K. (2003). Quantitative measures of memory malingering on the Wechsler Memory Scale—Third edition in mild head injury litigants. *Archives of Clinical Neuropsychology*, *18*, 181–197.

Lees-Haley, P. R. (1997). Attorneys influence expert evidence in forensic psychological and neuropsychological cases. *Assessment*, *4*, 321–324.

Mittenberg, W., Patton, C., Conyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology*, *24*, 1094–1102.

Rees, L. M., Tombaugh, T. N., & Boulay, L. (2001). Depression and the Test of Memory Malingering. *Archives of Clinical Neuropsychology*, *16*, 501–506.

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*, 10–20.

Reitan, R. M., & Wolfson, D. (1992). *The Halstead–Reitan neuropsychological test battery: Theory and clinical interpretation* (2nd ed.). Tucson, AZ: Neuropsychology Press.

Rey, A. (1964). *L'examen clinique en psychologie*. Paris: Presses Universitaires de France.

Reynolds, C. R. (Ed.). (1998). *Detection of malingering during head injury litigation: Critical issues in neuropsychology*. New York: Plenum Press.

Rohling, M. L., Green, P., Allen, L. M., & Iverson, G. L. (2002). Depressive symptoms and neurocognitive test scores in patients passing symptom validity tests. *Archives of Clinical Neuropsychology*, *17*, 205–222.

Rose, F. E., Hall, S., Szalda-Petree, A. D., & Bach, P. J. (1998). A comparison of four tests of malingering and the effects of coaching. *Archives of Clinical Neuropsychology*, *13*, 349–363.

Smith, G. E., Cerhan, J. H., & Ivnik, R. J. (2003). Diagnostic validity. In D. S. Tulsky & D. Saklofske (Eds.), *Clinical interpretation of the WAIS-III and WMS-III* (pp. 273–301). San Diego, CA: Academic Press.

Spreen, O., & Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms and commentary* (2nd ed.). London: Oxford University Press.

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

Tombaugh, T. N. (1997). The Test of Memory Malingering (TOMM): Normative data from cognitively intact and cognitively impaired individuals. *Psychological Assessment*, *9*, 260–268.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale: Administration and scoring manual* (3rd ed.). San Antonio, TX: Psychological Corporation.

Williams, J. M. (1998). The malingering of memory disorder. In C. R. Reynolds (Ed.), *Detection of malingering during head injury litigation: Critical issues in neuropsychology* (pp. 105–132). New York: Plenum Press.

Youngjohn, J. R. (1995). Confirmed attorney coaching prior to neuropsychological evaluation. *Assessment*, *2*, 279–283.