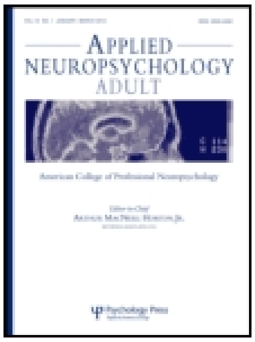

Applied Neuropsychology: Adult

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

Effort Test Performance in Clinical Acute Brain Injury, Community Brain Injury, and Epilepsy Populations

Natalie E. Hampson, Steven Kemp, Anthony K. Coughlan†, Chris J. A. Moulin & Bipin B. Bhakta

To cite this article: Natalie E. Hampson, Steven Kemp, Anthony K. Coughlan†, Chris J. A. Moulin & Bipin B. Bhakta (2014) Effort Test Performance in Clinical Acute Brain Injury, Community Brain Injury, and Epilepsy Populations, Applied Neuropsychology: Adult, 21:3, 183-194, DOI: 10.1080/09084282.2013.787425

To link to this article: https://doi.org/10.1080/09084282.2013.787425

Published online: 12 Sep 2013.

Submit your article to this journal

Article views: 440

View related articles

View Crossmark data

Citing articles: 7 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=hapn21

APPLIED NEUROPSYCHOLOGY: ADULT, *21*: 183–194, 2014
Copyright © Taylor & Francis Group, LLC
ISSN: 2327-9095 print/2327-9109 online
DOI: 10.1080/09084282.2013.787425

Effort Test Performance in Clinical Acute Brain Injury, Community Brain Injury, and Epilepsy Populations

Natalie E. Hampson
Department of Neuropsychology, Salford Royal NHS Foundation Trust, Salford, United Kingdom

Steven Kemp and Anthony K. Coughlan†
Department of Neuropsychology, St. James's University Hospital, Leeds, United Kingdom

Chris J. A. Moulin
LEAD CNRS UMR 5022, Dijon, France

Bipin B. Bhakta
Academic Department of Rehabilitation Medicine, Faculty of Medicine and Health, University of Leeds, Leeds, United Kingdom

Effort tests have become commonplace within medico-legal and forensic contexts and their use is rising within clinical settings. It is recognized that some patients may fail effort tests due to cognitive impairment and not because of poor effort. However, investigation of the base rate of failure among clinical populations other than dementia is limited. Forty-seven clinical participants were recruited and comprised three subgroups: acute brain injury ($N = 11$), community brain injury ($N = 20$), and intractable epilepsy ($N = 16$). Base rates of failure on the Word Memory Test (WMT; Green, 2003) and six other less well-validated measures were investigated. A significant minority of patients failed effort tests according to standard cutoff scores, particularly patients with severe traumatic brain injury and marked frontal-executive features. The WMT was able to identify failures associated with significant cognitive impairment through the application of profile analysis and/or lowered cutoff levels. Implications for clinical assessment, effort test interpretation, and future research are discussed.

Key words: *effort, malingering, symptom validity, WMT*

For the interpretation of neuropsychological tests to be valid, the examinee must apply full effort; otherwise, clinicians risk making a Type I error by concluding that someone is brain-damaged when this is not the case. There is currently considerable interest in the potential for feigning and exaggeration of symptoms as a threat to the validity of neuropsychological test results, particularly where personal gain is involved (e.g., Rogers, 2008a). To date, the majority of the literature has focused on populations in which feigning is thought to be most prevalent, such as medico-legal settings and disability benefit assessments.

---

†Dr. Tony Coughlan sadly died during the preparation of this article. The authors are extremely grateful for his contribution to this research.

Address correspondence to Natalie E. Hampson, Department of Neuropsychology, Clinical Sciences Building, Salford Royal Hospital, Stott Lane, Salford, M6 8HD, United Kingdom. E-mail: natalie.hampson@srft.nhs.uk

Although difficult to establish with certainty, numerous studies have sought to quantify the base rate of malingering across a variety of populations, with suggested rates of approximately 40% in those with mild traumatic brain injury (TBI) receiving disability payments or in litigation (Larrabee, 2003), up to 90% in criminal assessments (Ardolf, Denney, & Houston, 2007), and 8% in medical and psychiatric cases (Mittenberg, Patton, Canyock, & Condit, 2002), suggesting the economic burden of such false claims could be substantial. There are also clinical considerations: If feigning cannot be accurately identified, it may prevent or delay treatment for those genuinely in need. A label of malingering is also highly pejorative, and wrongful accusation or diagnosis can have a wide-ranging impact on the individual and their life.

Recently developed clinical guidelines have highlighted the importance of standardized symptom validity tests (commonly referred to as "effort" tests) within neuropsychological assessment. For example, the National Academy of Neuropsychology (NAN) position paper views such tests as central to understanding participant responses in neuropsychological assessment (Bush et al., 2005), and the British Psychological Society (BPS) also emphasizes the value and use of such assessments within UK practice (McMillan et al., 2009).

Belanger, Curtiss, Demery, Lebowitz, and Vanderploeg's (2005) meta-analysis of factors moderating outcome in mild TBI revealed that people not in litigation recovered within 3 months of injury, whereas those in litigation often continued to report symptoms or got worse over time. The recent Rohling et al. (2011) meta-analysis also showed no measurable effect of mild TBI at 3 months postinjury on neuropsychological tests, which are much more difficult than tests of effort. Effort tests are typically memory tests designed to appear demanding, but they are actually simple to complete, even for people who have substantial cognitive impairments. Therefore, it is not credible that a mild TBI could cause people to fail extremely easy symptom validity subtests or cause more impairment than that seen in children with developmental disabilities (Green, Flaro, & Courtney, 2009).

Numerous studies have identified that more people with mild TBI fail effort tests and score lower on other neuropsychological tests compared with those with more severe brain injury who pass such tests (e.g., Constantinou, Bauer, Ashendorf, Fisher, & McCaffrey, 2005; Green, Iverson, & Allen, 1999; Green, Rohling, Lees-Haley, & Allen, 2001; Meyers, Volbrecht, Axelrod, & Reinsch-Boothby, 2011; Moss, Jones, Fokias, & Quinn, 2003; Stevens, Friedel, Mehren, & Merten, 2008; West, Curtis, Greve, & Bianchini, 2011). For example, Fox (2011) reported no correlation between commonly used neuropsychological tests and objectively determined brain damage among those who failed effort tests. However, for those who passed effort tests, the expected relationship between cognitive test data and brain damage was found.

Therefore, poor effort seems to be the only reasonable explanation for people with a mild TBI to fail well-validated effort tests. However, questions remain regarding the level of failure in genuinely impaired populations. Although the recent NAN and BPS position papers provide guidance, there is still no gold standard to assess test effort. Without such a standard, it is difficult to assess validity adequately. Therefore, it is vital to know the accuracy of published cutoff scores if appropriate decisions are to be made regarding whether or not someone is putting forth their best effort. However, cutoff scores on effort tests have been criticized for increasing the probability that a person with genuine impairment could be classified as applying suboptimal effort (e.g., Greve, Ord, Curtis, Bianchini, & Brennan, 2008; Haines & Norris, 1995). Although most well-established symptom validity test manuals acknowledge the problem of high failure rates in truly impaired populations, this issue has often been managed by removing such populations from validation studies. This minimizes the very real concern that some people with genuine impairments will fail, and clinicians might wrongly conclude poor effort if tests are to be routinely used and guidance for their interpretation is not followed accurately. Therefore, it is important that the base rate of failure among clinical groups with more severe impairments is available.

The Word Memory Test (WMT) is one of the most popular and well-investigated measures of symptom validity currently available, with the authors stating that differences in scores cannot be explained on the basis of actual cognitive deficits apart from those with the most extreme forms of learning and memory impairments, such as those with dementia (Green, 2003; Green et al., 2001). Despite increasingly widespread use, relatively few studies have provided base rate data in clinical populations. Some researchers have also implied that people with less extreme impairments could fail effort tests for cognitive reasons (e.g., Batt, Shores, & Chekaluk, 2008; Bowden, Shores, & Mathias, 2006). However, such studies have been criticized due to their design and nonstandard administration procedures (e.g., Flaro, Green, & Robertson, 2007; Rohling & Demakis, 2010), including the fact that the correct procedure for interpreting the WMT was not applied appropriately, as only the first trial was utilized, so judgments about failure rates are likely to be inaccurate.

In addition, alongside the research involving those in litigation, some studies have attempted to provide data from clinical populations who have significantly impaired cognitive functioning. For example, Merten, Bossink, and

Schmand (2007) reported a base rate failure of 50% to 58% when using standard cutoff scores and no additional investigations in 24 people with "clinically obvious symptoms," as judged by a clinician as demonstrating bradyphrenia, repetitive speech, and word-finding difficulties.

Despite Flaro et al. (2007) and Rohling and Demakis (2010) providing evidence that people with milder injuries cannot fail the WMT for cognitive reasons, and that incomplete WMT administration and interpretation may be inflating failure rates in genuinely impaired participants, the Bowden et al. (2006) and Merten et al. (2007) studies do indicate a need for further research in more severely impaired clinical populations, and their results have led to recent attempts to refine the cutoff scores and methods of test interpretation for certain clinical groups.

Greve et al. (2008) studied the classification accuracy of the WMT, the Portland Digit Recognition Test (Binder, 1993), and the Test of Memory Malingering (TOMM; Tombaugh, 1996) in detecting malingering in 109 people with TBI. The authors concluded that the WMT misclassified 19 of these participants, and that although the measure was sensitive to malingering when using the established cutoffs, this led to reduced specificity and unacceptably high rates of false positives. As such, adjusted cutoff scores were proposed associated with 90% and 98% specificity levels. Axelrod and Schutte (2011) also identified lower specificity in a clinical sample of 153 participants when using a shorter version of the WMT (Medical Symptom Validity Test [MSVT]; Green, 2004) in comparison to the TOMM and California Verbal Learning Test-Second Edition list-learning task (Delis, Kramger, Kaplan, & Ober, 2000), with reduced cutoff levels again being suggested.

However, the Greve et al. (2008) sample included participants with mild TBI, so it is possible that those who were judged as false positives are actually false negatives, given that many of the participants were not severely impaired and that some may have had incentives to malinger that were not considered within the study criteria. It could also be argued that the TOMM may be less sensitive than the WMT, rather than the WMT having lower specificity than the TOMM. For example, Blaskewitz, Merten, and Kathmann (2008) reported the TOMM only identified 68% of children who were simulating, versus a detection rate of 90% with the MSVT, indicating different sensitivity rates across the two measures. Green (2011) and Armistead-Jehle and Gervais (2011) have also identified similar findings using the nonverbal version of the MSVT (NV-MSVT; Green, 2008).

As such, other researchers have proposed an alternative approach to evaluating performance on the WMT and related measures (i.e., the MSVT and NV-MSVT) in people with established impairments that includes the retention of published cutoff levels and analysis of the profile of scores across the various subtests (Howe & Loring, 2009; Singhal, Green, Ashaye, Shankar, & Gill, 2009). For example, Henry, Merten, Wolf, and Harth (2010) identified a maximum false-positive rate of 1 out of 65 in 44 people with neurological conditions and 21 people with dementia diagnoses when applying profile analysis to the NV-MSVT. More recently, Green, Montijo, and Brockhaus (2011) also highlighted using profile analysis with the WMT in people being screened for dementia, and they reported much fewer failures when compared with interpretation using standard cutoff scores alone.

Although profile analysis is now included in the advanced interpretation computer program (Green, 2009) designed to allow users to automatically apply these methods to the interpretation of WMT data, profile analysis represents a recent finessing of effort test interpretation and is not yet routinely used clinically. For example, Axelrod and Schutte's (2010) argument that profile analysis is not useful was based on a potentially incomplete understanding of profile analysis, as a significant number of their participants had a diagnosis of mild TBI with no neurological impairment. Only when a person has an established clinical condition known to produce severe cognitive impairments would a "genuine cognitive impairment profile" be interpreted as reflecting legitimate severe impairment. Although Axelrod and Schutte (2011) do go on to acknowledge this as an issue, by extension, profile analysis may be incorrectly applied, or not applied at all, in clinical practice with genuinely impaired populations.

For good reasons, use of symptom validity testing is becoming standard clinical practice, with recent UK and U.S. guidelines promoting their use. There is limited research regarding the performance of clinical populations with established diagnoses other than dementia. Hence, we know more about the sensitivity of these measures than we know about their specificity, and clinicians are now at risk for making a Type II error (i.e., assuming insufficient test effort and overlooking brain damage).

The current research is an exploratory study intended to provide initial data regarding effort test performance among three distinct and clinically prevalent groups. We compared performance relative to standard cutoffs on the WMT (Green, 2003), along with several other less well-established measures that have been applied as tests of effort. Additionally, the study considers whether applying adjusted cutoff scores or "genuine cognitive impairment profile" analysis affects the failure rate, as suggested by a number of recent authors.

METHOD

Participants

A total of 47 participants (32 men, 15 women) were recruited during a 10-month period, encompassing three groups: Group 1 (acute brain injury) included 11 inpatients on a National Health Service postacute neurological rehabilitation ward; Group 2 (community brain injury) included 20 people in residential community rehabilitation services run by a registered charity; Group 3 (intractable epilepsy) included 16 outpatients attending a National Health Service regional center for epilepsy surgery (see Table 1).

Participants were recruited in accord with the stipulations of the ethics committee that granted study approval. Namely, potential participants were identified by the treating consultant following inclusion/exclusion criteria. All participants were required to be older than 18 years of age and have a definitive diagnosis of brain injury or epilepsy (identified through computed tomography and/or magnetic resonance imaging brain scans and/or brain electroencephalography [EEG]). Participants were also required to have a good grasp of the English language. Potential participants were excluded if the clinical team determined that they did not have the mental capacity to consent, if the team judged them as not capable of participating in the study for any reason, if they were using any substances that might influence cognitive test scores (e.g., current drug or alcohol misuse), or if they had experienced a seizure in the last 24 hr. The presence of comorbidity judged by the clinical team to be a significant influence (e.g., serious somatic or psychiatric illness) or visual, motor, or language dysfunction that precluded administration of computerized tasks (e.g., hemiparesis, aphasia) were also used as the basis for exclusion from the study.

TABLE 1
Brain Injury Data for the Three Participant Groups

	Group 1 (Inpatient) N = 11	*Group 2 (Community) N = 20*	*Group 3 (Epilepsy) N = 16*
Time since diagnosis			
Mean	1.1 months	11.9 years	22.9 years
Range	0–2 months	1–30 years	2–43 years
Nature of brain injury			
Hypoxic injury	—	3 (15.0%)	—
Traumatic injury	4 (36.3%)	12 (60.0%)	—
Stroke	2 (18.2%)	1 (5.0%)	—
Intracerebral hemorrhage	1 (9.1%)	1 (5.0%)	—
Subarachnoid hemorrhage	1 (9.1%)	1 (5.0%)	—
Infection	1 (9.1%)	2 (10.0%)	—
Tumor	2 (18.2%)	—	6 (37.5%)
Cavernoma	—	—	1 (6.2%)
Hippocampal sclerosis	—	—	9 (56.3%)
Brain lesion			
Diffuse	1 (9.0%)	9 (45%)	Bitemporal = 1 (6%)
Right hemisphere	2 (18.2%)	2 (10%)	Right Temporal = 7 (44%)
Left hemisphere	4 (36.4%)	2 (10%)	Left Temporal = 7 (44%)
Primarily frontal	4 (36.4%)	7 (35%)	Frontal = 1 (6%)

Exclusion criteria also included ongoing litigation (assessed via direct questioning of participants and/or confirmation by a treating professional). However, none of the participants required exclusion on this basis, likely due to all participants presenting for clinical reasons to UK National Health Services or being involved in receiving rehabilitation from a registered charity. None of the participants in the inpatient group were receiving any benefits for their injury, but all of the patients in the epilepsy group were receiving or had received some form of state benefit, ranging from disability payments to free prescription medications. The community participants had either already been awarded prior litigation payouts for their injuries, which were now being used to fund their care, or were having their care fees paid by local statutory services. In addition to having a confirmed significant neurological condition and no current litigation, the participants showed no signs of suboptimal performance, uncooperativeness, or negative response bias within their clinical presentation or history. This is in line with Henry et al.'s (2010) definition of "bona fide" neurology patients.

Potential participants were contacted by the research team, provided with the study information sheet, and given time to consider whether to participate in the study. A total of 56 participants were identified by their treating consultant; of the 9 who declined to participate, 4 were from Group 1, 1 was from Group 2, and 4 were from Group 3. No external incentive was provided to any of the participants.

As can be seen in Table 1, location of brain injury varied depending on the participant's subgroup. The epilepsy sample had focal lesions (mainly left or right temporal lesions), whereas the community brain injury and acute brain injury subgroups had more diffuse and anterior brain damage. Due to the obvious overlap between brain injury and epilepsy, only those participants with epilepsy as their primary presenting problem, or due to an in-situ tumor resulting in monitoring at an epilepsy clinic, were included in the epilepsy group. None of the patients in the acute brain injury group were judged to be in posttraumatic amnesia (PTA) by the treating consultant. All patients with TBI in Group

1 and Group 2 had PTA for more than 24 hr and Glasgow Coma scores <12/15.

All inpatient participants were receiving 24-hr supervision. All community participants were also receiving significant amounts of support and supervision, with most in 24-hr supported accommodation. None of the participants with epilepsy were receiving any additional organized support or supervision during their daily activities, though family members were often providing informal support.

Measures Used to Assess Effort

The Word Memory Test. The WMT (Green, 2003) is based on a computerized administration procedure involving an Immediate Recognition trial and a Delayed Recognition trial, and the Consistency Rating can be calculated on the basis of these two results. The Multiple Choice, Paired Associate, and Free Recall tasks can also be administered to assess memory ability and to calculate the easy–hard difference used in profile analysis.

The Coin-in-Hand Test. In the Coin-in-Hand Test (Kapur, 1994), participants are shown a coin, and are then asked to recall which hand it was in following a short distraction task. The procedure is repeated for 10 trials, and sensitivity and specificity data have been reported when comparing controls simulating brain injuries and those with legitimate brain injuries (Cochrane, Baker, & Meudell, 1998; Kelly, Baker, van den Broek, Jackson, & Humphries, 2005).

Autobiographical Memory Index. This test is a modified version of Wiggins and Brandt's (1988) original test (Kelly et al., 2005). It contains 11 questions concerning general information about the participant, scored as either plausible or incorrect, and was found to differentiate between simulators and nonsimulating participants.

Digit–Symbol Coding. This measure is a subtest of the Wechsler Adult Intelligence Scale-Third Edition (Wechsler, 1997a) and involves participants copying symbols paired with numbers within a time limit. It has been used to assess effort in research by Trueblood (1994) and Inman and Berry (2002).

Mental Control. The Mental Control Test is an optional measure from the Wechsler Memory Scale-Third Edition (Wechsler, 1997b) and includes eight tasks involving the retrieval and manipulation of overlearned information. It has been successfully utilized as a measure of effort by Kelly et al. (2005) and Kemp et al. (2008).

The Short Recognition Memory Test for Faces. The Faces subtest of the Camden Memory Test (Warrington, 1996) requires presentation of 30 faces, and recognition is assessed using a two forced-choice format. Cutoff for failure on the task was taken as below the fifth age-related percentile for the oldest normative age group. This is in line with the research of Kemp et al. (2008), who used the Short Pictorial Recognition Memory Test from the Camden battery as a measure of effort.

Standard Neuropsychological Tests

Premorbid intellectual functioning. The Wechsler Test of Adult Reading (WTAR; Wechsler, 2001) assesses premorbid intellectual ability, with participants asked to correctly pronounce 50 phonetically regular and irregular words.

Speed of information processing. The Information Processing subtest of the BIRT Memory and Information Processing Battery (Coughlan, Oddy, & Crawford, 2007) involves presenting participants with a sheet of paper containing rows of numbers, and they must mark the second highest number in each row for 4 min. A test of hand speed is then administered to control for any motor difficulties.

Anxiety and depression. The Hospital Anxiety and Depression Scale (HADS) is a self-report questionnaire measure of emotional well-being consisting of 14 items used to screen for depression and anxiety (Zigmond & Snaith, 1983).

Procedure

Group 1 participants with acute brain injury were all tested on the hospital ward, and Group 2 community participants were all tested in their own home. Participants in Group 3 were attending an outpatient epilepsy clinic and were offered testing either at the hospital or during a home visit; 11 were tested at home and 5 were tested in the clinic. All participants were tested by the same examiner (NEH).

Each participant was administered the battery of tests in a standard order. For the community and epilepsy groups, administration order was: WTAR, HADS, WMT Immediate Recognition, Coin-in-Hand Test, Mental Control, Speed of Information Processing, Autobiographical Memory Index, Digit–Symbol Coding, WMT Delayed Recognition, WMT memory components (Multiple Choice, Paired Associate, Free Recall), and the Camden Memory Test for Faces. The total length of testing was approximately 60 min. The acute brain injury group received an abbreviated test battery, and effort tests other than the principal effort measures (i.e., WMT Immediate and Delayed Recognition trials) were omitted. The decision to shorten testing with the inpatients was a pragmatic one governed by clinical presentation and the constraints of the ward environment. However, even abbreviated effort test data were thought to be valuable given the recency of the brain injury and that these patients were too unwell for discharge home.

Data were analyzed using the Statistical Package for the Social Sciences Version 16 for Windows.

RESULTS

Parametric statistics are reported for all available data, as all findings based on non-normally distributed data were supported by both parametric and nonparametric analyses. Variables were divided into discrete categories based on the median value when continuous scores were utilized.

As Table 2 shows, no participants were identified as within the learning disability range ($IQ < 70$) prior to injury, as assessed by the WTAR. The community group had poorer memory scores than did the epilepsy group on the WMT Paired Associate, $t(32) = -2.43$, $p = .021$, and Free Recall tests, $t(32) = 3.14$, $p = .004$. The overall mean scores for anxiety and depression are below cutoff levels for caseness. However, 23.4% of the total sample scored at the cutoff or greater for significant levels of anxiety (acute = 27.3%; community = 20%; epilepsy = 62.5%), and 36.2% of the overall sample scored at the cutoff or greater on measures of depression (acute = 27.3%; community = 15%; epilepsy = 31.2%). More people in the epilepsy group scored above the clinical cutoff for anxiety compared with the other two groups, *likelihood ratio,* $p = .030$.

TABLE 2
Mean, Standard Deviation, and Range of Scores for Demographic, Psychological, and Neuropsychological Data

Variable	*Group 1 (Inpatient)* *Mean (SD)* *Range*	*Group 2 (Community)* *Mean (SD)* *Range*	*Group 3 (Epilepsy)* *Mean (SD)* *Range*
Age (in years)	54.6 (13.7) 26–73	47.0 (12.4) 26–66	36.6 (10.3) 19–57
Education (in years)	13.6 (2.4) 12–18	12 (1.5) 12–17	12.8 (1.4) 12–17
Anxiety	6.4 (5.3) 1–20	5.3 (4.5) 0–16	7.8 (2.5) 3–12
Depression	6.7 (5.0) 2–20	4.0 (3.2) 0–10	4.7 (3.9) 0–12
Premorbid IQ	99.7 (8.7) 82–115	99.2 (11.6) 81–115	91.5 (9.4) 81–109
Processing Speed	—	45.7 (14.8) 19–68	56.4 (20.7) 30–100
WMT-MC (%)	—	61.4 (18.8) 35–100	71.9 (14.2) 45–100
WMT-PA (%)	—	46.1 (21.8) 15–90	63.1 (18.7) 30–90
WMT-FR (%)	—	25.1 (15.3) 5–57.5	42.8 (17.5) 12.5–72.5

WMT-MC = Word Memory Test Multiple Choice measure; WMT-PA = Word Memory Test Paired Associate measure; WMT-FR = Word Memory Test Free Recall measure.

No participants scored below chance (using 90% confidence bands for a binomial distribution) on any effort measures used, and the majority of the participants passed the effort tests. However, a wide range of performance was observed. When using the standard cutoff scores provided in the manuals, the number of false-positive results produced low specificity rates (i.e., <90%) for all but the Coin-in-Hand test, with 13 of the 47 participants failing the Immediate and/or Delayed Recognition trials on the WMT.

Table 3 contains the pass and fail rates on measures used to assess effort across all participant subgroups. Overall failure rates on the WMT Immediate and/or Delayed Recognition trials were: 27.3% in the acute group, 35.0% in the community group, and 18.8% in the epilepsy group. Including the Consistency Rating as an effort test measure substantially increased these rates to 36.4% in the acute group, 45% for the community group, and 37.5% for the epilepsy group.

The community group failed the Autobiographical Memory Index at much higher rates than those of the epilepsy group, Fishers Exact Test, $p < .001$, and this pattern remained significant when scaled scores were analyzed, $t(32) = -5.78$, $p < .001$, with community participants significantly more likely to have lower scores compared with the epilepsy participants. This pattern of results was also found in relation to the Camden Memory Test for Faces, $t(32) = -3.08$, $p = .004$. As previously noted, participants in the acute group were only administered a subset of tests.

Within the acute brain injury group, complete WMT Immediate Recognition, Delayed Recognition, and Consistency Rating scores were unavailable for 1 participant, and 7 out of the remaining 10 participants passed all three trials of the WMT. Within the community group, complete scores were unavailable for 3 participants; 2 out of 17 passed all six of the administered effort tests, 9 out of 17 failed one test, 1 out of 17 failed two tests, and 5 out of 17 failed three or more tests. However, the results for this group are skewed by the disproportionately high number of people who failed the Autobiographical Memory Index. Finally, within the epilepsy group, only 1 participant had missing data. Therefore, 5 out of 15 passed all effort tests, 6 out of 15 failed one test, 4 out of 15 failed two tests, and none of the patients in the epilepsy group failed three or more tests.

TABLE 3
Performance on Effort Tests Across the Three Participant Subgroups

	Group 1 (Inpatient)	*Group 2 (Community)*	*Group 3 (Epilepsy)*
Test	*Pass (%)/Fail (%)* *Mean (SD)* *Range*	*Pass (%)/Fail (%)* *Mean (SD)* *Range*	*Pass (%)/Fail (%)* *Mean (SD)* *Range*
WMT-IR	8 (72.7%)/3 (27.3%)	14 (70.0%)/6 (30.0%)	15 (93.8%)/1 (6.2%)
	87.3 (11.0)	84.9 (15.4)	93.9 (5.3)
	60.0–100.0	42.5–100.0	82.5–100.0
WMT-DR	8 (72.7%)/2 (18.1%)	15 (78.5%)/4 (21.1%)	14 (87.5%)/2 (12.5%)
	90.5 (10.1)	87.9 (9.3)	92.2 (8.4)
	70.0–97.5	65.0–100.0	70.0–100.0
WMT-CR	7 (70.0%)/3 (30.0%)	11 (57.9%)/8 (42.1%)	10 (62.5%)/6 (37.5%)
	85.5 (8.8)	81.8 (13.9)	88.3 (10.4)
	67.5–97.5	47.5–100.0	67.5–100.0
C-I-H	—	19 (95.0%)/1 (5.0%)	15 (93.8%)/1 (6.2%)
		9.6 (1.1)	9.8 (0.7)
		5–10	8–10
MC	—	14 (73.7%)/5 (26.3%)	11 (73.3%)/4 (26.7%)
		18.2 (6.7)	20.5 (7.9)
		8–29	8–36
D–SC	—	13 (65.0%)/4 (20.0%)	12 (75.0%)/4 (25.0%)
		5.8 (2.3)	5.8 (1.6)
		3–12	4–9
AMI	—	4 (22.2%)/14 (77.8%)	16 (100.0%)/0 (0.0%)
		9.3 (1.1)	10.9 (0.3)
		8–11	10–11
CMTF	—	13 (72.2%)/5 (27.8%)	15 (93.8%)/1 (6.2%)
		19.5 (3.6)	22.8 (2.4)
		12–25	17–25

WMT-IR = Word Memory Test Immediate Recognition Trial; WMT-DR = Word Memory Test Delayed Recognition Trial; WMT-CR = Word Memory Test Consistency Rating; C-I-H = Coin-in-Hand Test; MC = Mental Control Test; D–SC = Digit–Symbol Coding Test; AMI = Autobiographical Memory Index; CMTF = Camden Memory Test for Faces.

Scores across the measures of effort were found to relate to a number of the demographic and injury variables. Particularly, those who scored above the cutoff for depression scored significantly lower on the Mental Control task, $t(32) = 2.51$, $p = .018$, and slower processing speed was also significantly associated with lowered scores on the Mental Control task, $t(31) = -2.67$, $p = .012$) and Digit–Symbol Coding task, $t(31) = -4.93$, $p < .001$).

Lowered scores on the Camden Memory Test for Faces were significantly associated with low scores on the WMT Multiple Choice, $t(32) = -2.40$, $p = .020$, and Paired Associates tasks, $t(32) = -2.92$, $p = .006$. Lower scores on the Autobiographical Memory Index were also significantly related to low WMT Multiple Choice scores, $t(32) = -2.40$, $p = .017$.

In addition, a relationship was identified between lower scores on the WMT Immediate Recognition trial and low memory scores on the WMT Multiple Choice, $t(32) = -4.43$, $p < .001$, Paired Associates, $t(32) = -6.79$, $p < .001$, and Free Recall trials, $t(32) = -3.52$, $p = .001$. Similar relationships were also identified between lower scores on the WMT Delayed Recognition trial and the Multiple Choice, $t(32) = -4.85$, $p < .001$, and Paired Associates tasks, $t(32) = -4.41$, $p < .001$.

WMT Interpretation Using Contemporary Adjusted Calculations

As Table 4 shows, analyses of the current data using the cutoffs associated with specificity levels of approximately 98% for TBI (Greve et al., 2008) resulted in significantly reduced WMT failure rates across all of the current groups; Immediate Recognition failures were observed in two participants from the community group and one patient from the acute brain injury group. None of the patients in the epilepsy group failed the Immediate Recognition trial when using this cutoff. Only one participant from the community group failed on the basis of the Consistency Rating, and no participants across any groups failed the Delayed Recognition trial.

WMT pass and fail rates within the current data using Green et al.'s (2011) profile algorithm alongside the standard cutoffs can also be found in Table 4. The algorithm for the WMT is calculated only for people

TABLE 4
WMT Failure Rates Utilizing Standard Cutoffs, Adjusted Cutoffs, and Profile Analysis

	Group 1 (Inpatient)			*Group 2 (Community)*			*Group 3 (Epilepsy)*		
WMT Pass/Fail Calculation	*WMT-IR*	*WMT-DR*	*WMT-CR*	*WMT-IR*	*WMT-DR*	*WMT-CR*	*WMT-IR*	*WMT-DR*	*WMT-CR*
Standard cutoff	3	2	4	6	4	9	1	2	6
Adjusted 90%	1	2	1	4	4	5	0	1	2
Adjusted 98%	1	0	0	2	0	1	0	0	0
Profile analysis	—	—	—	0	0	0	0	1	1

WMT-IR = Word Memory Test Immediate Recognition Trial; WMT-DR = Word Memory Test Delayed Recognition Trial; WMT-CR = Word Memory Test Consistency Rating.

who fail the WMT easy components by comparing the mean of the easy effort components to the mean of the more difficult memory components. If the WMT is failed, but an easy–hard difference of >30 points is observed, then participants are judged as displaying a "genuine cognitive impairment profile" when they have a neurological condition known to be associated with significant cognitive problems.

Where full WMT data from both the easy and hard subtests were available (18 participants from the community group and 16 participants from the epilepsy group), profile analysis revealed that all but 1 of the participants who failed the Immediate/Delayed Recognition trials and/or Consistency Rating of the WMT also met the criteria for a genuine cognitive impairment profile. Only 1 participant from the epilepsy group continued to fail the WMT when applying profile analysis; however, this participant's effort test scores were generally variable, with the participant failing the Delayed Recognition Trial while passing the Immediate Recognition Trial, even at standard cutoffs.

DISCUSSION

Tests of effort are now commonplace within clinical neuropsychology and there are both American and UK guidelines to govern their use. If effort tests are to be increasingly used in clinical settings, it is imperative that data are available relating to how clinical cases presenting to rehabilitation services may perform on such tests, so clinicians do not risk making inaccurate judgments regarding the legitimacy of genuine deficits. However, limited data exist on the base rate of failure among clinical populations, or what may influence such failures. The participants in the current research were all diagnosed with an established neurological condition known to produce significant cognitive impairment. Although the exploratory nature of the study cautions against overinterpretation of the results, this research does suggest that clinicians need to be cautious when using established cutoff scores to interpret effort test data among brain-damaged groups without applying the more recently developed adjustments to such scores.

The rates of failure on the WMT identified in the current research when using standard cutoffs are broadly consistent with the levels of failure found in other clinical studies not applying reduced cutoff scores or profile analysis (e.g., Batt et al., 2008; Bunnage, Eichinger, Pearce, Duckworth, & Newson, 2008; Dodrill, 2008; Merten et al., 2007), whereas failure rates across most of the additional effort tests appear higher than those cited in other studies (e.g., Inman & Berry, 2002; Kelly et al., 2005).

A potential reason for the differential failure rates across the current subgroups is the varying level of cognitive and neurobehavioral impairments. The community participants were generally more severely impaired than were the other participants based on their clinical history and presentation, particularly when compared with those from the epilepsy group. In addition, because the current participants were clinical cases with objectively identified and confirmed brain injuries, some of the tests are likely to be measuring genuine impairment in these populations. As suggested by the statistical analyses, Mental Control and Digit–Symbol Coding could be measuring processing speed, while the Camden Memory Test for Faces and the Autobiographical Memory Index may be measuring memory. Failure on the Coin-in-Hand Test was uncommon in the current study, supporting the proposition that this is an extremely simple task (Kapur, 1994) with very high specificity rates but low sensitivity. The significant number of failures on the Autobiographical Memory Index within the community sample also suggests increased cognitive impairment in this group and that this measure is not suitable for use in this population. Although none of the participants scored at chance levels, the very high cutoff of <10.5 is likely to explain this finding and is probably due to the nature of the questions included in the test. For example, the question that asks participants to recite their telephone number is of reduced importance for this population following increased use of mobile phones that can be utilized as a memory aid and the fact that they are receiving more assistance from others to organize and manage their day-to-day functioning. However, following the recent emergence of WMT adjusted cutoff scores and profile analyses, we applied these procedures to the current data. Applying these techniques resulted in much lower failure rates on the WMT and provided evidence to support the use of such techniques in clinical populations.

Most participants in the current study passed the WMT when applying the 90% and 98% specificity-adjusted WMT scores proposed by Greve et al. (2008), but several participants did continue to fail. However, Greve et al.'s data were based upon a population that contained people with mild TBI, whereas no person in the current research would have been classified as having this level of injury and other diagnoses in addition to TBI were also included in the present study. The participants who continued to fail even at the 98% specificity cutoffs were those whose testing needed to be halted prematurely, again indicating that other factors were likely influencing performance.

When applying profile analyses (Green, 2009) to the WMT scores in the current research, a genuine cognitive impairment profile was observed in all but one of the participants who failed the WMT, indicating application of this method is important to reduce the risk for inaccurate conclusions being drawn regarding the performance of impaired populations. However, application of a profile analysis calculation may be meaningless without further consideration of clinical information. Indeed, Green et al. (2011) state that the profile alone cannot be relied upon to make judgments about poor effort; only someone with clear clinical signs of severe impairment or a diagnosis associated with such impairment would be considered as putting forth full effort after failing the WMT but displaying such a profile. Therefore, consideration of the clinical presentation, diagnosis, and application of neuropsychological knowledge and expertise in establishing evidence for impairment remains critical in accurately interpreting failures on the WMT where a genuine cognitive impairment profile is identified.

Green et al. (2011) have suggested that specificity would be higher for people who are less cognitively impaired than their probable dementia sample, which would include all those with mild TBI and most of those with moderate or severe TBI. Although the participants in the current study all have genuine diagnoses known to be associated with cognitive impairment, many of them would likely not be considered as severe as someone with dementia or as mild as someone with depression. However, the current study indicates that people with other neurological conditions in addition to dementia can display high rates of failure on the WMT that are potentially explainable through the application of profile analyses. Additional research and increased guidance is still required regarding the specific types of neurological conditions that may potentially lead to failure on the WMT at standard cutoffs and whether profile analysis can then identify these cases appropriately.

As previously suggested, the qualitative aspects of our participants' presentations appeared significant to test interpretation. Although all participants had brain lesions that would be expected to produce neuropsychological deficits, many of the community participants also had clinically evident neurobehavioral symptoms consistent with their more severe and diffuse/anterior brain injuries. Such participants were prone to impulsive responding and distractibility, and this could be a factor in the relatively high rates of effort test failure among this group when using standard cutoffs without adjustments. In comparison, the epilepsy group had focal rather than diffuse brain lesions and presented with fewer neurobehavioral difficulties. Given that the participants with epilepsy also reported higher anxiety levels than did the other groups, anxiety does not seem a likely factor for the WMT results in the present study.

The structure of the tests may be another reason for the differential failure rates. For example, Batt et al. (2008) suggested variance in failure rates could be due to the use of semantically related words in the WMT, which could create confusion and more errors due to "semantic interference" (Wehner, Ahlfors, & Mody, 2007) even when applying full effort. Budson et al. (2002) also found that unlike controls, people with frontal lesions could not reduce their rates of false recognition across successive trials. Further studies could look in more detail at the structure of the WMT, as it may be that people with frontal lesions may struggle on this test as a result of the semantic relatedness of test items. Batt et al. also suggest that if the resulting brain injury has affected the ability to monitor performance and learn from feedback, then this could also contribute to test results. Such suggestions are consistent with the notion that more anterior or diffuse brain damage could be contributing to scores on effort tests in the current study. Two patients were reported in the Henry et al. (2010) study as impulsive responders, and the researchers highlight that empirical evidence is still lacking to show whether such behaviors may lead to failure on effort tests. Future research could attempt to assess attention/executive functioning skills in more detail to establish the extent of their influence.

Although it was assumed that the clinical participants in the current study were all applying good effort given their voluntary participation, established neurological diagnosis, and lack of litigation, people in UK clinical rehabilitation settings are often receiving external benefits of some form, including disability payments, discounted taxation rates, or free prescription medication.

It is possible that some cases within the current research had idiosyncratic reasons for not putting forth full effort. For example, some may have illogically believed that they would lose benefits or access to health services if they did well on tests. However, conscious feigning or exaggerating seems unlikely. None of the participants were presenting to services for explicit assessments of whether they should receive benefits, and feigning symptoms is implausible when presenting to services for monitoring of a brain tumor or when residing in 24-hr community care or acute hospital wards. In addition, all but one person who failed the WMT displayed a genuine cognitive impairment profile. The large number of such profiles suggests more severe cognitive impairment, as we would not expect as many genuine impairment profiles if the scores were solely due to poor effort.

Rogers (2008b) has also noted that many issues can affect test performance in addition to monetary gain, including environmental and individual factors. It is generally accepted that underperformance may also be a nonconscious response to maintain autonomy, communicate distress, be taken seriously, or preserve a sense of self-worth. Therefore, consideration must be given to psychological factors when interpreting low scores on effort tests in clinical populations. Particularly, the inpatient participants were on an acute ward after experiencing a major disruption to their lives and might therefore be failing to appreciate the context of the assessment and not applying the full effort required. A lack of incentive to perform well may also have led to reduced motivation. Therefore, we cannot be sure that all of the effort test failures were for cognitive reasons because the participants' motivations are not well understood, and such analyses were beyond the scope of this study. However, all participants did undergo a lengthy consent process to ensure they understood the confidential and voluntary nature of participation, indicating that their primary motivation for agreeing to participate was to do something helpful.

Future research would benefit from more clearly controlling for the presence or absence of external incentives within the clinical populations assessed. For example, it would be interesting to include an additional group involved in litigation and matched for type and severity of brain injury. It would also be useful to assess whether participants with severe brain injuries similar to those seen in the current study can pass effort tests at higher rates when provided with a direct external incentive contingent upon doing well on the tasks, as this would have implications for assessments and rehabilitation interventions in our standard clinical practice.

Although the current research included a variety of measures, the limited neuropsychological test data are a drawback. Research including more comprehensive test data will be of particular importance when attempting to establish clear reasons for effort test failure. However, the present work presents data on recently and severely brain-damaged patients who were difficult to recruit and difficult to test. The authors believe that these base rate data provide a valuable contribution to the emerging literature on effort test interpretation.

A number of the effort tests used in the present study are also not considered to be among the most sensitive or specific (McCarter, Walton, Brooks, & Powell, 2009). It would be valuable to evaluate effort test performance in the current groups with another effort measure that has received equal research and publicity to the WMT, such as the TOMM (Tombaugh, 1996). In addition, Green (2003) recommends inclusion of WMT Consistency Ratings as a measure of effort. The Consistency Ratings were used in the current research to calculate effort test failure rates and to apply the easy–hard difference for profile analysis. However, the rating was not used in any statistical analyses due to the expected inherent variability of scores in people with severe cognitive problems and because the Consistency Rating is a composite score based on the directly measured Immediate and Delayed Recognition trials. There is also a lack of research specifically addressing the incremental validity of these ratings given that they are composite measures. Tables 3 and 4 do indicate that some participants only failed the Consistency component of the WMT within the current clinical populations, and inclusion of this score did increase the number of failures at standard cutoffs. It would be beneficial to include the Consistency scores in further research to identify whether it is appropriate to apply in populations with established neurological diagnoses known to be associated with significant cognitive impairment.

In conclusion, although the majority of participants passed effort tests, the current data suggest that people with severe and enduring brain injury, recent brain injury, and intractable epilepsy (as confirmed via brain scans and/or EEG) fail certain effort tests at higher rates than previously thought when using standard cutoff scores. Premorbid IQ, anxiety, and external incentives are unlikely explanations for the findings. Cognitive impairment is a potential explanation for the observed failure rates, and this was particularly obvious with the WMT when utilizing adjusted cutoffs developed with a genuine brain injury population (Greve et al., 2008) or when applying the profile algorithm suggested by Green et al. (2009, 2011). Only one person continued to fail the WMT when applying the profile algorithm, which supports high specificity when using profile analysis with people who have established neurological conditions other than dementia (Henry et al., 2010) and a low risk for concluding poor effort in participants similar to those assessed in the current research.

Further work on the sensitivity and specificity of effort tests in clinical populations is crucial in guiding clinicians in test interpretation and decision making. The recent trend of using lowered cutoffs (Axelrod & Schutte, 2011; Greve et al., 2008) or employing profile analyses (Green et al., 2011; Howe & Loring, 2009) appear promising, and more widespread clinical application of these techniques is recommended.

REFERENCES

Ardolf, B. R., Denney, R. L., & Houston, C. M. (2007). Base rates of negative response bias and malingered neurocognitive dysfunction among criminal defendants referred for neuropsychological evaluation. *The Clinical Neuropsychologist*, *21*, 899–916.

Armistead-Jehle, P., & Gervais, R. O. (2011). Sensitivity of the test of memory malingering and the nonverbal medical symptom validity test: A replication study. *Applied Neuropsychology*, *18*, 284–290.

Axelrod, B. N., & Schutte, C. E. (2010). Analysis of the Dementia Profile on the Medical Symptom Validity Test. *The Clinical Neuropsychologist*, *24*, 873–881.

Axelrod, B. N., & Schutte, C. E. (2011). Concurrent validity of three forced-choice measures of symptom validity. *Applied Neuropsychology*, *18*, 27–33.

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*, 1074–1080.

Belanger, H. G., Curtiss, G., Demery, J. A., Lebowitz, B. K., & Vanderploeg, R. D. (2005). Factors moderating neuropsychological outcomes following mild traumatic brain injury: A meta-analysis. *Journal of the International Neuropsychological Society*, *11*, 215–217.

Binder, L. M. (1993). *Portland Digit Recognition Test manual* (2nd ed.). Portland, OR: Author.

Blaskewitz, N., Merten, T., & Kathmann, N. (2008). Performance of children on symptom validity tests: TOMM, MSVT, and FIT. *Archives of Clinical Neuropsychology*, *23*, 379–391.

Bowden, S. C., Shores, A. E., & Mathias, J. L. (2006). Does effort suppress cognition after traumatic brain injury? A re-examination of the evidence for the Word Memory Test. *The Clinical Neuropsychologist*, *20*, 858–872.

Budson, A. E., Sullivan, A. L., Mayer, E., Daffner, K. R., Black, P. M., & Schacter, D. L. (2002). Suppression of false recognition in Alzheimer's disease and in patients with frontal lobe lesions. *Brain*, *125*, 2750–2765.

Bunnage, M., Eichinger, C., Pearce, N., Duckworth, A., & Newson, M. (2008). Criterion validity of the Word Memory Test: An audit of a sample of patients assessed for clinical, not litigious, reasons. [Abstract]. *Journal of the International Neuropsychological Society*, *14*(Suppl. 1), 138–139.

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

Cochrane, H. J., Baker, G. A., & Meudell, P. R. (1998). Simulating a memory impairment: Can amnesics implicitly outperform simulators? *British Journal of Clinical Psychology*, *37*, 31–48.

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*, 191–198.

Coughlan, A. K., Oddy, M. J., & Crawford, J. R. (2007). *BIRT Memory and Information Processing Battery (BMIPB)*. London, England: Brain Injury Rehabilitation Trust.

Delis, D. C., Kramger, J. H., Kaplan, E., & Ober, B. A. (2000). *California Verbal Learning Test* (2nd ed.). San Antonio, TX: Psychological Corporation.

Dodrill, C. B. (2008). Do patients with psychogenic nonepileptic seizures produce trustworthy findings on neuropsychological tests? *Epilepsia*, *49*, 369–370.

Flaro, L., Green, P., & Robertson, E. (2007). Word Memory Test failure 23 times higher in mild brain injury than in parents seeking custody: The power of external incentives. *Brain Injury*, *21*, 373–383.

Fox, D. D. (2011). Symptom validity test failure indicates invalidity of neuropsychological tests. *The Clinical Neuropsychologist*, *25*, 488–495.

Green, P. (2003). *Word Memory Test for Windows: User's manual and program*. Edmonton, AB, Canada: Green's Publishing.

Green, P. (2004). *Medical Symptom Validity Test for Windows: User's manual and program*. Edmonton, AB, Canada: Green's Publishing.

Green, P. (2008). *Green's Nonverbal Medical Symptom Validity Test (NV-MSVT) for Microsoft Windows: User's manual 1.0*. Edmonton, AB, Canada: Green's Publishing.

Green, P. (2009). *The advanced interpretation program for the WMT, MSVT, NV-MSVT and MCI*. Edmonton, AB, Canada: Green's Publishing.

Green, P. (2011). Comparison between the Test of Memory Malingering (TOMM) and the Nonverbal Medical Symptom Validity Test (NV-MSVT) in adults with disability claims. *Applied Neuropsychology*, *18*, 18–26.

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*, 741–750.

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

Green, P., Montijo, J., & Brockhaus, R. (2011). High specificity of the Word Memory Test and Medical Symptom Validity Test in groups with severe verbal memory impairment. *Applied Neuropsychology*, *18*, 86–94.

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

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*, 896–918.

Haines, M. E., & Norris, M. P. (1995). Detecting the malingering of cognitive deficits: An update. *Neuropsychology Review*, *5*, 125–148.

Henry, M., Merten, T., Wolf, S. A., & Harth, S. (2010). Nonverbal Medical Symptom Validity Test performance of elderly healthy adults and clinical neurology patients. *Journal of Clinical and Experimental Neuropsychology*, *32*, 19–27.

Howe, L., & Loring, D. (2009). Classification accuracy and predictive ability of the Medical Symptom Validity Test's Dementia Profile and Genuine Memory Impairment Profile. *The Clinical Neuropsychologist*, *23*, 329–342.

Inman, T. H., & Berry, D. T. R. (2002). Cross-validation of indicators of malingering: A comparison of nine neuropsychological tests, four tests of malingering, and behavioral observations. *Archives of Clinical Neuropsychology*, *17*, 1–23.

Kapur, N. (1994). The Coin-in-the-Hand Test: A new 'bedside' test for the detection of malingering in patients with suspected memory disorder. *Journal of Neurology, Neurosurgery and Psychiatry*, *57*, 385–386.

Kelly, P. J., Baker, G. A., van den Broek, M. D., Jackson, H., & Humphries, G. (2005). The detection of malingering in memory performance: The sensitivity and specificity of four measures in a UK population. *British Journal of Clinical Psychology*, *44*, 333–341.

Kemp, S., Coughlan, A. K., Rowbottom, C., Wilkinson, K., Teggart, V., & Baker, G. (2008). The base rate of effort test failure in patients with medically unexplained symptoms. *Journal of Psychosomatic Research*, *65*, 319–325.

Larrabee, G. J. (2003). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist*, *17*, 410–425.

McCarter, R. J., Walton, N. H., Brooks, D. N., & Powell, G. E. (2009). Effort testing in contemporary UK neuropsychological practice. *The Clinical Neuropsychologist*, *23*, 1050–1066.

McMillan, T. M., Anderson, S., Baker, G., Berger, M., Powell, G. E., & Knight, R. (2009). *Assessment of effort in clinical testing of cognitive functioning for adults*. Leicester: British Psychological Society.

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*, 308–318.

Meyers, J. E., Volbrecht, M., Axelrod, B. N., & Reinsch-Boothby, L. (2011). Embedded symptom validity tests and overall neuropsychological test performance. *Archives of Clinical Neuropsychology*, *26*, 8–15.

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

Moss, A., Jones, C., Fokias, D., & Quinn, D. (2003). The mediating effects of effort upon the relationship between head injury severity and cognitive functioning. *Brain Injury*, *17*, 377–387.

Rogers, R. (2008a). *Clinical assessment of malingering and deception* (3rd ed.). New York, NY: Guilford.

Rogers, R. (2008b). An introduction to response styles. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed., pp. 3–13). New York, NY: Guilford.

Rohling, M. L., Binder, L. M., Demakis, G. J., Larrabee, G. J., Ploetz, D. M., & Langhinrichsen-Rohling, J. (2011). A meta-analysis of neuropsychological outcome after mild traumatic brain injury: Re-analyses and reconsiderations of Binder et al. (1997), Frencham et al. (2005), and Pertab et al. (2009). *The Clinical Neuropsychologist*, *25*, 608–623.

Rohling, M. L., & Demakis, G. J. (2010). Bowden, Shores, & Mathias (2006): Failure to replicate or just failure to notice. Does effort still account for more variance in neuropsychological test scores than TBI severity? *The Clinical Neuropsychologist*, *24*, 119–136.

Singhal, A., Green, P., Ashaye, K., Shankar, K., & Gill, D. (2009). High specificity of the Medical Symptom Validity Test in patients with very severe memory impairment. *Archives of Clinical Neuropsychology*, *24*, 721–728.

Stevens, A., Friedel, E., Mehren, G., & Merten, T. (2008). Malingering and uncooperativeness in psychiatric and psychological assessment: Prevalence and effects in a German sample of claimants. *Psychiatry Research*, *157*, 191–200.

Tombaugh, T. N. (1996). *TOMM, Test of Memory Malingering*. New York, NY: Multi Health Systems.

Trueblood, W. (1994). Qualitative and quantitative characteristics of malingered and other invalid WAIS-R and clinical memory data. *Journal of Clinical and Experimental Neuropsychology*, *16*, 597–607.

Warrington, E. K. (1996). *The Camden Memory Test Battery*. Hove East Sussex, UK: Psychology Press.

Wechsler, D. (1997a). *Wechsler Adult Intelligence Scale* (3rd ed.). San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997b). *Wechsler Memory Scale* (3rd ed.). San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2001). *Wechsler Test of Adult Reading*. San Antonio, TX: The Psychological Corporation.

Wehner, D. T., Ahlfors, S. P., & Mody, M. (2007). The influence of semantic processing on phonological decisions in children and adults: A magnetoencephalography (MEG) study. *Journal of Speech, Language, and Hearing Research*, *50*, 716–731.

West, L. K., Curtis, K. L., Greve, K. W., & Bianchini, K. J. (2011). Memory in traumatic brain injury: The effects of injury severity and effort on the Wechsler Memory Scale-III. *Journal of Neuropsychology*, *5*, 114–125.

Wiggins, E., & Brandt, J. (1988). The detection of simulated amnesia. *Law and Human Behaviour*, *12*, 57–78.

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, *67*, 361–370.