

Assessment of Feigned Cognitive Impairment: A Cautious Approach to the Use of the Test of Memory Malingering for Individuals with Intellectual Disability

Author: Christopher L. Ray, Ph.D., ABPP, Department of Mental Health & Behavioral Sciences, Columbus VA Chalmers P. Wylie Ambulatory Care Center, Department of Veterans Affairs, Columbus, Ohio, United States of America. Email: christopher.ray5@va.gov

Abstract

The detection of feigned cognitive impairment is an important issue to consider in forensic assessment. The Test of Memory Malingering (TOMM) (Tombaugh, 1996) is a frequently administered measure to assess feigned cognitive impairment. There is disagreement in the research literature on the TOMM regarding its appropriateness in cases when the defendant has intellectual disability (formally known as mental retardation). The importance of considering all available data in interpreting the TOMM is recommended to minimize the chance of defendants with bona fide intellectual disability being misclassified as feigning cognitive impairment. Recommendations regarding alternative effort tests for individuals with intellectual disability are provided.

Keywords: feigned cognitive impairment, Test of Memory Malingering, intellectual disability, psychological tests, forensic assessment

The Importance of Malingering Assessment

There is a growing consensus that the assessment of malingering is a fundamental component of forensic evaluations that should be considered and systematically assessed with each referral (Heilbrun, 2001; Rogers, 1997; Vitacco, Rogers, Gabel & Munizza, 2007). Malingering (American Psychiatric Association, 2000) is the deliberate fabrication or gross exaggeration of psychological or physical symptoms for the fulfillment of an external goal. Mittenberg, Patton, Canyock, and Condit (2002) surveyed 131 neuropsychologists on the prevalence of malingering encountered in their practice. They reported probable malingering prevalence rates of 19% of criminal, 29% of personal injury, 30% of disability, and 8% of medical cases. Utilizing Slick, Sherman, and Iverson's (1999) criteria, Ardolf, Denney, and Houston (2007) demonstrated that the combined rate of probable and definite malingered neurocognitive dysfunction in the criminal setting for pre-sentence males referred for neuropsychological assessment as a part of a court-ordered evaluation was over 54%. Moreover, Greve, Ord, Bianchini, and Curtis (2009) found that the prevalence of malingering in 508 patients with chronic pain who had a financial incentive was between 20% and 50%, depending on the diagnostic system used and the statistical model's underlying assumptions. Some

Ray, C. L. (2012). Assessment of feigned cognitive impairment: A cautious approach to the use of the Test of Memory Malingering for individuals with intellectual disability. *Open Access Journal of Forensic Psychology*, 4, 24-50.

factors associated with the medico-legal context such as the jurisdiction of a worker's compensation claim or attorney representation were associated with slightly higher malingering rates.

Feigning is a term described by Rogers, Sewell, Martin, and Vitacco (2003) as the deliberate fabrication or gross exaggeration of psychological or physical symptoms without any assumptions about its goals. According to Rogers and Bender (2003), available tests typically assess feigning, because they are unable to evaluate supposed goals required for the classification of malingering or the diagnosis of factitious disorders.

Defendants can attempt to deceive forensic evaluators by fabricating or exaggerating mental disorder symptoms or by convincing examiners that their efforts to succeed are sincere and that their ostensible impairments are genuine. According to Heilbronner et al. (2009), various terms such as 'insufficient effort,' 'inadequate effort,' and 'poor effort' have been used by researchers and clinicians to describe the behaviors of interest in the identification of intentionally exaggerated symptoms and diminished or reduced capability. Heilbronner et al. suggest that a useful descriptor of problematic effort is one that clearly conveys a substantial negative impact that has the potential of invalidating measurement of ability. Heilbronner et al. further indicate that measures used to identify problematic effort are often identified as effort tests. Such tests are considered to be in a category of measures that evaluate the validity of symptoms and are known as symptom validity tests (SVTs).

Pankratz and Binder (1997) suggest that feigned cognitive impairment, which is a term often used to describe effortful failure, is observed in a wide range of settings with a spectrum of presentations. Currently, the size of the neuropsychological literature related to assessment of feigned cognitive impairment in the measurement of abilities and the overall quality of this literature is substantial and well developed (Boone, 2007; Heilbronner et al., 2009; Larrabee, 2007; Morgan & Sweet, 2009; Rohling et al., 2011). According to Heilbronner et al., there is now abundant research evidence that stand-alone cognitive effort tests are extremely useful within forensic evaluations, which have been shown to be associated with a high risk of invalid responding. Rohling et al. agree that cognitive effort tests can be used to evaluate many important questions, including prospects for malingering in defendants involved in forensic assessments such as insanity claims.

The Value of the Routine Use of Effort Tests

Rogers and Bender (2003) advocate a specialized perspective in the detection of malingering, a perspective that focuses on the use of measures specifically designed for the assessment of response styles. In situations when the primary issue is mental illness, investigators such as Vitacco et al. (2007) suggest using malingering screeners on a routine basis to bring to the attention of the examiner individuals who may be fabricating or exaggerating their clinical condition. Malingering screeners provide

clinicians with a relatively brief, cost-efficient means of identifying individuals who might be malingering or engaging in other forms of dissimulation (Ray, 2009). The regular administration of cognitive effort tests would appear to be similarly beneficial in the identification of individuals showing signs of cognitive impairment. Heilbronner et al. (2009) contend, in fact, that *not* administering such measures in the context of a forensic evaluation requires strong justification by the examiner.

According to Iverson and Binder (2000), SVTs are needed to measure effort because the amount of effort a person makes has been shown to affect neuropsychological test scores to a major degree. There are clear recommendations now that a thorough assessment of symptom validity should be routinely included in all forensic and clinical neuropsychological evaluations (Boone, 2007; Bush et al., 2005; Iverson, 2006). According to Green (2007), effort tests are used mainly because they help determine whether effort is sufficient to produce valid data and, hence, whether other test data are likely to be valid. Because of the substantial impact pretrial forensic evaluations have on defendants, then, cognitive effort tests are worthwhile instruments to consider administering to identify individuals who are feigning or exaggerating cognitive deficits. Such an approach is consistent with Vitacco et al.'s (2007) recommendation to systematically assess individuals who might be feigning mental illness.

Brief Review of Cognitive Effort Tests

There are a variety of cognitive effort tests to choose from, and, as noted by Bush et al. (2005), the extra time and cost associated with administering such instruments is justified by the benefits associated with the detection of potential malingerers. For individuals with intellectual disability (formally referred to as mental retardation), however, the administration of cognitive effort tests requires careful consideration. A complete review of the multitude of SVTs is beyond the scope of this article. What follows is a brief examination of some widely used effort measures.

The Validity Indicator Profile (VIP) (Frederick, 1997) employs a two-choice format for the assessment of suboptimal effort. According to Rogers and Bender (2003), the VIP is distinguished from other cognitive measures by its use of multiple strategies focused predominantly on unexpected patterns. The strategies include three estimates of response consistency and five estimates of performance curve. The VIP classifies profiles as either "valid" or "invalid." Frederick suggests that invalid profiles are then sorted into three categories including inconsistent (minimal or inconsistent effort but motivated to perform well), irrelevant (intention to perform poorly and likely responding without regard to item content) and suppressed (high effort to perform poorly with an attempt to feign cognitive deficits). Rogers and Bender reported that using the broad categories of valid and invalid, classification rates are moderately high. The VIP nonverbal subtest (VIP-NV), for example, has a sensitivity rate of 73.5% and a specificity rate of 85.7% while the VIP verbal test (VIP-V) has a sensitivity rate of 67.3% and a specificity rate of 83.1% (Frederick, 2003).

The VIP manual (Frederick, 1997) reported a sample of 40 individuals with historically demonstrable mental retardation who were in assisted living in the community. On both VIP subtests, Frederick indicated that the mean total score for those individuals with a bona fide history of mental retardation was higher than that for computer-generated random protocols but lower than that for all other participant groups, including the “noncompliant” coached normal and patients at risk for malingering. According to Frederick, only 20% received a compliant classification on the VIP-NV and 25% on the VIP-V. Frederick further suggested that an irrelevant classification was received by about a third of this sample on the VIP-NV and by half of the group on the VIP-V. Frederick concluded that the VIP is not useful in evaluations of individuals with historically demonstrable mental retardation or continuing severe cognitive impairment.

A peer reviewer noted that, even if an individual’s ability is quite low, the performance curve on the VIP would be shifted, but it should still have the same shape and characteristics with easier items passed and a gradual fall off as items become harder. The reviewer indicated that it is not apparent why this would differ for mentally retarded persons. According to Frederick (1997), the relationship between IQ and VIP response-style classification in low IQ individuals suggests that those with extremely low ability may stop trying after they encounter more difficult items mixed among easy ones. Frederick added that the organization of items on the VIP test is randomized with respect to difficulty, and some individuals may assume, after trying the first few, that they will not know the answers to any items. Accordingly, Frederick suggested they may simply fill in the answer blanks randomly after a few items.

Although the VIP does not appear useful in evaluating individuals with historically demonstrable mental retardation, Frederick (2003) indicated that he has commonly been asked whether the VIP can be used to evaluate individuals *suspected* of mental retardation. Frederick suggested that the VIP is useful in such cases. According to Frederick, individuals who feign severe mental impairment generally produce suppressed or irrelevant performance curves. R. Frederick (personal communication, June 4, 2012) noted that the VIP is intended for use when there is a legitimate question about whether somebody is impaired, especially when they are presenting with intellectual ability at the level of mental retardation.

The Rey 15-Item Test (FIT) was developed by Andre Rey (1964). This test has been described in prominent neuropsychology texts (e.g., Lezak, 1995) and, according to Nitch and Glassmire (2007), survey data show that it is one of the most commonly administered effort measures. In Rabin, Barr, and Burton’s (2005) survey of practicing neuropsychologists, the FIT was the only effort measure other than the TOMM in the list of the 40 most commonly administered instruments in memory assessment. Nitch and Glassmire indicate that the FIT was developed as a technique for assessing the validity of visual memory complaints and consists of 15 simple, highly redundant items presented in five rows of three characters each. The examinee is shown the stimulus sheet for 10 seconds, after which the sheet is removed and the examinee is asked to

reproduce as many of the 15 items as he or she can remember on a blank piece of paper.

Nitch and Glassmire (2007) suggested that the FIT uses the floor-effect principle as it is thought to be so simple that all but severely brain damaged or intellectually disabled individuals could perform the task easily. Lezak (1995) indicated that individuals who are not significantly deteriorated can recall at least three of the five character sets, or nine items. According to Nitch and Glassmire, sensitivity levels for a cutoff of fewer than nine items average slightly less than 50%, while specificity values are typically above 90%. However, an examinee's FIT performance may be adversely affected by the presence of borderline or lower IQ (Goldberg & Miller, 1986). Hayes, Hale, and, Gouvier (1997) investigated the utility of the FIT along with two other effort indicators for predicting the malingering status of 37 intellectually disabled defendants. The results showed that the malingerers produced better memory scores than the groups with true intellectual disability. The authors concluded that the FIT should not be used among defendants with intellectual disability.

The Dot Counting Test (DCT) was originally developed by Andre Rey (1941) as a technique to detect malingered cognitive symptoms. It consists of 12 cards, each containing different numbers of dots. The first six cards contain ungrouped dots, and the final six contain dots grouped in various ways. Participants are asked to count the number of dots on each card as quickly and accurately as possible, with the examiner recording response time and number of dots counted for each card. Nitch and Glassmire (2007) indicated that development of a combination score incorporating both response times and errors showed adequate sensitivity in detecting suspect effort (averaging 78.8%), with detection rates likely higher in a correctional setting. However, specificity rates vary considerably across clinical comparison groups. Nitch and Glassmire noted that in its current form, the DCT appears contraindicated for use in the determination of actual versus feigned mental retardation and moderate dementia presentations.

Hurley and Deal (2006) administered the DCT, along with three other instruments assessing malingering, to 39 individuals with intellectual disability. Only one participant exceeded the cut off score of 180 seconds, suggesting that this test may hold promise as a brief screening device for individuals of below average intelligence. Further study of the DCT using alternative strategies for scoring the data besides total time was recommended.

In an unpublished study, Proto (2008) found in a simulation design that malingerers performed significantly worse than individuals with intellectual disability on the DCT. The DCT total-time variable provided an excellent level of sensitivity (100%) in a group of individuals with mild mental retardation. However, specificity was low. The results suggest that total number of errors on the DCT might prove to be a useful indicator of suspect effort in individuals thought to be malingering intellectual disability.

The Letter Memory Test (LMT) (Inman et al., 1998) is a forced-choice test of motivation. This test uses letters as stimuli in 45 trials that cross two factors, which include the number of stimuli to be remembered (3, 4, or 5) and the number of foils (1, 2, or 3) presented with the target stimulus. According to Graue et al. (2007), reported sensitivity rates at the recommended cutting score of <93% correct have ranged from .58 to .90 and specificity rates from .88 to 1.00. Graue et al. noted, however, that the LMT had not been validated with intellectually disabled samples.

Graue et al. (2007) compared results from tests of intelligence, psychiatric feigning, and neurocognitive faking in a group of 26 mildly mentally retarded defendants and 25 demographically matched community volunteers asked to feign mental retardation. The LMT significantly discriminated a group of community volunteers who were malingering from both honest community volunteers and an intellectually disabled group with robust *d* scores for the crucial contrasts. Application of the recommended cutting score for the LMT produced a moderate hit rate of .73. Although the sensitivity rate of .88 was comparable to somewhat higher than those reported in other studies, the specificity rate was poor at 58%. A revised cutting score of <70% on the LMT resulted in sensitivity of .76, specificity of .96, and a hit rate of .88. Of note, however, is that in a study by Shandera et al. (2010), the modified (lower) cutting score for the LMT failed to cross-validate successfully.

The Word Memory Test (WMT) (Green, Allen & Astner, 1996; Green & Astner, 1995) is a verbal memory test with multiple subtests and with self-contained effort measures, which serve to check whether the test taker's test scores are valid or not. According to the WMT manual (Green, 2003), the primary WMT effort or symptom validity subtests are very sensitive to poor effort or to exaggeration of cognitive difficulties. However, they are insensitive to all but the most extreme forms of cognitive impairment.

The WMT contains a number of normative comparison groups including, but not limited to, normal adult control volunteers; parents seeking custody of their children; neurological patients with brain tumors, strokes, ruptured aneurysms, multiple sclerosis, and other diseases; children with fetal alcohol syndrome; adults with major depression who passed the WMT effort subtests and were on disability or making a claim for disability; adults with major depression but who failed the WMT effort tests; adults with mild head injuries; and children with neurological diseases. According to Green (2003), the WMT is failed more often by people with mild than with severe brain injury.

Green (2007) indicated that it is important to determine how well one effort test compares with another in predicting whether or not other test data are valid. Some studies have been performed in which several effort tests have been used with the same subjects to ensure that sample characteristics cannot explain differences between results on one SVT versus another. Tan, Slick, Strauss, and Hultsch (2002), for example, did a study comparing three well-known effort tests including the WMT, TOMM (Tombaugh, 1996), and the Victoria Symptom Validity Test (Slick, Hopp, Strauss, & Spellacy, 1996). The WMT was observed to be the most accurate and

achieved 100% effectiveness in discriminating between simulators and good-effort volunteers.

The WMT includes validation data in its manual for an adult population with intellectual disability. Adults with Verbal IQs in the intellectually disabled range scored a mean of 96% (SD = 5) on the WMT, which is well above the standard cutoff of 85% (Green, 2003). However, Victor and Boone (2007) noted that specificity rates were not reported and the authors (Green et al., 1996) indicated that future studies should determine WMT scores for motivated individuals of given VIQ ranges and some adjustment to cut-off scores may be needed for people of very low VIQ. Victor and Boone (2007) indicated that the WMT may be helpful in assessing effort in individuals with intellectual disability, but additional data is needed.

Brockhaus and Merten (2004) conducted two studies to determine the effectiveness of a German version of the WMT in assessing suboptimal cognitive performance. The first study compared the performance of a healthy experimental group of 100 malingerers with 27 healthy controls after standard test instructions. The WMT achieved a 100% correct identification of group membership. In the second study, the WMT was likewise able to identify high performance motivation in almost all cases among a sample of 32 intellectually disabled patients.

The findings of Brockhaus and Merten (2004) appeared to indicate good performance by intellectually disabled individuals on the WMT. Marshall and Happe (2007), however, reported a personal communication from Brockhaus, which noted some important limitations of Brockhaus and Merten's study. According to Marshall and Happe, participants selected for Brockhaus and Merten's study had to be able to understand and do the Spatial Span and Picture Completion tests. Moreover, patients were pre-selected only if they were thought to be emotionally stable enough to comply with and do the testing. Marshall and Happe suggested that only 18 of 75 patients with intellectual disability in a psychiatric hospital were preselected for the study. The other 12 patients in the study were reportedly pre-selected by the program directors at a workshop for individuals with intellectual disability. Marshall and Happe concluded that Brockhaus and Merten's study sample is not a representative sample of individuals with intellectual disability, which limits conclusions that can be drawn regarding the percentage of actual patients who could pass the WMT.

Recently, Green, Flaro, Brockhaus, and Montijo (2012) examined performance on the WMT in children with intellectual disability. These authors indicated that only 20 cases, or 5.3% of a sample of 380 intellectually disabled children, were classified as showing poor effort on the WMT. A specificity of 94.7% in this sample of intellectually disabled children was obtained. Also, of the 38 intellectually disabled children who were non-readers, only 5 cases showed a possible poor effort profile (i.e., 13% of the non-readers).

A peer reviewer suggested that even though Green et al. (2012) described their sample as “developmentally disabled,” there appeared to be few mentally retarded individuals included and a variety of individuals with other conditions such as attention-deficit/hyperactivity disorder (ADHD). An examination of Green et al.’s data confirms that of the 380 developmentally disabled individuals examined, only 21 were actually diagnosed with mental retardation. Moreover, there were individuals with a wide variety of clinical conditions such as ADHD, Asperger’s disorder, language impairment, oppositional defiant disorder, epilepsy, conduct disorder, learning disability, schizophrenia, personality disorder, obsessive compulsive disorder, traumatic brain injury, depression, developmental coordination disorder, Tourette’s syndrome, reading disorder, autism, reactive attachment disorder, nonverbal learning disability, dyslexia, fledgling psychopath, disruptive behavior disorder, posttraumatic stress disorder, hearing impairment, bipolar disorder, fetal alcohol spectrum disorder, and even one intellectually gifted individual. As such, the findings reported by Green et al. are derived from a relatively small sample of individuals who can truly be considered intellectually disabled.

The MSVT (Medical Symptom Validity Test) (Green, 2004) and the NV-MSVT (nonverbal MSVT) (Green, 2008) contain subtests that have been shown to be very easy for children with severe brain injury (Carone, 2008) and for healthy children (Blaskewitz, Merten, & Kathmann, 2008). Moreover, Singhai, Green, Ashaye, Shankar, and Gill (2009) compared 10 elderly dementia patients to 10 adult volunteers who were asked to try to simulate dementia. The authors found that combined profile analysis yielded 80% sensitivity to suboptimal effort in simulators and 100% specificity for genuine impairment in dementia, which exceeded the performance of either test alone. The results of this study replicated prior findings in dementia patients reported by Green (2008) and in a study by Henry, Merten, Wolf, and Harth (2009), which showed 100% specificity for the NV-MSVT in dementia.

Limitations of Singhai, Green, Ashaye, Shankar, and Gill’s (2009) study include its small sample size and the fact that it did not assess sensitivity to suboptimal effort in clinical cases with actual incentives to exaggerate memory impairment. This research is notable, however, because of the possible applicability of these tests to intellectually disabled populations. According to Singhai et al., further research with intellectually disabled children would be important to determine how many fail the easy subtests of the MSVT and the NV-MSVT and with what profiles.

In a comparison of 18 university students to 33 children tested clinically who had issues including ADHD and conduct disorder, fetal alcohol syndrome, and mental retardation, Green (2004) indicated that there were no differences in MSVT effort scores.

In addition to studying the WMT, Green et al. (2012) examined performance on the MSVT and NV-MSVT in children with intellectual disability. MSVT mean scores from 265 intellectually disabled children indicated 2.6% with a poor effort profile. According to the authors, a specificity of 97.45% was obtained. In an examination of NV-MSVT

mean scores from 217 intellectually disabled children, 4.1% had a poor effort profile. The authors added that the specificity of the NV-MSVT in intellectually disabled children making a good effort would be a minimum of 91.2%. Using profile analysis, Green et al. suggested that the maximum possible false-positive rate in intellectually disabled children was, on average, only 4% for the three tests (i.e., the WMT, MSVT & NV-MSVT), which is equal to a 96% rate of specificity.

The individuals included in the MSVT and NV-MSVT studies came from the same sample used in Green et al.'s (2012) WMT study. As noted, this sample of individuals had a wide variety of clinical diagnoses with relatively few individuals with actual mental retardation. Green et al.'s small sample size limits the generalizability of their findings to a broader population of individuals with intellectual disability.

Test of Memory Malingering

Another commonly administered cognitive effort test that has been described as a promising and useful instrument in examining feigned cognitive impairment (e.g., Delain, Stafford, & Ben-Porath, 2003; Heinze & Purisch, 2001) is the Test of Memory Malingering (TOMM) (Tombaugh, 1996). The TOMM is a 50-item recognition test for adults, which includes two learning trials and a Retention Trial. Two decision rules are used in the interpretation of the TOMM. The first rule suggests that scoring lower than chance on any trial indicates the possibility of malingering. The second rule is that any score lower than 45 on Trial 2 or the Retention Trial indicates the possibility of malingering. According to Tombaugh, performance on Trial 2 is very high for non-malingers regardless of age, neurological dysfunction, or psychological symptom.

Within the forced-choice testing paradigm, the material on the TOMM is designed to be so easy that well-motivated individuals with brain damage can perform in the normal range (D'Amato & Denney, 2008). According to Tombaugh (1997), the TOMM has been designed to be sensitive to malingering, but insensitive to neurological impairments. The TOMM, in other words, is set up to detect individuals who deliberately choose wrong answers in order to appear more cognitively impaired than is actually the case. Its simple design, ease of use, and relatively short administration time make it an ideal candidate to utilize as a measure of feigned cognitive impairment.

Usefulness of the TOMM with Special Populations

There are a number of studies in the research literature on the TOMM that suggest it is a useful instrument for various special populations. For example, Rees, Tombaugh, and Boulay (2001) examined how inpatients diagnosed with major depression performed on the TOMM. The results show that the TOMM is unaffected by affective state. All of the depressed individuals scored above cutoff levels on Trial 2 and the Retention Trial of the TOMM, achieving a 99% accuracy rate. Yanez, Fremouw, Tennant, Strunk, and Coker (2005) examined whether 20 participants with high levels of depression, as measured by the Beck Depression Inventory 2nd Edition (BDI-II) and with

current diagnoses of Major Depressive Disorder, would perform significantly worse on the TOMM than a control group. The results showed that the depressed and control groups did not have significant mean group differences on TOMM performance. Of the 20 depressed participants, only 2 on Trial 2 and 1 on the Retention Trial scored below the cutoff of 45, while none of the control participants performed in this range. The results indicate that the TOMM can be used with even severely depressed participants with only slight caution.

Ashendorf, Constantinou, and McCaffrey (2004) examined a sample of community-based older adults (55–75) to determine whether scores on the TOMM are influenced by the presence of symptoms of depression or anxiety, as measured by the Beck Depression Inventory (BDI) and State-Trait Anxiety Inventory (STAI), respectively. The results indicate that, regardless of BDI or STAI scores, all subjects scored above 45 correct out of 50 on TOMM Trial 2. These findings demonstrate that depression and anxiety levels in an older community-dwelling sample do not negatively affect performance on the TOMM.

In addition to being a useful instrument for detecting cognitive feigning for inpatients with depression and older adults with depression or anxiety, the research literature suggests that the TOMM is effective for individuals with cognitive disorder associated with psychosis (Duncan, 2005). Duncan analyzed the scores of inpatient psychiatric patients with varying degrees of cognitive impairment who met criteria for a psychotic disorder on a number of tests including the TOMM. Psychotic disordered patients with and without significant concentration deficits were compared. Although psychotic patients with concentration problems showed greater variability across TOMM trials, both groups obtained mean scores above 45 on Trial 2 and the Retention Trial on the TOMM. These findings indicate that cognitive impairment associated with psychosis generally does not negatively impair TOMM performance to such a level that would produce a false positive on the TOMM.

Performance on the TOMM was evaluated in a sample of 100 consecutively referred 6- to 16-year-old children with a wide range of clinical diagnoses (Donders, 2005). In the complete sample, 97 children met actuarially defined criteria for sufficient effort on the TOMM. Two children were correctly identified as providing suboptimal effort and only one case was a possible false positive. Performance on the second trial of the TOMM did not vary with gender, ethnicity, parental occupation, performance on an independent memory test, or length of coma. Although younger children tended to be somewhat less efficient on the TOMM than older children, more than 90% of children in the 6-8 years range met criteria originally developed for adults for sufficient effort on the TOMM. Donders concluded that the TOMM is a potentially useful measure of effort in the clinical neuropsychological evaluation of school-age children.

Teichner and Wagner (2004) examined 78 individuals referred for evaluation of memory complaints. The TOMM's performance was examined in a sample of cognitively intact elderly, dementia patients, and older adults with cognitive impairment who did not meet

diagnostic criteria for dementia. Normal and cognitively impaired individuals performed similarly on the TOMM independent of their cognitive status. The results were analogous to those documented in the TOMM test manual (Tombaugh, 1996). Of note is that the dementia group performed poorly. The results suggest that the TOMM is a useful index for detecting the malingering of memory deficits, even in patients with cognitive impairment, but only when dementia can be ruled out.

Applicability of the TOMM to Individuals with Dementia

Dean, Victor, Boone, Philpott, and Hess (2008) examined archival data from 214 non-litigating patients with dementia on 18 effort indices including the TOMM. The findings were similar to those found in prior research on effort testing in dementia. Previous specificities for demented groups on the second trial of the TOMM ranged from a high of 82% (Greve et al., 2009) to a low of 24% (Teichner & Wagner, 2004). Dean et al.'s specificity for the TOMM was 45%. Of note is that Dean et al. found that the TOMM retained greater than or equal to 90% specificity with a revised cut off of Trial 2 < 28. This finding suggested that, even in demented groups, incorrectly identifying half or less of the items is quite uncommon. Dean et al. concluded that only a minority of existing effort indicators were useful in differentiating between actual and feigned dementia when using standard cut off scores. However, Dean et al. also indicated that cut-off adjustments can be used as guidelines in interpreting effort test scores obtained in clinical practice.

Overall, the data presented by Teichner and Wagner (2004) and Dean et al. (2008) indicate that the TOMM is not a useful measure to assess test motivation for persons with dementia. Moreover, Green (2011) noted that, out of the 37 testable patients with dementia described in the test manual for the TOMM (Tombaugh, 1996), 27% failed the TOMM using the usual criterion, yielding only 73% specificity. As is the case with dementia, questions have been raised in the research literature regarding how well the TOMM works for individuals with intellectual disability.

Applicability of the TOMM to Individuals with Intellectual Disability

There is little in the general literature addressing the issue of using effort measures such as the TOMM with individuals with intellectual disability. According to Marshall and Happe (2007), the validity of tests for identifying inadequate effort when the test taker may be intellectually disabled is largely unknown, as there is very limited research on the use of effort tests with those diagnosed as having intellectual disability. Victor and Boone (2007) agree that individuals with intellectual disability are typically excluded from effort-test validation samples and that great caution should be used in effort-test interpretation for individuals with intellectual disability as the likelihood of false-positive error is probably quite high.

The main problem with the TOMM as well as other effort measures with respect to their applicability to individuals with intellectual disability has to do with the potential of a

false-positive identification. A false-positive identification occurs when an intellectually disabled defendant is thought to be not putting forth maximum effort on an effort test when, in fact, the defendant's poor performance is actually due to his or her cognitive impairment. As noted by Simon (2007), intellectually disabled individuals may well perform more poorly on malingering tests than do individuals of normal intelligence, even when making their best effort. Green (2011) agrees that most effort tests are limited by their low specificity in people with severe impairment, such as individuals with intellectual disability. According to Green, this means that there is a high risk of false positives in people with intellectual disability and other forms of severe impairment (e.g., dementia or other brain diseases).

Tombaugh (1997) conducted four experiments in order to establish the validity of the TOMM as an appropriate instrument in detecting individuals who exaggerate or deliberately fake memory impairment on psychological tests. The experiments also sought to establish a criterion score to assist in determining whether a person is malingering. Tombaugh concluded that the neurologically impaired individuals he studied showed exceptionally accurate scores on Trial 2 and on the Retention Trial of the TOMM regardless of age, education, or neurological dysfunction. None of the experiments, however, included individuals with intellectual disability. There is a concern, then, that a defendant involved in a pretrial forensic evaluation with an intellectual disability might score low on Trial 2 or the Retention Trial even though she or he is actually putting forth maximum effort.

A number of investigators have demonstrated that the TOMM in particular may not be appropriate for individuals with intellectual disability. Weinborn et al. (2003), for example, attempted to validate the TOMM using a differential prevalence design. In their review of Weinborn et al.'s study, Victor and Boone (2007) indicate that of the 61 study participants, 18% had intellectual disability and 15% were considered borderline intellectual functioning. Victor and Boone add that 27% of the 11 individuals in Weinborn et al.'s study with intellectual disability were misclassified by the TOMM. This finding suggests that the TOMM may not be appropriate for individuals with significant subaverage intelligence despite its applicability to other populations. Victor and Boone point out, however, that there were too few subjects in Weinborn et al.'s study to thoroughly evaluate the hypothesis that the TOMM may not be appropriate for individuals with intellectual disability.

Hurley and Deal (2006) studied a number of instruments, including the TOMM, designed to detect feigned cognitive impairment. They examined a sample of 39 individuals with varying degrees of intellectual disability. Hurley and Deal found that 41% of their sample of individuals with significant subaverage intellectual functioning scored below the cutoff for malingering (45) on Trial 2 of the TOMM. Based on these results, Hurley and Deal concluded that considerable caution should be exercised when using instruments such as the TOMM with individuals who have below-average intelligence. Similarly, in their review of instruments used to identify feigned intellectual disability, Victor and Boone (2007) concluded that the use of standard cutoffs for

instruments such as the TOMM may lead to an unacceptable rate of false-positive identifications within intellectually disabled samples.

In a study by Proto (2008), as expected, malingerers performed significantly worse than controls on the TOMM and three other malingering measures. However, Proto noted that the TOMM misidentified 36% of effortful intellectually disabled responders as potentially malingering. Thus, the results of this study support the notion that use of the TOMM Trial 2 cut-off score misclassified unacceptably high levels of intellectually disabled individuals. Weaknesses of this unpublished study are that all non-intellectually disabled participants were students attending LSU, which reduces the study's applicability to other populations. Also, the intellectually disabled group had a significantly larger proportion of African-Americans than the non-intellectually disabled group. As such, comparisons between intellectually disabled individuals and high-ability malingerers and controls cannot be fully attributable to the experimental conditions and are confounded by race. There was also a low sample size of intellectually disabled individuals, most of whom had dually-diagnosed conditions, which suggests that the intellectually disabled individuals' performance could have been confounded by their psychological disorders.

There are mixed findings in a study by Graue et al. (2007) regarding the applicability of the TOMM with intellectually disabled individuals. Utilizing the traditional cutoff score of 45, the TOMM significantly discriminated community volunteer malingerers from both community volunteer honest responders and an intellectually disabled group. Of note is that there were robust *d* scores for the crucial contrasts. Moreover, application of Tombaugh's recommended cutting score produced moderate hit rates, ranging from .75 for Trial 2 to .80 for the Retention Trial. The sensitivity rates were .80 for both trials and are comparable to somewhat higher than those reported in other studies. However, the specificity rate of .69 for Trial 2 is less than optimal.

Despite the aforementioned research, some studies have indicated that established TOMM cutoffs can be used with an intellectually disabled sample. Heinze and Purisch (2001), for example, examined a criminal population, many of whom had intellectual disability, who were suspected of feigning incompetence to stand trial. Although these authors found the TOMM and other effort tests to be valid in terms of their applicability to an intellectually disabled population, the specificity or rate of false-positive identifications was not examined.

Simon (2007) administered the TOMM to 21 adjudicated forensic inpatients who had been diagnosed with intellectual disability. The participants attained a mean score of 48.7 on Trial 2, with only 1 participant scoring below the standard cutoff for malingering. The participants attained a mean score of 49.4 on the Retention Trial, with no participants falling below the cutoff for malingering. According to Simon, the results of his study indicate that mildly intellectually disabled forensic patients can perform well on the TOMM. The specificity rate was 95.2% for Trial 2 and 100% for the Retention Trial.

Some weaknesses associated with Simon's (2007) investigation include the small sample size and that Simon did not report the TOMM's sensitivity rate. Moreover, the subjects involved in Simon's research had already been adjudicated. Thus, it is unclear the extent to which these research findings can be generalized to intellectually disabled defendants with active charges. It may be that selection bias was the determining factor in Simon's (2007) findings. A peer reviewer noted that, because the intellectually disabled patients in Simon's research were preselected to be competent (adjudicated), this decreases the likelihood that they would fail any cognitive measure.

Simon (2007) specifically mentioned that his findings are discrepant from the aforementioned study by Hurley and Deal (2006). According to Simon, one of the reasons for the discrepancy in these two investigations may have been that, in Hurley and Deal's study, the participants all resided in residential facilities for the intellectually disabled. However, in Simon's study, the participants were living in the community prior to their arrests. Simon postulated that there may be some intrinsic differences in intellectually disabled individuals who require residential care versus those that have the skills necessary to live in the community. Moreover, even though Simon's participants were told the test results would be used only for research purposes, Simon hypothesized that many may have believed that a good performance would be viewed positively by their treatment team. Simon also indicated that Hurley and Deal administered a battery of four malingering tests, which took as long as two hours to administer, whereas in his study only the TOMM was administered. Hurley and Deal's subjects' TOMM performance, then, might have been impacted by fatigue.

Use of Alternate Cut Scores

As noted, the problem with the TOMM and many other commonly used effort measures is the risk of false identification of intellectually disabled individuals as not putting forth maximum effort when they have actually tried their best. Victor and Boone (2007) reviewed the validity of using standard measures of effort with individuals with intellectual disability. A number of studies (e.g., Goldberg & Miller, 1986; Hayes, Hale, & Gouvier, 1997; Hayes, Hale, & Gouvier, 1998; Hurley & Deal, 2006; Reznick, 2005) have called into question the validity of using standard cutoffs for individuals with intellectual disability for the Rey 15-Item Test (Rey, 1964). Victor and Boone indicated that the use of standard cutoffs for most effort measures may lead to an unacceptable rate of false-positive identifications within samples of individuals with intellectual disability.

Marshall and Happe (2007) reported that it would not be surprising if the currently established cutoff scores on commonly used effort tests such as the TOMM were not valid. According to Marshall and Happe, the kinds of cognitive deficits commonly found in individuals with intellectual disability, such as deficits in attention, short-term memory, sequential processing, executive functions, and working memory, are very apt to adversely affect performance on many of these effort tests. For example, Marshall and

Happe lowered the cutoff for the Rey 15-Item Test to < 3 . However, this still incorrectly classified 17% of their intellectually disabled sample.

As a possible solution to the problem of false-positive identification, some researchers have proposed using alternative cut scores on effort tests. Theoretically, lowering an effort test's cut score would make the test easier to "pass," thus lowering the risk of an intellectually disabled person being falsely identified as not putting forth maximum effort when the individual was, in fact, trying his or her best. Based upon the aforementioned high false-positive rates among intellectually disabled individuals who are administered the TOMM, Hurley and Deal (2006) suggested using data from individuals with intellectual disability to determine cut scores for detecting feigning of an intellectual disability.

Graue et al. (2007) followed Hurley and Deal's (2006) recommendation by identifying alternative cut scores for individuals with intellectual disability. A review of the data indicated the same revised cut score (30) for both Trial 2 and the Retention Trial for the TOMM. Application of this cut score resulted in sensitivities of .56 for Trial 2 and .60 for the Retention Trial, with specificity at .96 for both trials and hit rates of .80 & .82, respectively. Given their small sample size, Graue et al. recommended that their results be cross-validated using known-groups methodologies.

Shandera et al. (2010) examined the TOMM and a number of other measures of effort with a sample of individuals with intellectual disability. On the basis of the TOMM's standard cut score of 45, specificity rates were .88 for Trial 2 and .92 for the Retention Trial. As was the case in Graue et al.'s (2007) research, in Shandera et al.'s study the TOMM again significantly discriminated community volunteer malingerers from community volunteer honest responders and an intellectually disabled group. Moreover, there were robust *d* scores for the crucial contrasts. Shandera et al. also cross validated the aforementioned study by Graue et al. using a modified cut score of 30 and found even better specificity. Unfortunately, the sensitivity rate of .24 (Trial 2) and .20 (Retention Trial) for the revised cut score are less than optimal. The fact that the intellectually disabled group had only 24 participants represents an important limitation of Shandera et al.'s study.

The research by Graue et al. (2007) and Shandera et al. (2010) suggests, then, that lowering the cut score of the TOMM from 45 to 30 decreased the number of false positive identifications. However, doing so reduced the TOMM's ability to correctly identify intellectually disabled individuals who are actually not putting forth their best effort.

Cut Scores as General Guidelines

In an article concerning experiments to validate the TOMM, Tombaugh (1997) indicated that in making a diagnosis of malingering, the criterion score should not be regarded as a rigid or fixed cutoff. Instead, Tombaugh suggested that the criterion score should be

viewed as a guideline, with the likelihood of malingering increasing as the score deviates further from the normative baseline for each diagnostic group. According to Tombaugh, although a low score on the TOMM may play a critical role in determining whether the results from a neurological assessment represent an individual's optimal level of performance, the diagnosis of malingering should never be made exclusively on the basis of a test score.

Marshall and Happe (2007) also warn against depending too much on test scores in the assessment of malingering. Other sources of information that should be considered include observed behaviors in testing and in the community, reliable information from collateral informants, and documented medical, occupational, and educational background history. As noted, research by Graue et al. (2007), which was cross-validated by Shandera et al. (2010), lends empirical support to Tombaugh's (1997) admonition against strict adherence to cutoff scores in the determination of cognitive malingering. These studies underscore the importance of being cautious in making broad conclusions about cognitive malingering for TOMM scores, particularly when examining individuals with intellectual disability. With this population, it is especially important for forensic clinicians to make use of multiple sources of data in the assessment of cognitive malingering.

Conclusions

Forensic clinicians who choose not to administer effort testing or other means of objectively examining feigned cognitive impairment place themselves at risk of making preventable mistakes in their conclusions and interpretations. In general, forensic examiners should not assume that defendants who claim to have neurocognitive dysfunction are not feigning. One way to reduce inaccuracies in the identification of feigned cognitive impairment in forensic evaluations is to regularly administer effort tests such as the TOMM. Such tests are important tools that forensic examiners should consider administering given the aforementioned prevalence rates of feigned cognitive impairment among individuals evaluated in forensic settings.

Although assessing feigned cognitive impairment in forensic settings is warranted, the current research review indicates that there are no effort tests that have been specifically designed to examine a lack of effort in individuals with intellectual disability. The TOMM is an instrument that has been established as an effective means of identifying a lack of effort in certain populations, but the applicability of this measure to individuals with intellectual disability has been questioned in the literature. As noted, the problem with the TOMM and many other effort tests used by forensic clinicians is that an unacceptably large number of intellectually disabled individuals are being falsely identified as not putting forth maximum effort when they are actually trying their best.

The floor effect in statistics refers to when data cannot take on a value lower than some particular number, called the floor (Everitt, 2002). A ceiling effect is when all participants reach the high end of the distribution (Pedhazur & Schmelkin, 1991).

Because of the floor effect, the absence of differences between individuals not putting forth maximum effort and those who are intellectually disabled on a measure like the TOMM may result from limits in the range of scores obtained on that measure (Kazdin, 1992). As noted by a peer reviewer, the most plausible explanation for why tests like the TOMM have too many false positives is that the floor of the test is too high, which results in a compressed distribution. The TOMM, in other words, is relatively difficult for individuals with intellectual disability. It is possible that differences between individuals not putting forth maximum effort and intellectually disabled individuals would be evident if the TOMM permitted a greater spread of scores.

Given the mixed findings in the research concerning using the TOMM with individuals who have intellectual disability, one option is to simply not administer this instrument or other effort tests to this population of defendants. This approach may reduce the chances of falsely identifying intellectually disabled defendants who appear to be feigning but who are actually putting forth maximum effort. However, an important disadvantage of not administering the TOMM is that despite the mixed research concerning the usefulness of this instrument with individuals who are intellectually disabled, scores that are equal to or exceed this cutoff can still provide important information to the forensic clinician.

When examinees earn scores at or above the cutoff of 45 on a floor-effect test such as the TOMM, it cannot be concluded that they displayed good effort, but only that they did not display bad effort (Faust & Ackley, 1998). Based on the available research, it appears prudent to conclude that intellectually disabled defendants who score at 45 or higher on the TOMM have not displayed poor effort on this instrument. Based on this information, it would be tempting to make the broader conclusion that the defendant must not be exhibiting poor effort on questions concerning legal matters. However, a peer reviewer noted that even though there is likely to be a correlation, a lack of poor effort on the TOMM does not necessarily imply a lack of poor effort on a test of legal knowledge. Nonetheless, the finding of a lack of poor effort on the TOMM provides useful information in forming hypotheses regarding the defendant's effort when asked questions about legal information.

Another reason forensic clinicians may want to consider administering an effort test such as the TOMM to intellectually disabled defendants is because the reviewed research suggests that scores that fall below chance responding appear likely to indicate a lack of effort even among individuals with intellectual disability (see Dean et al., 2008). Based upon the reviewed literature, it appears appropriate for forensic examiners to conclude that intellectually disabled defendants who score below 18 on the TOMM are probably not answering honestly. Tombaugh (1997) indicated that any score of lower than 25 represents below-chance performance and implies that the person actually knew some of the pictures were correct but intentionally picked the incorrect picture. In the TOMM manual, Tombaugh (1996) indicated that application of the binomial distribution shows that the 95% confidence interval for chance performance

on the TOMM ranges from 18 to 32. Thus, scores below 18 are unlikely to occur by chance.

As noted, Graue et al. (2007) and Shandera et al. (2010) suggested a new TOMM cutoff score of 30 for Trial 2 and the Retention Trial. This new cutoff score had high specificity even though there was a decrease in sensitivity. Based on Graue et al. and Shandera et al.'s research, intellectually disabled defendants who score below 30 on Trial 2 or the Retention Trial are likely putting forth inadequate effort.

The sample sizes in Graue et al. (2007) and Shandera et al.'s (2010) studies are relatively small and more cross validation with larger samples of individuals with intellectual disability is needed. Despite this caveat, when administering the TOMM to an intellectually disabled defendant, taking into account Graue et al. and Shandera et al.'s revised cut score enables the forensic examiner to make a more informed decision regarding the defendant's effort level.

The available research, then, provides forensic examiners with guidelines regarding the interpretation of TOMM scores at 45 and above, below 18, and between 18 and 29. However, there are an unacceptably high number of potential false positive identifications of a lack of effort when defendants obtain TOMM scores from 30 to 44. Given this state of affairs, it would appear prudent to conclude that a defendant's level of effort is indeterminate when his or her TOMM score is from 30 to 44.

Forensic examiners should always consider relevant collateral data regardless of the nature of the defendant's intellectual deficits. However, based on the current literature review, when examining intellectually disabled defendants, collateral data such as prior intellectual testing scores, educational history, and other relevant information as pointed out by Marshall and Happe (2007) is particularly important in order to make accurate conclusions about the defendant's level of effort.

Salekin and Doane (2009) indicated that historical factors should play a critical role in the assessment of feigned intellectual disability because, unlike many disorders, intellectual disability does not have a sudden onset. According to these authors, by definition, intellectual disability is a condition that begins during the developmental period and persists into adulthood. Thus, most individuals will show signs of the disorder prior to age 18. Collateral data from multiple sources can assist the forensic examiner in making more accurate determinations regarding the progression of a defendant's intellectual disability.

Until recently, no effort tests were available that have been extensively validated on intellectually disabled individuals. New research, however, suggests some promising results regarding alternative effort testing instruments for individuals with intellectual disability. As noted in the current literature review, Green et al. (2012) suggested that intellectually disabled individuals rarely fail the easy subtests on newer effort measures such as the WMT, MSVT and NV-MSVT. These instruments appear to be easy enough

such that intellectually disabled individuals usually do quite well. Based upon the aforementioned issues concerning the low number of actual mentally retarded individuals who were examined in Green et al.'s studies, however, additional research is needed to determine if the findings by Green et al. can be replicated.

Based on the information presented in this review of the literature on the TOMM, when a defendant involved in a forensic evaluation has an intellectual disability, the following steps are recommended to assess whether or not the person is putting forth inadequate effort:

1. Administer the TOMM.
2. Interpret TOMM scores on Trial 2 or the Retention Trial for individuals with known or suspected intellectual disability in the following manner:
 - 45 or above → inadequate effort unlikely
 - 30 to 44 → indeterminate
 - 18 to 29 → inadequate effort strongly suspected
 - 0 to 17 → inadequate effort very likely
3. Avoid strict adherence to cutoff scores when interpreting results from effort tests such as the TOMM.
4. Particularly in cases when the TOMM Trial 2 or Retention Trial score is in the indeterminate range, consider administering one or more of the alternative effort instruments mentioned in this article that have recently been validated on intellectually disabled populations.
5. Through interviews and pertinent collateral data, identify whether or not the defendant's cognitive deficits emerged before age 18.
6. Make a determination of malingering based on a variety of sources including effort testing and relevant collateral information.
7. Carefully communicate the defendant's testing results to the court, taking into account the defendant's status as an intellectually disabled individual.

Making use of these guidelines will assist forensic clinicians in better serving courts by improving the accuracy of interpretations of effort test scores for individuals with intellectual disability.

Acknowledgements

I want to thank my wife, Dorothy Ray, for her steady support. I also appreciate information provided by Paul Green, Richard Frederick, and Gregg Nigl. The views expressed in this article are those of the author and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the United States government.

Received January 22, 2012; revision submitted July 28, 2012; accepted July 30, 2012

References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- 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. doi: 10.1080/13825580600966391
- Ashendorf, L., Constantinou, M., & McCaffrey, R. J. (2004). The effect of depression and anxiety on the TOMM in community-dwelling older adults. *Archives of Clinical Neuropsychology, 19*, 125–130.
- 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. doi:10.1016/j.acn.2008.01.008
- Boone, K. B. (Ed.). (2007). *Assessment of feigned cognitive impairment: A neuropsychological perspective*. New York, NY: Guilford Press.
- Brockhaus, R. & Merten, T. (2004). Neuropsychological assessment of suboptimal performance: The Word Memory Test. *Der Nervenarzt, 75*, 882-887. doi: 10.1007/s00115-003-1593-4
- Bush, S. S., Ruff, R. M., Troster, A., Barth, J., Koffler, S. P., Pliskin, N. H., . . . Silver, C. H. (2005). NAN position paper: Symptom validity assessment: Practice issues and medical necessity. *Archives of Clinical Neuropsychology, 20*, 419–426. doi:10.1016/j.acn.2008.04.001
- Carone, D. (2008). Children with moderate/severe brain damage/dysfunction outperform adults with mild to no brain damage on the Medical Symptom Validity Test. *Brain Injury, 22*, 960-971. doi: 10.1080/02699050903133962
- D'Amato, C. P., & Denney, R. L. (2008). The diagnostic utility of the Rarely Missed Index of the Wechsler Memory Scale-Third Edition in detecting response bias in an adult male incarcerated setting. *Archives of Clinical Neuropsychology, 23*, 553–561. doi:10.1016/j.acn.2008.05.007
- Dean, A. C., Victor, T. L., Boone, K. B., Philpott, L. M., & Hess, R. A. (2008). Dementia and effort test performance. *The Clinical Neuropsychologist, 23*, 133-152. doi: 10.1080/13854040701819050
- Delain, S. L., Stafford, K. P., & Ben-Porath, Y. S. (2003). Use of the TOMM in a criminal court forensic assessment setting. *Assessment, 10*, 370–381. doi: 10.1177/1073191103259156

- Donders, J. (2005). Performance on the Test of Memory Malingering in a mixed pediatric sample. *Child Neuropsychology (Neuropsychology, Development and Cognition: Section C)*, *11*, 221-227. doi: 10.1080/09297040490917298
- Duncan, A. (2005). The impact of cognitive and psychiatric impairment of psychotic disorders on the Test of Memory Malingering. *Assessment*, *12*, 123-129. doi: 10.1177/1073191105275512
- Everitt, B. S. (2002). *The Cambridge dictionary of statistics* (2nd ed). Cambridge: Cambridge University Press.
- 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 (pp. 1-54). In C. R. Reynolds (Ed.). *Detection of Malingering During Head Injury Litigation*. New York: Plenum.
- Frederick, R. I. (1997). *Validity Indicator Profile manual*. NCS Assessments, Minnetonka, MN.
- Frederick, R. I. (2003). Review of the Validity Indicator Profile. *Journal of Forensic Neuropsychology*, *2*, 125-145. doi: 10.1300/J151v02n03_06
- Goldberg, J. O., & Miller, H. R. (1986). Performance of psychiatric inpatients and intellectually deficient individuals on a task that assesses the validity of memory complaints. *Journal of Clinical Psychology, Vol 42*, 792-795. doi: 10.1002/1097-4679(198609)42:5<792::AID-JCLP2270420519>3.0.CO;2-8
- Graue, L. O., Berry, D. T. R., Clark, J. A., Sollman, M. J., Cardi, M., Hopkins, J., & Werline, D. (2007). Identification of feigned mental retardation using the new generation of malingered detection instruments: Preliminary findings. *Clinical Neuropsychologist*, *21*, 921-942. doi: 10.1080/13854040600932137
- Green, P. (2003). *Word Memory Test for Windows: User's manual and program*. Edmonton, Alberta, Canada: Author. (Revised 2005).
- Green, P. (2004). *Test manual for the Medical Symptom Validity Test*. Edmonton: Green's Publishing.
- Green, P. (2007). Spoiled for choice: Making comparisons between forced-choice effort tests. In K. B. Boone (Ed.) *Assessment of feigned cognitive impairment: A neuropsychological perspective*, (pp. 50-77). New York: Guilford.
- Green, P. (2008). *Test manual for the Nonverbal Medical Symptom Validity Test*. Edmonton: 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. doi: 10.1080/09084282.2010.523365
- Green, P., Allen, L. M., & Astner, K. (1996). *The Word Memory Test: A user's guide to the oral and computer-administered forms, US version 1.1*. Durham, NC: CogniSyst.
- Green, P., & Astner, K. (1995). *Manual for the Oral Word Memory Test*, Cognisyst, N.C.
- Green, P., Flaro, L., Brockhaus, R., & Montijo, J. (2012). Performance on the WMT, MSVT & NV-MSVT in children with intellectual disabilities and in adults with mild traumatic brain injury. In C. R. Reynolds & A. Horton (Eds.). *Detection of malingering during head injury litigation*, 2nd ed. (pp. 201-219). New York: Plenum Press.
- Greve, K. W., Ord, J. S., Bianchini, K. J., & Curtis, K. L. (2009). Prevalence of malingering in patients with chronic pain referred for psychologic evaluation in a medico-legal context. *Archives of Physical Medicine and Rehabilitation, 90*, 1117-1126. doi: 10.1016/j.apmr.2009.01.018
- Hayes, J. S., Hale, D. B., & Gouvier, W. D. (1997). Do tests predict malingering in defendants with mental retardation? *Journal of Psychology, 131*, 575-576. doi: 10.1207/s15324826an0501_4
- Hayes, J. S., Hale, D. B., & Gouvier, W. D. (1998). Malingering detection in a mentally retarded forensic population. *Applied Neuropsychology, 5*, 33-36. doi: 10.1207/s15324826an0501_4
- Heilbronner, R. L., Sweet, J. J., Morgan, J. E., Larrabee, G. J., Millis, S. R., & Conference Participants (2009). American Academy of Clinical Neuropsychology consensus conference statement on the neuropsychological assessment of effort, response bias, and malingering. *The Clinical Neuropsychologist, 23*, 1093-1129. doi: 10.1080/13854040903155063
- Heilbrun, K. (2001). *Principles of forensic mental health assessment*. New York: Kluwer Academic/Plenum Publishers.
- Heinze, M. C., & Purisch, A. D. (2001). Beneath the mask: Use of psychological tests to detect and subtype malingering in criminal defendants. *Journal of Forensic Psychology Practice, 1*, 23-51. doi: 10.1300/J158v01n04
- Henry, M., Merten, T., Wolf, S., & 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. doi: 10.1080/13803390902791653
- Hurley, K. E., & Deal, W. P. (2006). Assessment instruments measuring malingering used with individuals who have mental retardation: Potential problems and issues. *Mental Retardation*, 44, 112–119. doi: [http://dx.doi.org/10.1352/0047-6765\(2006\)44\[112:AIMMUW\]2.0.CO;2](http://dx.doi.org/10.1352/0047-6765(2006)44[112:AIMMUW]2.0.CO;2)
- Inman, T. H., Vickery, C. D., Berry, D. T. R., Lamb, D., Edwards, C., & Smith, G. T. (1998). Development and initial validation of a new procedure for evaluating adequacy of effort during neuropsychological testing: The Letter Memory Test. *Psychological Assessment*, 10, 128–139.
- Iverson, G. L., & Binder, L. M. (2000). Detecting exaggeration and malingering in neuropsychological assessment. *Journal of Head Trauma Rehabilitation*, 15, 829-858. doi: [10.1037/1040-3590.10.2.128](https://doi.org/10.1037/1040-3590.10.2.128)
- Iverson, G. L. (2006). Ethical issues associated with the assessment of exaggeration, poor effort, and malingering. *Applied Neuropsychology*, 13, 77–90. doi: 10.1207/s15324826an1302_3
- Kazdin, A. E. (1992). *Research design in clinical psychology* (2nd ed.). Boston: Allyn & Bacon.
- Larrabee, G. J. (Ed.). (2007). *Assessment of malingered neuropsychological deficits*. New York, NY: Oxford.
- Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.
- Marshall, P., & Happe, M. (2007). The performance of individuals with mental retardation on cognitive tests assessing effort and motivation. *The Clinical Neuropsychologist*, 21, 826-840. doi: 10.1080/13854040600801001
- Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. *Journal of Clinical and Experimental Neuropsychology*, 24, 1094-1102. doi: 10.1076/jcen.24.8.1094.8379
- Morgan, J. E., & Sweet, J. J. (2009). *Neuropsychology of malingering casebook*. New York, NY: Psychology Press.
- Nitch, S. R., & Glassmire, D. M. (2007). Non-forced-choice measures to detect noncredible cognitive performance. In K. B. Boone (Ed.) *Assessment of feigned cognitive impairment: A neuropsychological perspective*, (pp. 78-102). New York: Guilford.

- Pankratz, L., & Binder, L. M. (1997). Malingering on cognitive measures. In R. Rogers (Ed.) *Clinical assessment of malingering and deception (2nd ed.)*, (pp. 223-236). New York: Guilford.
- Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Proto, D. A. (2008). *The detection of malingered mental retardation in high and low cognitive ability individuals* (Master's thesis). Retrieved from <http://etd.lsu.edu/docs/available/etd-04092008-154108/>
- Rabin, L. A., Barr, W. B., & Burton, L. A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA Division 40 members. *Archives of Clinical Neuropsychology*, *20*, 33-65. doi:10.1016/j.acn.2004.02.005
- Ray, C. L. (2009). The importance of using malingering screeners in forensic practice. *Journal of Forensic Psychology Practice*, *9*, 138-146. doi: 10.1080/15228930802575508
- Rees, L. M., Tombaugh, T. N., & Boulay, L. (2001). Depression and the Test of Memory Malingering. *Archives of Clinical Neuropsychology*, *16*, 501–506. doi: [10.1016/S0887-6177\(00\)00064-0](https://doi.org/10.1016/S0887-6177(00)00064-0)
- Rey, A. (1941). L'Examen psychologie dans las cas d'encephalopathie traumatique. *Archives de Psychologie*, *23*, 286–340.
- Rey, A. (1964). *L'examen Clinique en psychologie*. Paris: Presses Universitaires de France.
- Reznick, L. (2005). The Rey 15-Item Memory Test for malingering: A meta-analysis. *Brain Injury*, *19*, 539-543.
- Rogers, R. (1997). *Clinical assessment of malingering and deception (2nd ed.)*. New York: The Guilford Press.
- Rogers, R., & Bender, S.D. (2003). Evaluation of malingering and deception. In: A.M. Goldstein and I.B. Weiner, (Eds.) *Handbook of psychology: Forensic psychology* Vol. 11, (pp. 109-129). Wiley, New York.
- Rogers, R., Sewell, K. W., Martin, M. A., & Vitacco, M. J. (2003). Detection of feigned mental disorders: A meta-analysis of the MMPI-2 and malingering. *Assessment*, *10*, 160-177. doi: 10.1177/1073191103010002007

- Rohling, M. L., Larrabee, G. J., Greiffenstein, M. F., Ben-Porath, J. S., Lees-Haley, P., Green, P., & Greve, K. W. (2011). A misleading review of response bias: Comment on McGrath, Mitchell, Kim, and Hough (2010). *Psychological Bulletin*, *137*, 708-712. doi: 10.1016/j.acn.2003.12.004
- Salekin, K. L., & Doane, B. M. (2009). Malingering intellectual disability: The value of available measures and methods. *Applied Neuropsychology*, *16*, 105-113. doi: 10.1080/09084280902864485
- Simon, M. J. (2007). Performance of mentally retarded forensic patients on the Test of Memory Malingering. *Journal of Clinical Psychology*, *63*, 339-344. doi: 10.1002/jclp.20351
- 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. doi: 10.1093/arclin/acp074
- Shandera, A. L., Berry, D. T., Clark, J. A., Schipper, L. J., Graue, L. O., & Harp, J. P. (2010). Detection of malingered mental retardation. *Psychological Assessment*, *22*, 50-56. doi 10.1037/a0016585
- Slick, D. J., Hopp, G., Strauss, E., & Spellacy, F. J. (1996). Victoria Symptom Validity Test: Efficiency for detecting feigned memory impairment and relationship to neuropsychological tests and validity scales. *Journal of Clinical and Experimental Neuropsychology*, *18*, 911-922. doi: 10.1080/01688639608408313
- Slick, D. J., Sherman, E. M. S., Iverson, G. L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist*, *13*, 545-561.
- Tan, J. E., Slick, D. J., Strauss, E. & Hultsch, D. F. (2002). How'd they do it? Malingering strategies on symptom validity tests. *The Clinical Neuropsychologist*, *16*, 495-505. doi: dx.doi.org/10.1076/clin.16.4.495.13909
- Teichner, G., & Wagner, M. T. (2004). The Test of Memory Malingering (TOMM): normative data from cognitively intact, cognitively impaired, and elderly patients with dementia. *Archives of Clinical Neuropsychology*, *19*, 455-464. doi: 10.1016/S0887-6177(03)00078-7.
- Tombaugh, T. N. (1996). *Test of Memory Malingering*. North Tonawonda, 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. doi: [10.1037/1040-3590.9.3.260](https://doi.org/10.1037/1040-3590.9.3.260)

- Victor, T. L., & Boone, K. B. (2007). Identification of feigned mental retardation. In K. B. Boone (Ed.) *Assessment of feigned cognitive impairment: A neuropsychological perspective*, (pp. 310-345). New York: Guilford.
- Vitacco, M., Rogers, R., Gabel, J., & Munizza, J. (2007). An evaluation of malingering screens with competency to stand trial patients: A known-groups comparison. *Law and Human Behavior*, *31*, 249-260. doi: [10.1007/s10979-006-9062-8](https://doi.org/10.1007/s10979-006-9062-8)
- Weinborn, M., Orr, T., Woods, S. P., Conover, E., & Feix, J. (2003). A validation of the Test of Memory Malingering in a forensic psychiatric setting. *Journal of Clinical and Experimental Neuropsychology*, *25*, 979–990. doi: 10.1076/jcen.25.7.979.16481
- Yanez, Y. T., Fremouw, W., Tennant, J., Strunk, J., & Coker, K. (2005). Effects of severe depression on TOMM performance among disability-seeking outpatients. *Archives of Clinical Neuropsychology*, *21*, 161–165. doi: 10.1016/j.acn.2005.07.009