Chapter 4

Negative Response Bias and Malingering during Neuropsychological Assessment in Criminal Forensic Settings

ROBERT L. DENNEY

Interest in malingering within clinical psychology has exploded in the last 10 years. This trend is demonstrated with the results of PsychLit searches for "malingering" among peer reviewed journal articles (Figure 4.1). This trend is likely in response to the recognition that incentive to demonstrate mental impairment in forensic settings can be tremendous. Whereas much of the research has dealt with civil forensic issues, the motivation is no less significant in criminal settings. Rather than large monetary awards, criminal defendants have the understandable motivation to avoid long prison sentences or even execution. Indeed, research exists to suggest criminal defendants facing more serious charges are more likely than those facing lesser charges to exaggerate deficits (Weinborn, Orr, Woods, Conover, & Feix, 2003). Even sentenced inmates oftentimes have incentive to spend their prison sentence in hospitals. Female staff are more prevalent, and greater opportunity for acquiring drugs of abuse exists. Inmates may also find themselves facing dangers that force them to seek a move to another facility. Feigning psychiatric or physical disease and transferring to a hospital for specialized assessment is a possible solution.

There has also been increased recognition that for cognitive test results to be valid, effort on behalf of the examinee is required. Although this recog-

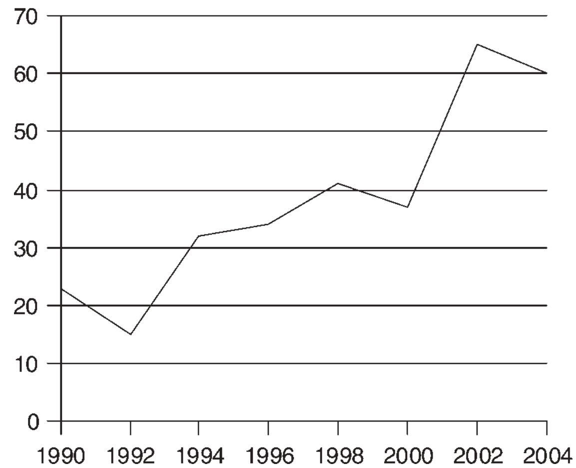

FIGURE 4.1. Number of peer-reviewed malingering-related papers published by year.

nition is not new (e.g., Anastasi, 1961), widespread appreciation of the effect of effort on cognitive test results is more recent. Green, Rohling, Lees-Haley, and Allen (2001) evaluated the performance of a series of 904 patients with head injury and neurological impairment on neuropsychological tests. They found a correlation (r = .73) between a well-validated symptom validity test (Word Memory Test [WMT]; Green, 2003) and the overall neuropsychological test battery mean. More strikingly, they found that poor effort suppressed the test battery mean 4.5 times more than did moderate to severe brain injury. In effect, effort has a much greater effect on test scores than does brain injury or a neurological condition. Neuropsychological tests are achievement oriented and require examinee effort to obtain valid results. Consequently, they are extremely susceptible to the influence of poor effort, exaggeration, and feigning.

Researchers also demonstrated that it is virtually impossible to detect malingering reliably without using indices designed to identify it (Faust, Hart, Guilmette, & Arkes, 1988; Heaton, Smith, Lehman, & Vogt, 1978). A new appreciation for the effects of effort on assessment resulted in admonishments to change the manner in which neuropsychological testing is performed in forensic settings. In the Foreword to the WMT (Green, 2003, p. iv), Paul Lees-Haley noted, "Neuropsychological assessments are no longer complete without evaluation of effort." This sentiment has risen to official levels, as evidenced by the position paper on symptom validity testing from the National Academy of Neuropsychology (Bush et al., 2005), which asserts the importance of objectively evaluating evaluee effort in all forensic-related neuropsychological assessments.

The purpose of this chapter is to review important issues regarding assessment of effort and malingering in the context of criminal forensic neuropsychological evaluations. There is great wealth of research dealing with psychiatric symptom exaggeration. Although relevant to neuropsychologists evaluating criminal defendants, the general psychiatric exaggeration literature is too broad to cover in this single chapter. For a broad review regarding detection of deception and malingering, the reader is referred to Boyd, McLearen, Meyer, and Denney (2007) and the classic work by Rogers (2008). For a more comprehensive review and analysis of poor effort and malingering in neuropsychology, I recommend both Boone (2007) and Larrabee (2007). Here I discuss the nature of exaggeration and malingering as it relates to neuropsychology in general, but with specific emphasis on these issues in the criminal forensic setting. I present what is known about the prevalence of malingering; review the accepted classification system and specific measures and strategies for detecting poor effort; address issues of remote memory loss; and conclude with areas of concern and recommendations for future research.

The Nature of Malingering

Malingering, as it relates to neuropsychology, is a clinical determination based on test results and behaviors within a contextual framework. Malingering can be viewed as requiring two components: response bias and conscious intention. *Response bias* is a systematic pattern of performance, such that obtained results do not accurately reflect what the tests are purported to measure. Response bias discloses nothing about the reason for the atypical performance. An example of negative response bias could be something as simple as fatigue. In neuropsychology, we are typically concerned about the possibility of negative response bias rather than positive response bias, because people cannot "fake good" on achievement-oriented testing (it is called "cheating"). Once negative response bias is documented, the conclusion of malingering requires the determination that significant secondary gain is influencing evaluee performance. Because it is a two-part process, I address aspects of negative response bias, then turn to malingering.

Negative Response Bias

Negative response bias (NRB) is a systematically poor performance that is not consistent with genuine neurocognitive compromise. NRB makes no reference to conscious or unconscious motivations. It may arise due to fatigue, anxiety, or the presence of significant psychiatric disturbance. NRB may also arise because of intentional attempts to misrepresent abilities. NRB specifically impacts achievement-oriented test results. Intelligence, academic abilities, and neurocognitive functions, such as attention, concentration, language, visual–spatial learning, memory, abstract reasoning, and problem solving, could all be impacted.

NRB can also take the form of exaggerated self-reports of cognitive impairments or the influence of claimed cognitive impairments of daily functioning. This form of NRB can be identified through inconsistencies between self-report and real-world functioning, and through inconsistencies between real-world functioning and normative-based results on self-report measures.

NRB can also be applied to functioning outside the area of cognitive abilities. Exaggeration of psychosis, anxiety, depression, and dissociation can be presented through self-report measures and overt behavior as well. Studies of the Minnesota Multiphasic Personality Inventory–2 (MMPI-2; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) among forensic populations reveals that this type of exaggeration is less common for neuropsychological evaluations in civil forensic settings (e.g., personal injury), where individuals oftentimes minimize mental health difficulties while highlighting somatic and neurocognitive disability (Lees-Haley, Iverson, Lange, Fox, & Allen, 2002). Nonetheless, NRB can occur with any combination of intellectual or cognitive dullness, specific neurocognitive deficits (e.g., learning or memory), psychiatric manifestations, and general somatic concerns (e.g., neurological and pain). Larrabee (2003b) demonstrated correlations between somatic concerns and cognitive complaints among individuals seeking personal injury claims.

Within neuropsychology, test procedures that have proven reliability and validity may yield scores that are not reliable or valid for a particular individual. Questions can arise because of significant inconsistencies in the test data. These inconsistencies may occur among the following areas:

- Neuropsychological domains (e.g., impaired attention with normal memory).
- Test scores and suspected etiology (e.g., impaired IQ with normal memory in hypoxia).
- Test scores and documented severity of injury (e.g., performance levels characteristic of prolonged coma in traumatic brain injury [TBI] with no actual loss of consciousness).
- Test scores and behavioral presentation (e.g., failure on measures of recent and remote memory, but ability to report accurate clinical history).

A valuable conceptual model developed by Frederick and presented by Frederick, Crosby, and Wynkoop (2000) is displayed in Figure 4.2. It reveals the fourfold nature of an individual's potential performance across two continua, effort and motivation. Individuals motivated to perform well, and who give their maximum effort, are considered to have provided a compliant and valid performance (upper right quadrant). The remaining three quadrants each represent some form of NRB and are not considered valid reflections of true ability. Individuals who may be motivated to perform well but do not put forth their best effort are classified as *careless*. Their carelessness represents NRB, but in most instances does not represent intentional misrepresentation of ability. Rather, this quadrant likely reflects NRB caused by excessive fatigue or distraction, such as that caused by psychosis or severe headache. However, it may reflect poor effort. The remaining two NRB quadrants indicate individuals who are motivated to perform poorly. Many such individuals do not assert much effort, and their results are best classified as *irrelevant*. These individuals present a picture of cooperation by completing the test, but they are responding in a manner independent of item content (e.g., a random response pattern on the MMPI-2 [Butcher et al., 1989]; irrelevant response pattern on the Validity Indicator Profile [VIP; Frederick, 2003]). Such individuals may later be classified as malingering. At the very

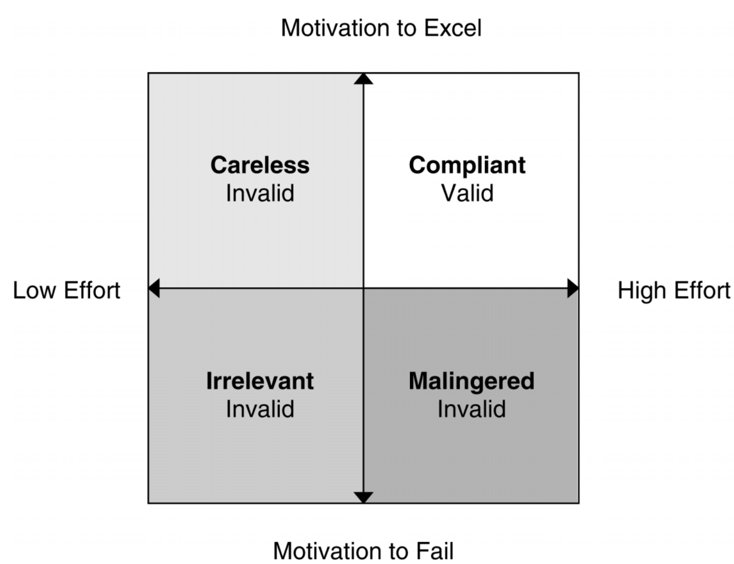

FIGURE 4.2. Fourfold nature of examinee performance. From Frederick (2003). Copyright 2003 by Pearson Assessments. Reprinted by permission.

least, they are noncooperative with the evaluation. The final quadrant represents malingering in its truest form. These individuals are not only motivated to perform poorly but they also put forth effort to perform poorly. This pattern of performance is most clearly demonstrated with below-random responding on forced-choice tests.

Variable effort and motivation can sometimes be secondary to factors outside the examinee's conscious intent or control, such as those found in severe depression and anxiety, as well as somatoform disorder. They can also be under conscious intent and control, such as factors found in factitious disorder and malingering. Factitious disorder and malingering share intentional, volitional distortion or misrepresentation of symptoms, but factitious disorder requires the determination that the subject has a psychological need to assume the sick role (primary gain). Malingering requires the contextual determination that the subject's motivation is for secondary gain, such as obtaining financial compensation or avoiding criminal prosecution.

Malingering

In contrast to the other forms of NRB, *malingering* is the "intentional production of false or grossly exaggerated physical or psychological symptoms, motivated by external incentives such as avoiding military duty, avoiding work, obtaining financial compensation, evading criminal prosecution, or obtaining drugs" (American Psychiatric Association, 2000, p. 739). Rogers (2008) provided a discussion of the possible explanatory models, noting a history of believing that individuals who malinger are mentally disordered. This *pathological* model arose from the explanations of behavior espoused by psychoanalytic thinkers. Rogers also described a second explanatory model that focused on individual character. In contrast to the pathological model, in which malingerers deceive because of some intrapsychic pathology, the *criminological* model emphasizes the propensity for antisocial and psychopathic personalities to lie, cheat, and steal.

The pathological model arose because of the clinical observation that many people considered to be malingering were also noted to have antisocial or sociopathic personality disorders. This view espouses that people with antisocial personality disorder are more prone to malinger because of their propensity to lie. The model, which appears to lack substantial empirical support, appears to be the current view of contributors to the *Diagnostic and Statistical Manual of Mental Disorders*, fourth edition, text revision (DSM-IV-TR; American Psychiatric Association, 2000), where it is suggested that malingering should be "strongly suspected" when any combination of the following are noted: (1) a medicolegal context; (2) marked discrepancy between subjective claims and objective findings; (3) lack of cooperation; and (4) presence of antisocial personality disorder (p. 739). In contrast to these pathological and "bad" models of malingering, Rogers (2008) also proposed an *adaptational* model, whereby malingering individuals are considered to be basically normal people attempting to meet their needs in adversarial circumstances. These individuals perform a cost–benefit analysis when confronted with an assessment perceived as indifferent to or in opposition to their needs. Research seems to support this finding, because the base rates for NRB and malingering appear to be high in adversarial contexts such as personal injury litigation (Larrabee, 2003a), Social Security Disability litigation (Chavez, Abrahams, & Kohlmaier, 2007), and criminal litigation (Ardolf, Denney, & Houston, 2007; Denney, 2007). Furthermore, research suggests that as the potential for greater gains increases, the rates of NRB increase (Bianchini, Curtis, & Greve, 2006). Personal stakes can be high in the adversarial context, and for many people there are no other perceived viable alternatives than to feign or exaggerate illness.

Prevalence Rates

Diagnostic accuracy is dependent, in large part, on one's awareness of prevalence rates for the suspected condition in the relevant population (Baldessarini, Finkelstein, & Arana, 1983; Elwood, 1993). It is no less a concern in the area of malingering detection (Gouvier, Hayes, & Smiroldo, 1998; Rosenfeld, Sands, & Van Gorp, 2000). Indeed, ignorance of such information appears to contribute to the difficulty of malingering detection and misclassification in general (Labarge, McCaffrey, & Brown, 2003; Rosenfeld et al., 2000). There should be no expectation that malingering styles or base rates will remain stable across evaluation settings. Furthermore, there should be no expectation that rates are the same in civil and criminal areas. Rates may even vary within the criminal setting depending on the context (sentenced inmates seeking hospitalization vs. criminal pretrial defendants) and severity of crime and sentence (potential 5-year vs. life sentence). This type of "dose–response relationship" has been found in the civil litigation (workers compensation) arena (Bianchini et al., 2006). Before addressing what little is known about malingering prevalence rates in the criminal setting, I briefly review what is known about those rates in the civil setting.

NRB in Civil Forensic Settings

Mittenberg, Patton, Canyock, and Condit (2002) surveyed members of the American Board of Clinical Neuropsychology regarding the number of times they diagnosed malingering over the course of the previous year. Mean estimated rates of "probable malingering or symptom exaggeration" were between 28 and 33% for personal injury, disability, and workers' compensation cases. This survey appears to have relied on clinicians' recollection of the number of times they identified NRB. This finding may be weakened by fading memory, as well as possibly less than thorough NRB evaluation efforts by a portion of the membership in the first place. Larrabee (2003a) combined results of 11 different malingering studies of over 1,350 civil litigants (for more information about these studies, see Ardolf et al., 2007). He found an overall 40% base rate of NRB among civil forensic cases, with a range from 15 to 64.3%. Consistent with those findings, Chavez and colleagues (2007) evaluated 232 consecutive Social Security Disability referrals and found an NRB rate of 55.8% using the Test of Memory Malingering (TOMM; Tombaugh, 1996), and 61.4% using the Medical Symptom Validity Test (MSVT; Green, 2004). These studies suggest substantial rates of NRB in the civil forensic arena, but much less literature exists regarding NRB in the criminal forensic setting, particularly related to neurocognitive concerns.

NRB in Criminal Forensic Settings

Little is known about the prevalence of NRB in the criminal forensic arena, and much of that information deals with exaggeration of general psychiatric issues. Rogers (1986) identified rates of 20% for suspected malingering and 4.5% for "definite malingering" among criminal defendants. Cornell and Hawk (1989) considered 8% of their pretrial criminal defendants to be feigning psychosis. Lewis, Simcox, and Berry (2002) identified a rate of 31.4% for feigned psychiatric presentation among pretrial criminal defendants.

These studies dealt with general psychiatric presentation. Even fewer studies exist for exaggerated neurocognitive deficit in the criminal population. The Mittenberg and colleagues (2002) survey included estimates of neuropsychological exaggeration for criminal referrals as well, with the mean falling between 19 and 23%, depending on whether cases were referred by the defense or the prosecution. Rates of malingering conclusions were lower when cases were referred by the defense. Frederick and Denney (1998) estimated a 25% prevalence of below-random responding assessed by forced-choice recognition tests for individuals claiming amnesia in a criminal forensic setting. Ardolf and colleagues (2007) reviewed data from 105 presentence criminal defendants who had been referred by the U.S. District Courts for mental health evaluation of their competency to undergo criminal proceedings. These cases were unique in that all had some question regarding their cognitive status and underwent neuropsychological evaluation. These cases varied greatly in their referral diagnoses, as one would expect in a consecutive series of cases. Each defendant was evaluated with multiple measures of NRB (free-standing and imbedded validity indices). Although all cases did not receive all of the same measures, use of multiple measures allowed for use of various NRB and malingering classification schemes, including Slick, Sherman, and Iverson's (1999) multimodal, multidimensional classification of malingered neurocognitive dysfunction (MND).

Ardolf and colleagues (2007) found that 89.5% of their criminal defendants scored positively on at least one measure of NRB, 70.5% of the defendants scored positively on two or more indicators, and 53% of cases scored positively on three or more indicators. Use of the Slick and colleagues (1999) criteria resulted in a 54% rate of MND classification (combined 32% probable and 22% definite; i.e., below-random performance). The Slick and colleagues classification takes into account possible false-positive findings, as well as indications of invalidity of test results apart from the validity indices (e.g., discrepancy between test data and observed behavior, and indications of invalidity due to self-report). As a result, the classification system attempts to avoid both false positives and false negatives, and provides what appears to be a robust estimate of neurocognitive malingering in the criminal forensic venue.

I (Denney, 2007) returned to the previous neuropsychological data, which had increased from 105 to 118 pretrial criminal defendants during the intervening time. I then extracted only those cases with complaints of mild to moderate TBI and obtained data on 67 of them (22.4% moderate and 77.6% mild). As with the larger database, defendants were administered a variety of NRB detection strategies: 73.1% of the defendants were positive on two or more NRB indices (previously 70.5%). Use of the Slick and colleagues (1999) classification system resulted in a combined rate of probable and definite MND of 62.7% rather than the 54.3% rate of the previous heterogeneous pathology data. These rates are similar to the NRB rates found in Social Security Disability evaluations by Chavez and colleagues (2007). Although one can argue that the Slick and colleagues classification system has yet to be validated as an accurate predictor of MND, the results clearly indicate a greater than 50% rate of NRB occurrence in presentence criminal defendants referred for neuropsychological evaluation.

Why Is the Base Rate of Malingering Important?

Diagnostic accuracy is clearly an important endeavor. Psychological tests help in this pursuit of accuracy; however, even the most reliable and valid test has error. That error interacts with the base rates of the condition in any given population. Diagnostic accuracy, then, requires an understanding of the predictive characteristics of the tests we use within a specific population (Baldessarini et al., 1983; Gouvier et al., 1998; Rosenfeld et al., 2000). Neglecting base rate information appears to contribute to the difficulty of malingering detection and misclassification in general (Labarge et al., 2003; Rosenfeld et al., 2000). Identifying the base rate for malingering among criminal defendants referred for neuropsychological evaluations is important, because it sheds light on the predictive value of the malingering detection method used. Establishing a base rate in one's setting may be difficult initially, because it requires (1) the use of NRB indicators on a regular basis over an extended period of time, and (2) a systematic review of that data. Results from my NRB test data (Ardolf et al., 2007; Denney, 2007) suggest that the base rate in the population of male criminal defendants referred for neuropsychological assessment is over 50%.

The importance of a rise in prevalence rates beyond 50% cannot be overstated, because it has relevance for classification confidence. The difference in diagnostic certainty, as it relates to classification accuracy between a low base rate phenomenon and a high base rate phenomenon, is easily demonstrated. First, we must remember test sensitivity (i.e., cases with the diagnosis that have a positive test finding[1]) and specificity (nonimpaired cases that have a negative test finding[2]) are unique to the instrument and do not vary based on prevalence of disorder. The hit rate index refers to the overall correct classification ability of the instrument and is identified as (true positives + true negatives)/*N*. Predictive value statistics, such as positive predictive value (PPV) and negative predictive value (NPV), incorporate the prevalence of disorder (Baldesserini et al., 1983). PPV is determined by true positives/(true positives + false positives). NPV is determined by true negatives/(true negatives + false negatives). Baldesserini and colleagues (1983) provide these formulas for computing PPV and NPV, incorporating test sensitivity (x), specificity (y), and base rate (prevalence, p):

$$\text{PPV} = (px)/[(px) + (1 - p)(1 - y)]$$
$$\text{NPV} = [(1 - p)y]/[(1 - p)y + p(1 - x)]$$

Although seemingly complicated, it is really not difficult. Let us assume a reasonably accurate malingering instrument, or combination of instruments, that have a sensitivity of .80 and specificity of .90. In a base rate setting of 20% positive findings, PPV is .667 and NPV is .947. Here, one has more diagnostic confidence in the negative findings. This confidence significantly changes as the base rate changes, however. In a base rate setting of 70% positive findings, PPV is .949 and NPV is .659. Here we have much more confidence in our positive findings, and strikingly less confidence in negative findings. This change in our confidence occurred because the prevalence of the condition (here, NRB) changed from well below 50% to above 50%. As prevalence nears 50%, our confidence in positive and negative findings more closely approximates the sensitivity and specificity of the measure without the influence of prevalence. Knowledge of prevalence rates does not change the accuracy of the test, but it does change *confidence* in the test findings. My data suggest that the base rate of MND among criminal defendants with neurocognitive concerns is well over 50% (possibly over 70%). In this setting, positive findings on NRB indices are quite convincing. Negative results, however, are not as definitive.

Malingered Neuropsychological Dysfunction Classification

Slick and colleagues (1999; hereafter called Slick criteria) proposed diagnostic criteria for MND that includes possible, probable, and definite classifications. This multidimensional approach incorporates several criteria: presence of substantial external incentive; evidence from neuropsychological testing, including negative response bias; and evidence from self-report. Finally, it incorporates a rule-out criterion for the evaluator to consider possible psychological, neurological, or psychiatric reasons for the unusual behavior. Evidence coming from neuropsychological testing is considered for NRB. The finding of NRB does not equate to malingering, it is one of the requirements for malingering. MND requires the presence of a substantial external incentive. *Probable MND* is defined as positive findings on one or more well-validated psychometric tests or indices designed to measure exaggeration or fabrication of cognitive deficits, such that it is consistent with feigning along with inconsistencies from other sources. *Definite MND* is defined as below-chance performance ($p < .05$) on one or more forced-choice measures of cognitive function. Millis (2004) reviewed the Slick criteria and developed a decision tree to assist evaluators in applying the multidimensional model in the assessment setting (Figure 4.3). Larrabee, Greiffenstein, Greve, and Bianchini (2007) discuss recommendations for improvement and validation of this classification system.

Below-Random Performance and Definite Malingering

The Slick criteria differentiate between NRB and malingering. This distinction is important, because the conclusion of malingering must always be clinical; that is, it takes into consideration contextual aspects of the evalua-

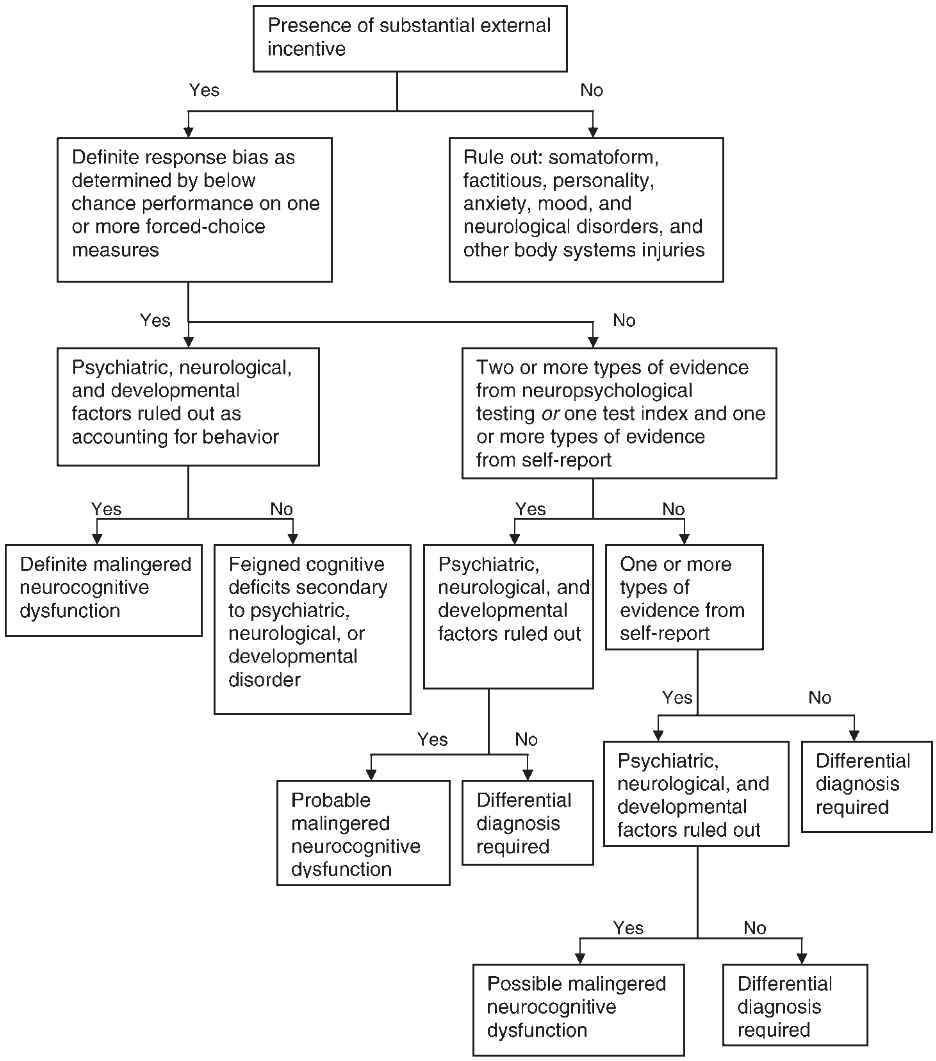

FIGURE 4.3. Diagnostic decision tree for malingered neurocognitive dysfunction. From Millis (2004). Copyright 2004 by W. B. Saunders. Adapted by permission.

tion and clinical judgment. Definite NRB is demonstrated from test results (i.e., below-random performance on a forced-choice measure), whereas definite MND requires a clinical conclusion, even in the presence of below random performance. See Frederick and Speed (2007) for a cogent discussion about the meaning of below random findings on forced-choice tests. The field is currently debating whether below random performance *definitively* indicates willful misrepresentation. Boone (2007) notes the history of the paradigm and points out that it was initially developed to identify conversion disorders. She also notes that published clinical cases demonstrated severe cognitive impairment and no incentive to feign deficits, but significantly below random performance on two-alternative, forced-choice testing. Her argument raises two issues: (1) It is theoretically possible for individuals to perform below random because of purportedly unconscious processes; (2) severely compromised individuals can perform below random occasionally. The second issue might be straightforward and has a statistical explanation. Using $p < .05$ as the cutoff for below-random performance indicates there will be five false-positive findings for every 100 performances from individuals who have *truly no cognitive capability* on the test. We must keep in mind, however, that we never administer these tests to individuals who truly have no ability. The issue of unconscious motivation is understandably much more murky. Results still reveal misrepresentation, but the question remains whether or not it was willful.

I (Denney, 1999) noted that below-random performance demonstrates ability and suggested that it was the most definitive indication of malingering in the context of litigation. I also raised the concern that because "symptom validity testing was originally designed to detect conversion syndromes, one cannot automatically cry malingering when suppressed scores occur" (p. 16). There is some indication, however, that the "conversion disordered" individuals used in the development of the paradigm (Brady & Lind, 1961; Grosz & Zimmerman, 1965; Theodor & Mandelcorn, 1973; Zimmerman & Grosz, 1966) may have been feigning, but the behavior was considered unconsciously motivated given the *zeitgeist* of the time. Pankratz (1979; Pankratz, Fausti, & Peed, 1975) was one of the first to adapt the strategy to neurocognitive functioning and described below-random results as a possible "smoking gun of intent" (Pankratz & Erickson, 1990, p. 385). Bianchini, Mathias, and Greve (2001) suggested that such suppression was clearly indicative of feigning. Larrabee and colleagues (2007) reviewed the issue again and concluded "the interpretation of a significantly below-chance result as definitive evidence of *intentional* exaggeration of cognitive deficits even in the context of objective pathology has become well established in the neuropsychological literature" (p. 346, original emphasis). Given the age-old difficulty with defining and measuring unconscious processes, the contribution of unconscious motivation to below-random performance will undoubtedly remain unclear. Nonetheless, current literature indicates that the field is largely in agreement regarding the significance of below-random performance from individuals in litigation.

Thankfully, both Boone (2007; Boone & Lu, 2003) and Larrabee (2003a, 2005) have demonstrated that below-random performance on two-alternative, forced-choice testing may not be the only definitive test finding regarding NRB classification. Effort tests are developed to have very low false-positive rates (i.e., their specificity rates are kept high, generally greater than 90%). Because the false-positive rate is at or less than 10%, tests can be used in concert to increase diagnostic certainty substantially. Presuming a minimal amount of correlation between tests, positive findings on multiple tests significantly decrease the possibility of false-positive findings (e.g., $.10 \times .10 \times .10 = .001$). In other words, positive results on three NRB indices suggest a less than 1 in 1,000 chance of a false-positive finding. Although it is clear that some level of correlation occurs between NRB tests (Nelson, Boone, Dueck, Wagener, Lu, & Grills, 2003), these correlations are likely kept to a minimum by incorporation of multiple measures from different cognitive domains or testing paradigms (e.g., verbal vs. visual memory, free-standing measures vs. imbedded indices). Accuracy of specificity rates is limited when tests have not been validated using individuals with severe, and specific, cognitive pathology. In this regard it is important for the clinician to understand how well the proposed pathology of the criminal defendant corresponds to validation samples of the particular NRB index. I believe this is the greatest area of weakness regarding the assessment of malingering at this time.

Assessment of Malingering

The detection of deception is a difficult endeavor under any circumstance (Boyd et al., 2007). However, it is made extremely difficult when neuropsychologists rely solely on behavioral clues (Ekman, 1992). Similarly, limitations of subjective clinical judgment necessitate use of objective measures of invalidity and effort during the assessment of neurocognitive function (Bigler, 1990; Faust & Guilmette, 1990; Faust et al., 1988; Heaton et al., 1978). Consequently, neuropsychologists must incorporate objective measures when assessing malingering of neurocognitive and psychiatric impairment. I first review free-standing measures of NRB, then address NRB indices within common neuropsychological tests and batteries.

Free-Standing Measures of NRB

Each of the following measures of NRB has different characteristics. An important consideration is the apparent face validity of each. For example, some measures are based on general intellectual ability, whereas others are based on visual or verbal memory. Although some of these procedures may be termed "floor effect" tests in that they actually measure only a modest level of ability, they, nonetheless, have face validity to the evaluee. In this regard, subjects who wish to exaggerate deficits in certain neurocognitive functions (consistent with their understanding of what true pathology looks like) will likely not exaggerate in every functional domain. This distinction becomes important when psychologists evaluate the meaningfulness of NRB test results. Furthermore, this distinction may explain why certain types of procedures appear to be more sensitive than others. Individuals attempting to deceive an evaluator will likely not suppress every task they perform, because it may too easily appear disingenuous. They suppress the function they see as most likely to appear legitimate based on their understanding of their proposed pathology, as well as specific forensic demands (e.g., competency to stand trial related abilities). Finally, NRB measures differ in their level of transparency as a malingering detection device, which likely affects their sensitivity as well.

General Intellectual Ability

VALIDITY INDICATOR PROFILE

The Validity Indicator Profile (VIP; Frederick, 2003) is unique for several reasons. It is a measure based on verbal and nonverbal intelligence; it combines two-alternative, forced-choice structure with performance curve analysis due to progressive item difficulty; and it establishes a fourfold classification system of performance (compliant, inconsistent, irrelevant, and suppressed). Developed with items adapted from intelligence tests, it assesses general, intellect-related functions and not memory. The nonverbal test can be administered separately from the verbal test for those who cannot read. It uses six different classification strategies developed with a sample of more than 1,000 clinical and nonclinical subjects. It was cross-validated with an independent sample of 312 individuals in five criterion groups: TBI patients; suspected malingerers; normal subjects; simulators; and random responders. In his own review of the test, Frederick (2002) reported sensitivity/specificity rates of 73.5/85.7 for the nonverbal subtest and 67.3/83.1 for the verbal subtest during cross-validation. Frederick suggested that these rates were underestimates of test performance because of criterion group contamination. The VIP is a breakthrough in attempting to understand the nature of malingering as more than simply a dichotomous occurrence, because it takes into consideration motivation and effort. In addition, there are no other free-standing NRB indices specific to general intellectual function.

As a result of this last consideration, the measure can actually provide an estimate of a person's intellectual range of ability. This estimate can be provided when the subject is motivated to perform well and provides adequate effort; it is also available when subjects are motivated to perform poorly and suppress their scores below random performance. By the nature of the binomial theorem, below-random scores indicate just as much ability as above-random scores (Denney, 1999), which can then estimate intellectual ability in the case of the VIP. This characteristic of the VIP was validated in the development samples by correlating it with the Shipley Institute of Living Scale (Zachary, 1986). I have found this test helpful when criminal defendants are claiming mental retardation and are suppressing their scores below random. In these instances, defendants occasionally suppress their score so low that they actually demonstrate intellectual abilities well above that of mental retardation.

A word of caution is needed regarding the use of the VIP with individuals presenting with possible mental retardation. The VIP was validated on 40 nonlitigating subjects with mental retardation (MR) (Frederick, 2003). Subjects were determined to have IQs between 54 and 75, based on Shipley estimates (Zachary, 1986). Twenty subjects' IQs were below 65. There was a clear relationship between IQ and VIP classification, because those with Compliant classifications had higher IQs than those classified as Inconsistent. Those classified as Inconsistent had higher IQs than those classified as Irrelevant. Thirty-two of the 40 subjects with MR produced results considered invalid on the Nonverbal subtest, and 30 subjects produced invalid results on the Verbal subtest. Most obtained results were classified as Inconsistent and Irrelevant. Only one subject produced a suppressed result, and this occurred on the Verbal subtest. These findings have resulted in the recommendation not to use the VIP with individuals who have a bona fide history of MR. However, the manual indicates that it is appropriate to use the VIP with individuals who present themselves as having MR but have no historical documentation to support that conclusion.

DOT COUNTING TEST

The Dot Counting Test (DCT) was originally developed by André Rey (1941) as a method to identify individuals attempting to feign TBI. Frederick (2002) reviewed the history of the DCT and presented Rey's original ideas regarding the procedure. The procedure incorporates 12 cards with dots arranged in various patterns. The first set comprises six cards with dots arranged in random patterns, whereas the latter set has six cards with dots arranged in clear groupings. Individuals are asked to count the dots as quickly as they can as each card is presented. Classically, performance times were compared between grouped and ungrouped cards, on the assumption that disingenuous subjects inadvertently perform too slowly when counting grouped dots compared to ungrouped dots (Frederick, 2002; Frederick, Sarfaty, Johnston, & Powel, 1994). There have also been other variations in scoring methods, typically dealing with the number of errors in counting (Binks, Gouvier, & Waters, 1997; Lee et al., 2000; Paul, Franzen, Cohen, & Fremouw, 1992).

There have been numerous studies of the DCT with simulators, clinical, and forensic samples. Vickery, Berry, Inman, Harris, and Orey (2001) performed a meta-analysis of six studies and concluded the DCT was not particularly effective at differentiating between simulators and honest responders. Boone, Lu, and Herzberg (2002) developed a formula (mean ungrouped time + mean grouped time + number of errors) and found it to be a sensitive index for detecting individuals with noncredible symptoms from different settings. This formula was then formalized as the Effort Index, or "E-score," in the later standardized test publication (Boone et al., 2002). The studies on which this published version is based include a number of normal-effort clinical groups including persons with depression, schizophrenia, head injury, stroke, learning disability, and mild dementia. With the various groups come differing cutoff scores to maintain specificity rates near .90. With these groups, the DCT has an average sensitivity of 78.8%. It is currently not recommended for differentiating feigning persons from those with moderate dementia (Boone et al., 2002) or MR (Marshall & Happe, 2001; Victor & Boone, 2007).

Short-Term Memory

REY 15-ITEM MEMORY TEST

The Rey 15-Item Memory Test (FIT; Rey, 1958) is the most widely known of Rey's malingering tests. It is a simple memory procedure that takes little time to complete and has received a great deal of research attention. Frederick (2002) reviewed the procedure and found sensitivity rates to vary from 40 to 89% depending on the cutoff, with specificity generally placing in the middle to upper .90's. There appeared to be a difference in test performance between civil litigants and criminal defendants. Boone, Salazar, Lu, Warner-Chacon, and Razani (2002) reviewed the test and found sensitivities ranging from 7 to 72% for volunteer simulators, and from 5 to 72% for clinical samples of patients in litigation and those suspected of malingering. Reznek (2005) performed a meta-analysis and found the FIT to have an overall sensitivity of only 36% using the standard cutoff. Nitch and Glassmire (2007) reviewed studies pertaining to the FIT and found sensitivities between 5 and 86% (M = 41.13) using the same standard cutoff. The mean sensitivity for suspected malingerers in clinical settings was 44.2%. Only three of those studies dealt with criminal suspected malingerers (Frederick et al., 1994; King, 1992; Simon, 1994). Sensitivities among these criminals ranged from 43 to 86%, with a mean of 66.33%. It appears the test may be more sensitive among criminal defendants.

Boone and her colleagues (2002) developed a recognition test to add to the FIT, which raised the sensitivity to 71% over the rather modest sensitivity of 47% in a known groups, civil litigation design. This change in the procedure adds little additional time to a test that is already quite time-efficient. There have been no studies of the procedure's effectiveness in the criminal setting using this additional modification. Vallabhajosula and van Gorp (2001) suggested that the procedure, as originally developed, would not meet *Daubert* (*Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 1993) court admissibility standards because of low sensitivity, but Frederick (2002) suggested that it is a reasonable procedure to use so long as it is not used in isolation. Although sensitivity has been somewhat low, the procedure has been generally considered quite good.

Reznek (2005) performed a meta-analysis and found the FIT to have an overall specificity of 85% using the standard cutoff. A 90% specificity rate was obtained if cases with MR were removed. From a conceptual perspective, however, pooling various neuropathologies with psychiatric referrals and normal controls to obtain a global specificity rate is less than fully meaningful, because the procedure clearly has higher specificity rates with mild head injuries compared to severe injuries and dementing illnesses. Nitch and Glassmire (2007) reviewed the procedure's specificity using the standard cutoff over a wide variety of studies and patient groups. As expected, there was such striking variability between groups that the authors recommended caution when using the test to evaluate cases of strongly suspected genuine memory disorder, dementia, stroke, low intelligence, poor education, and psychosis.

FORCED-CHOICE DIGIT RECOGNITION TESTS

The Portland Digit Recognition Test (PDRT; Binder, 1990; Binder & Willis, 1991), the Computerized Assessment of Response Bias (CARB; Allen, Conder, Green, & Cox, 1997), and the Victoria Symptom Validity Test (VSVT; Slick, Hopp, Strauss, & Thompson, 1997) are all forced-choice, digit recognition tests patterned after the Hiscock Digit Memory Test (Hiscock & Hiscock, 1989). In each of these procedures individuals view a five-digit number and are required to recognize the correct number after a delay. There are three distinct series of presentations, each with longer delay times. The increased delays included in each series make the items appear more difficult. The PDRT incorporates a counting backward distraction task during the delays. The PDRT requires about 40 minutes to administer, although an abbreviated form has been developed for individuals who appear to be performing well (Binder, 1993). Binder (2002) reviewed the procedure and noted that sensitivity rates varied from 39 to 77% depending on the type of subject, whereas specificity rates held constant at 100%. In a meta-analytic review of six studies using the PDRT, Vickery and colleagues (2001) found the measure to have intermediate sensitivity (44%) but excellent specificity (97.3%). Greve and Bianchini (2006a) evaluated the effectiveness of the PDRT with a known-groups design that comprised 262 TBI referrals classified on Slick and colleagues (1999) criteria for MND. They concluded that the original cutoff scores for the PDRT were too conservative, because the sensitivity was only .20–.50. Using a lower total score cutoff, they were able to demonstrate a sensitivity rate of 70% while holding to a 95% specificity rate.

The CARB is similar to the PDRT, but it is computer administered. It has three forms, the 111-item original (Conder, Allen, & Cox, 1992), the 72-item CARB-97 (Allen et al., 1997), and a shorter length variation (Green & Iverson, 2001). It can also be set to provide correct–incorrect feedback via color and tone for each trial and even to discontinue after a perfect performance on the first trial. Subjects are asked to count backward from 20 during the delay periods. The CARB uses DOS-based software and has a poorly written manual (Wynkoop & Denney, 2001). The test publisher has a website that purports to provide a Windows version of the test and ongoing software updates; however, this site repeatedly did not function. Allen, Iverson, and Green (2002) reviewed the development and characteristics of the CARB but gave no specific sensitivity or specificity rates. The great benefit of the CARB, however, is that it includes norm references that allow the examiner to compare the subject's performance to known groups of patients with neurological, amnestic, and severe brain injury. Oftentimes, this comparison by itself provides substantial indication of poor subject effort due to presence of clear inconsistency between test results and activities of daily living. Studies supporting the validity of the CARB are not clearly understood beyond the striking finding that larger percentages of litigating persons with disability fail the procedure compared to subjects with more severe brain injuries. Given the research with individuals with moderate and severe TBI, the test appears rather insensitive to brain damage (Conder, Allen, & Cox, 1992). The mean total CARB score was 96.8% correct with moderate and severe brain injury. When Allen and Green (1999) administered the CARB to 56 patients with severe TBI, the total CARB score correct mean was 98.3%. Green and Allen (1999) also reported results for 40 patients with mixed neurological disease, with a mean total CARB score of 97.2%. Overall, these studies indicate that the CARB is not sensitive to brain damage when subjects are otherwise testable.

There are no specific reports of research directed at the sensitivity of the CARB. Gervais, Rohling, Green, and Ford (2004) compared the CARB with the WMT and TOMM using 519 non-head-injured personal injury and disability claimants. They found that twice as many individuals failed the WMT compared to the TOMM, with CARB results falling in between. They concluded that the CARB was less sensitive than the WMT, but more sensitive than the TOMM. Nitch and Glassmire (2007), in their review, conclude that the CARB appears to be "limited by its negative predictive power; consequently, false negatives may be common" (p. 64).

The VSVT (Slick et al., 1997) is also computer administered, but unlike the CARB it is published by a well-known test publisher and runs well in Windows. The VSVT program presents a total of 48 items over three trials of 16 items. The three trials have increasing delay times before the subject is required to recognize the presented five-digit number from a foil. Also unlike the PDRT or CARB, subjects do not have distraction activities during the delay periods. Uniquely, the VSVT incorporates what appears to be easy items and difficult items by having foils completely different from the stimuli for the easy items and foils only minimally different from the stimuli. While giving the appearance of increased difficulty to the examinee, research demonstrates no actual increase in difficulty between easy and difficult items with non-compensation-seeking neurological patients (Slick et al., 1997). Classification of test results is based solely on the binomial theorem by labeling performance above random (valid), random (questionable), or below random (invalid) for easy, difficult, and total scores. Finally, the program allows for printing of graphs with comparison groups.

There are few findings regarding test sensitivity with the VSVT, because most studies have used a differential prevalence design. Tan, Slick, Strauss, and Hultsch (2002) performed a simulation study with undergraduates. VSVT difficult item scores accurately classified 96% of malingerers. They found that combining easy and difficult items resulted in 100% classification of malingerers and controls.

The original development studies did not include individuals with severe brain injury, because the neurological group comprised non-compensation-seeking patients with seizure disorders. Slick and colleagues (2003) presented VSVT data on six non-compensation-seeking individuals with severe memory disorder (anterior cerebral artery aneurysm, Korsakoff syndrome, anoxia, and epilepsy). None of these subjects' performance was below 22 out of 24 correct for the difficult items, and all obtained 100% correct for the easy items. These results are clearly limited by small sample size. Macciocchi, Seel, Alderson, and Godsall (2006) administered the VSVT to 71 individuals with acute, severe brain injury and found that nearly all subjects (96%) scored better than 44 out of 48 correct on the combined easy and difficult items. They noted that only those individuals with severe visual–perceptual and verbal fluency deficits performed poorly. They suggested that the computer classification system was too conservative and provided alternative cutoffs. These studies suggest that the VSVT is quite resistant to the effects of brain damage in general.

In contrast to these results, Loring, Larrabee, Lee, and Meador (2007) presented VSVT data for various neurological cases evaluated in a medical setting (50 dementia, 38 cerebrovascular, 19 multiple sclerosis, and 27 mixed pathology). There were unacceptably high rates of test failure (i.e., less than .90 specificity) using the criteria of less than 18 out of 24 correct for difficult items among all their clinical groups (22% dementia, 18% clinical TBI, 15% memory complaints, 16% cerebrovascular, 11% multiple sclerosis, and 11% mixed neurological). However, they did find significantly poorer performances in the compensation-seeking cases. Grote and colleagues (2000) obtained a 93.3% specificity rate using 90% correct on difficult items with 30 non-compensation-seeking patients with epilepsy. Loring, Lee, and Meador (2005) found a 71.6% specificity rate among 120 patients with epilepsy using the same cutoff and a 75% specificity rate using the binomial theorem–based cutoff of "questionable validity." It was unclear how many of the Loring and colleagues cases may have been seeking compensation, because theirs was a retrospective study. The VSVT has been studied in Spain (Vilar-López et al., 2007) and appears to have similar effectiveness in a non-English-speaking population. Clearly, there is variability in VSVT findings with clinical groups. Additional known-group studies are needed to establish optimal cutoffs that maintain specificity rates of .90 or greater.

TEST OF MEMORY MALINGERING

The Test of Memory Malingering (TOMM; Tombaugh, 1996) is also a two-alternative, forced-choice test of memory, but it uses line drawings rather than digits. There are three trials of the TOMM. Trial 1 is a learning trial in which all pictures are presented and the subject immediately completes a recognition task. The pictures are then presented again in Trial 2, with the subject again performing an immediate recognition task. There is then a 15-minute delay recognition task (Retention). The TOMM manual suggests that evaluators may wish to forgo the Retention trial if subjects have performed well on Trial 2; however, Greve and Bianchini (2006b) found that doing so resulted in a 3% decrease in test sensitivity.

Initial development simulation studies with undergraduates revealed 84 and 88% detection rates for simulators, with a 100% specificity rate using the recommended cutoffs (Tombaugh, 1996, 2002). A follow-up study using patients with TBI in litigation resulted in 77% correct classification rate, and specificity dropped to still a respectable 90% when including severely impaired neurological patients (Tombaugh, 1996, 2002). Teichner and Wagner (2004) evaluated the performance of the TOMM in elderly patients with cognitive impairment and dementia. Patients with cognitive impairments did not differ significantly from cognitively intact patients (92.7 vs. 100% normal performance, respectively). The false-positive rate for patients with dementia, however, was considered unacceptably high, with false-positive rates of 76% using the standard cutoff for Trial 2. There appears to be little effect from depression and anxiety for community-dwelling older adults (Ashendorf, Constantinou, & McCaffrey, 2004; Yanez, Fremouw, Tennant, Strunk, & Coker, 2006) nor from moderate to severe pain (Etherton, Bianchini, Greve, & Ciota, 2005).

Delain, Stafford, and Ben-Porath (2003) reviewed the performance of the TOMM in a sample of pretrial criminal defendants, and results appeared to support the test's validity in that setting. In a differential prevalence design, Weinborn and colleagues (2003) found that criminal defendants referred for pretrial mental health evaluations were more likely to demonstrate below-cutoff performances on the TOMM than were criminal defendants adjudicated not guilty by reason of insanity (and presumably wanting to perform their best for possible release). Although a number of studies demonstrate the utility of the TOMM in the civil forensic arena, these results also suggest that it is a valid indicator of poor effort among criminal defendants. The TOMM appears to perform as well in at least some Spanish-speaking populations as it does in English-speaking North American populations (Vilar-López et al., 2007).

WORD MEMORY TEST

The WMT (Green, 2003; Green, Allen, & Astner, 1996) is unique in the area of malingering tests, in that it is a forced-choice, two-alternative procedure that also includes legitimate memory assessment for words and word pairs. This computer-administered test automatically calculates *z*-score comparisons to over 2,800 individuals of various ages and diagnoses that most closely correspond to characteristics of the subject's performance. Comparison groups can also be selected by the administrator. It can be computer-administered in nine languages. Green and colleagues (2002) noted the WMT demonstrates overall sensitivity rates of about 97% and specificity rates of 100% for simulation studies. Tan and colleagues (2002), in an undergraduate simulation study, found the WMT to have 92.6% sensitivity and 100% specificity at recommended cutoffs.

There have been numerous studies of the WMT with subjects with significant neuropathology; however, the findings are difficult to interpret (Green, Iverson, & Allen, 1999; Green et al., 2002; Green, Rohling, Lees-Haley, & Allen, 2001). Results of these studies seem to suggest that more individuals with mild brain injuries fail the WMT than do those with severe brain injuries. Furthermore, those with putatively mild injuries who failed WMT demonstrated significantly worse neuropsychological test scores overall than did those who had severe injuries but passed the WMT. These results strongly suggest that the WMT is not nearly as sensitive to brain injury as it is some other construct, presumably poor effort. Bowden, Shores, and Mathias (2006) sought to replicate the Green and colleagues (2001) finding that WMT failure rates appear higher in mild compared to more severe brain injuries among litigating subjects. They reported data from 100 consecutive subjects, ranging from age 6 to 74 years, who were involved in litigation related to TBI. The researchers found no indication of improvement in WMT (Immediate Recognition only) performance as injury severity increased. They found a significant, but likely inconsequential, effect of WMT Immediate Recognition performance on the outcome measure of delayed memory using 86 subjects ($p < .001$, $\eta^2 = 0.18$). It is difficult to determine the significance of Bowden and colleagues' results, because their research incorporated a much smaller *N* than did the Green and colleagues data; they did not identify how many cases were pediatric; and they only administered the Immediate Recognition portion of the WMT.

Many of the WMT studies have also included other malingering tests, such as the CARB and TOMM. Gervais and colleagues (2004) reviewed 519 compensation-seeking cases who were given the TOMM, the CARB, and the WMT. They found failure rates of 11% on the TOMM, 17% on the CARB, and 32% on the WMT Primary Effort subtests. These findings suggest that the WMT is a more sensitive measure of poor effort and malingering than the CARB or the TOMM.

Gorissen, Sanz, and Schmand (2005), using the Dutch and Spanish oral versions of the WMT with patients with schizophrenia, nonpsychotic psychiatric patients, neurological controls, and healthy controls as part of a larger neuropsychological assessment, confirmed that the WMT does not correlate with more customary measures of memory, such as delayed verbal and figural memory, regardless of which subject group was considered. Nonetheless, they found that 72% of the schizophrenia group, 25% of the psychiatric controls, 10% of the neurological controls, and none of the healthy controls failed the WMT. When they divided groups based on WMT performance, they found that those in the poor effort schizophrenia group performed more poorly on other neuropsychological tests than did the neurological patients. Additionally, poor WMT scores within the schizophrenia group correlated significantly with the negative symptoms scale of the Positive and Negative Syndrome Scale (PANSS; Spearman's $\rho = -0.42$; Kay, Opler, & Fiszbein, 1986), yet did not correlate with tests of executive function, such as the Trail Making Test–Part B, the Stroop Color–Word test, or Verbal Fluency test. The authors considered results to possibly reflect an amotivation syndrome. Other studies have demonstrated that incentives may increase cognitive performance among those with schizophrenia (Schmand et al., 1994). Results are difficult to interpret, because there was no apparent motivation for patients in this study to perform well. We are left not knowing how much the poorer performances simply reflected subjects' lack of desire to perform well.

Although the WMT initially received a cool review due to difficult computer software and lack of peer-reviewed and published research (Wynkoop & Denney, 2001), it has now become a widely researched tool with a strong empirical basis (Hartman, 2002; Wynkoop & Denney, 2005). With the new Windows version, the test is easy to use, includes numerous clinical and nonclinical comparison groups, and can be performed in multiple languages. Research with the instrument in different languages is starting to come out, and results suggest it performs just as well in German, Turkish, and Russian (Brockhaus & Merten, 2004; Brockhaus, Peker, & Fritze, 2003; Tydecks, Merten, & Gubay, 2006).

MEDICAL SYMPTOM VALIDITY TEST

The MSVT (Green, 2004) is, in essence, a simpler version of the WMT. Despite the term *medical* in the title, the test uses a verbal memory paradigm like the WMT, rather than any particular assessment of medical symptoms. The MSVT uses 10 word pairs rather the WMT's 20 word pairs, and the word pairs have a much stronger semantic relationship. Consequently, it appears to be a much easier test. As with the WMT, words are presented via the computer in word pairs. Subjects then complete an initial recognition test for each word. There is a 10-minute delay and the recognition test is presented again. There are then paired associate and free recall tasks. The entire test takes about 20 minutes to administer, including the 10-minute delay. The ease of the task is demonstrated by the fact that English-speaking adults and children who could not speak French scored nearly perfectly on the validity scales portion of the test, even when tested with French words (Richman et al., 2006). The computer program includes a feature that easily compares the subject's performance to a large number of comparison groups, including five simulator groups, healthy children and adults, child clinical groups (fetal alcohol syndrome, learning disability, conduct disorder, MR, and attention deficit disorder), and adult clinical groups (mild through severe TBI, neurological with impaired memory, Social Security Disability claimants, chronic pain, major depression, anxiety, soft-tissue insurance disability claimants, schizophrenia and bipolar disorder, and early and advanced dementia). In addition to a strict cutoff for the three validity scales in the measure, a dementia profile is designed to identify those who fail the MSVT, but do so in a manner consistent with genuine dementia (Green, 2004).

Merten, Green, Henry, Blaskewitz, and Brockhaus (2005) performed an analog study of the oral form of the MSVT in Germany. They compared performances between 18 simulators and 18 healthy adult controls. The simulators were warned that validity measures would be included in the test battery. The MSVT results were nearly perfect. Delayed Recognition and Consistency scores demonstrated 100% classification of both simulators and health adults. The Immediate Recognition classified all of the simulators, but one healthy subject's score fell on the cutoff (94.4% specificity). These specificity rates have little meaning, because the study did not include cases with known clinical conditions.

Teichner, Waid, and Buddin (2005) presented results of the computer-administered English-language MSVT with 294 clinical and forensic adult and pediatric subjects. They used the TOMM and WMT as indices within the Slick and colleagues (1999) classification for MND. They did not explain this classification further. All 102 children were considered to have provided good effort. Seventeen of the 192 adults were considered to have provided poor effort (8.85%). All of those poor effort cases were identified by the MSVT (100% sensitivity). The strength of this study lies in the fact that it included subjects with bona fide neuropathological conditions (e.g., TBI, stroke, carbon monoxide exposure, MR, mild dementia). Of the children, 7.8% were misclassified as giving poor effort (92.2 specificity). Of the adults, including those with dementia, 12% were misclassified as having poor effort (88% specificity). Extracting the dementia cases (N = 15) from the data reduced the misclassification to 9.4% (90.6% specificity). It is difficult to interpret this data without knowing the details of how the TOMM and WMT were used in classifying the subjects.

Howe, Anderson, Kaufman, Sachs, and Loring (2007) evaluated the performance characteristics of the MSVT among clinical referred memory disorders patients. They included data for 63 subjects, 11 of whom were considered to have a disability-related incentive to possibly perform poorly. Forty-five precent of the disability group failed one or more of the MSVT Validity scales. Two of them had clear indication of invalid responding, and three met the MSVT's valid dementia profile algorithm. Twenty-two percent of the subjects without disability failed one or more of the MSVT Validity scales. After considering the dementia profile algorithm, however, they found a false-positive rate of only 4.76% for their entire sample. Two of these three subjects were considered to have advanced dementia. The third was considered to have early dementia. These data indicate the MSVT is a powerful tool when differentiating valid from invalid performances among memory disorder referrals, including dementia, with a 95.24% specificity rate.

NONVERBAL MEDICAL SYMPTOM VALIDITY TEST

Green (2006) has introduced the preliminary research edition of the Nonverbal Medical Symptom Validity Test (NV-MSVT), a computer-administered test that is very similar in design to the MSVT; however, it uses visual stimuli rather than word pairs. It also takes about 20 minutes to complete and includes a 10-minute delay. In contrast to the MSVT, it incorporates variations during the delayed recognition task that, when compared to the paired associate task, appear to make it particularly capable in differentiating genuine from feigned dementia. At the time of this writing one can compare a subject's performance to a number of normative and clinical comparison groups (including children with fetal alcohol disorder, adults passing and failing WMT and MSVT, good-effort adult volunteers, four groups of adults with dementia, adult dementia simulators, adult male prisoners with end-stage renal disease, and four groups of children stratified by age from 7 to 18 years). A number of clinical researchers are currently collecting data on the NV-MSVT, so comparison groups will likely increase in number and variety.

Green reported results from 107 consecutive compensation-seeking subjects who were also administered other validity measures (Reliable Digit Span [RDS], TOMM, WMT, & MSVT). Failure rates were 32% for the WMT, 25% on MSVT, 20% on RDS, 20% on NV-MSVT, and 8% on TOMM. Preliminary results suggest a strikingly different pattern between simulators and those with genuine dementing conditions. Use of singular cutoffs does not appear to work nearly as well as the overall test result profile. Using an algorithm that compared recognition, paired associate, and free recall, Green reports a 96% correct classification for 128 subjects (40 good-effort adults, 40 adult dementia simulators, 19 good-effort children, and 29 individuals with dementia). When parsing out the 40 simulators and 29 dementia cases from the data provided in the handout and creating a hybrid simulator, known-group design, the NV-MSVT obtained a 95% sensitivity, a 90.6% specificity, and correct classification of 93%. These preliminary results are quite promising, particularly when dealing with the often difficult diagnosis of dementia. The manual should be published by the time this volume goes to press.

Embedded Indices of Neurocognitive NRB

Aside from tests and techniques designed specifically to assess neurocognitive malingering, many NRB indices are incorporated within widely used psychological and neuropsychological measures. These strategies are beneficial because they take little additional time to administer and the connection between poor performance on the validity scale and poor performance on the genuine test is clear. In addition, they can be calculated on testing protocols from evaluations done in the past. The negative aspect of these strategies is that their specificity will likely not be as strong as that of freestanding NRB measures, because they comprise tasks designed to measure actual effects of brain pathology. Vallabhajosula and van Gorp (2001) suggested that the best malingering detection strategies will be measures sensitive to feigning but not to genuine impairment. Nevertheless, the trend of embedding NRB detection methods within already established neuropsychological measures appears to be a reasonable pursuit. Examples of such malingering detection strategies based on established clinical measures include the Warrington Recognition Memory Test (for a review, see Millis, 2002), atypical pattern analysis on the Wechsler Scales, Wechsler Memory Scale—Revised, and Halstead–Reitan Battery (for a review of each, see Mittenberg et al., 2002), RDS (Greiffenstein, Baker, & Gola, 1994; Larrabee, 2003a; Meyers & Volbrecht, 1998; for a review, see Suhr & Barrash, 2007), Rarely Missed Index of the Wechsler Memory Scale–III (Killgore & DellaPietra, 2000; Suhr & Barrash, 2007), California Verbal Learning Test (Millis, Putnam, Adams, & Ricker, 1995; Slick, Iverson, & Green, 2000), Wisconsin Card Sorting Test (for a review of various methods, see Greve & Bianchini, 2007; Greve, Bianchini, Mathias, Houston, & Crouch, 2002), Finger Tapping (Heaton et al., 1978; Larrabee, 2003a), Benton Judgment of Line Orientation (Iverson, 2001; Meyers, Galinsky, & Volbrecht, 1999), Category Test (for reviews, see Greve & Bianchini, 2007; Sweet & King, 2002), test–retest changes on the Halstead–Reitan Battery (for review, see Reitan & Wolfson, 2002), and Benton Visual Form Discrimination (Larrabee, 2003a). Readers are referred to each article to learn specifics of computing these indices, because reprinting them here could compromise test security.

Using Multiple NRB Indices

Larrabee (2003a) pointed out that "assessment of effort in medicolegal settings must be multi-variate" (p. 422). Meyers and Volbrecht (2003) demonstrated that use of multiple *imbedded* NRB indicators within a standard neuropsychology battery can identify invalid performance. Meyers and Volbrecht found 83% sensitivity and 100% specificity in identifying NRB using a positive cutoff rule of 2 or more among a mixed group of clinical cases and analog simulators. There were no false positives using this method. Larrabee (2003a) found that using multiple NRB indicators and a positive cutoff of 2 or more resulted in an overall sensitivity of 87.8% and specificity of 94.4% for the combined samples of litigating and nonlitigating closed head injury evaluees classified based on Slick and colleagues (1999) criteria. He also demonstrates that use of multiple indicators decreases the chance of false-positive identification errors. These sensitivity rates rival, and even surpass, those of many free-standing measures of NRB. Using the aforementioned imbedded indices facilitates the identification of NRB in this regard, with no additional test administration time. Incorporating one or more free-standing indices of NRB would likely increase the sensitivity of this multivariate method without compromising specificity.

Self-Report Measures of Psychiatric Disturbance

Although this chapter is focused on exaggeration of neurocognitive dysfunction as it relates to criminal forensic evaluations, the neuropsychologist performing work in this setting will often face cases in which feigning of psychiatric impairment is also an issue. There is a large amount of literature regarding the detection of feigned psychosis; however, the subject is beyond the scope of this chapter. For a recent review of self-report measures related to assessment of feigned psychiatric impairment, see Berry and Schipper (2007).

Complaints of Remote Memory Loss

Significant retrograde amnesia is rare in the absence of substantial brain damage and typically raises concerns about psychogenic etiology (Parkin, 1996; Ross, 2000). Claims of remote memory loss, when they relate to alleged criminal activity, however, are not unusual (Schacter, 1986). It appears to occur more commonly with acts of violence. Reported amnesia in relation to homicide charges are estimated to range from 23 to 65%

(Bradford & Smith, 1979; Guttmacher, 1955; Evans, 2006; Leitch, 1948; Parwatikar, Holcomb, & Menninger, 1985). Hopwood and Snell (1933) reviewed 100 criminal cases and found that 90% of the amnestic claims pertained to murder or attempted murder charges. In a more recent review, only 8% of 120 cases of nonhomicide violent crimes included claimed amnesia, and there were no claims of amnesia among 47 individuals charged with nonviolent crimes (Taylor & Kopelman, 1984).

Most researchers suggest that a substantial portion of such claims are feigned (Adatto, 1949; Bradford & Smith, 1979; Hopwood & Snell, 1933; Lynch & Bradford, 1980; O'Connell, 1960; Parwatikar et al., 1985; Power, 1977; Price & Terhune, 1919). Schacter (1986) provided this viewpoint:

> In the large majority of criminal cases that involve amnesia, the loss of memory either has a functional origin or concerns only a single critical event. I have found no cases in the literature in which a patient afflicted with chronic organic amnesia has come before the courts on a serious criminal matter that is related to his or her memory disorder. Organic factors may play a role when concussion, alcohol intoxication, or epileptic seizure occurs during a crime, with subsequent limited amnesia for the crime itself, but in these cases memory problems typically do not exist prior to the crime. (p. 287)

Occasionally, criminal defendants experience a neurological condition either severe enough or close enough in time to the crime to hinder recall of events around the time of the alleged offense (Miller, 2003; *Wilson v. United States*, 1968; Wynkoop & Denney, 1999). It is also possible that individuals carrying out criminal activity while intoxicated may subsequently not recall these important events. Other instances could include individuals who experience a TBI during the crime or arrest, or experience a neurological insult, such as cerebrovascular stroke or hemorrhage, after the arrest but before legal proceedings are concluded (Denney & Wynkoop, 2000). Under such circumstances, it is not unreasonable to find loss of memory for events preceding the arrest, including the offense behavior. A criminal defendant's ability to recall events sufficiently to reconstruct his or her activities for the period of time around the offense may be an important aspect of his or her competency to proceed.

Symptom Validity Testing for Remote Memory

Identifying feigned memory loss for a specific period of time or for a specific event is not easy. Most methods used to evaluate amnesia (see Rubinsky & Brandt, 1986; Schacter, 1986; Wiggins & Brandt, 1988) do not avail themselves of the detection of specific past criminal events. One exception is symptom validity testing (SVT), which appears to work well in assessing claims of remote memory loss.

Originally designed to assess psychogenic sensory complaints as mentioned above, SVT was modified for learning and retention complaints (Grosz & Zimmerman, 1965; Haughton, Lewsley, Wilson, & Williams, 1979; Pankratz, 1979; Pankratz et al., 1975; Theodor & Mandelcorn, 1973). Frederick and Carter (1993; Frederick, Carter, & Powel, 1995) adapted this technique to assess memory for events surrounding an alleged offense for which a criminal defendant claimed amnesia. They developed two-alternative, forced-choice questions for events presented in the criminal investigative records for which the defendant claimed no recollection. Upon administration, the defendant performed statistically below expectations for an individual who had no memory for those events. They concluded the man was feigning his amnesia.

The procedure is based on the binomial theorem, which purports that when two possibilities of equal probability exist, results will fall around the mean in an expected bell-shaped curve (Siegel, 1956). Similarly, when an individual with no ability/knowledge is asked a number of questions with only two possible answers of approximately equal probability, results should fall in a random range. Knowledge is demonstrated to a particular level of statistical certainty when results fall outside the random range. Customarily, individuals with knowledge of events in question score well above the random range, therefore demonstrating their knowledge (see Marcopulos, Morgan, & Denney, Chapter 6, this volume, for an example of this occurrence). Likewise, results falling below the random range also demonstrate knowledge, albeit in the opposite direction.

When applied in a criminal context, questions are derived from investigative materials, medical records, or from interviews of witnesses, family members, or law enforcement personnel. Oftentimes, the facts of the case are well described in the indictment and supportive information. This information is combed to create questions about events and facts that the defendant should have known or experienced during the period of claimed amnesia. Enough detail is required to generate an ample number of questions (preferably more than 24); an increased number of items increases sensitivity and overall accuracy. In addition, the information on which the questions are based must be salient enough and be created in such a manner that an individual without significant memory loss would likely have remembered the information. Questions should be developed to include the correct answer and an equally plausible alternative. Questions should also be worded in such a manner that responses avoid direct admissions of guilt (e.g., "investigative records allege . . . " or "the prosecution claims . . . "). The procedure is designed to identify false claims of memory loss rather than to identify guilt.

Frederick and colleagues (1995) noted the difficulty in creating reasonably plausible alternative answers. Occasionally, the difficulty is in creating equally *implausible* alternative answers. For example, I (Denney, 1996) presented a case in which a male defendant dressed as a woman to rob a bank. Although developing questions can be difficult, research has demonstrated that unequal probability answers actually increase the conservative nature of the procedure (Frederick & Denney, 1998). As in the case of the bank robber, it was important that alternative answers not systematically present more likely possibilities. Biasing the test in this manner would inappropriately increase the possibility that a truly amnestic individual could select wrong answers more often than would occur by chance. It is helpful, therefore, to have a colleague review the questions before administration.

Defendants are instructed that, because of their memory concerns, they will be tested regarding their memory for those specific events in order to understand their memory problem more clearly. Each item is presented after the question of whether they remember this information or not. Items for which they claim recall of the information or "reason out" the solution are discarded, because the test is designed to measure their lack of recollection (Denney, 1996). Items not recalled or invalidated through deductive reason are administered to defendants with the instruction to choose the correct answer or simply guess to the best of their ability if they cannot remember. Subjects are told whether they are correct or incorrect, and the correct answer is noted. Often, the task is constructed so that succeeding queries are more specific variants of the preceding question (e.g., "Investigative records allege you did what, rob a bank or perform a drug deal? No, it alleges you robbed a bank. Was it First Interstate Bank or Seattle First Bank? Was it on 2nd Avenue or 4th Avenue?" etc.). Correct answers are totaled and applied to the following formula from Siegel (1956):

$$z = [(x \pm 0.5) - NP] / \sqrt{NPQ}$$

where z is the test statistic, x is the number of correct responses, N is the number of questions administered, P is the probability of a correct discrimination given no true ability (0.5); and Q represents $1 - P$ (probability of an incorrect discrimination). The correction (adding 0.5 when $x <$ NP; subtracting 0.5 when $x >$ NP) is made to correct for continuity as the binomial distribution involves discrete variables. A one-tail test is used to identify the exact probability using the Unit Normal Table (z table). The one-tail test is considered appropriate given the intent to identify suppressed performance

(Larrabee, 1992; Siegel, 1956). A z-score of –1.65 is significant at p = .05 and –2.33 is significant at p = .01.

Denney (1996) and Frederick and Denney (1998) demonstrated that individuals with no knowledge of events in question perform predominantly within the random range, with scores clustering around the mean (50%). Results demonstrated that the procedure was actually slightly more conservative than that spelled out by the binomial distribution. Furthermore, Frederick and Denney performed computer simulations in which P and Q were progressively not equal to 0.5, and demonstrated that the test statistic performed even more conservatively as response options moved away from 0.5 probability. The increased variability led to a decrease in sensitivity, thereby lessening the likelihood of labeling a true amnestic as a malingerer. While the SVT procedure can be time-intensive in terms of acquiring investigative material and developing questions, it has proven itself as an effective tool in identifying feigned claims of remote memory loss.

Malingering and the Criminal Courts

By definition, the term *malingering* indicates intentional misrepresentation for secondary gain. In the correctional environment, this may include feigning mental or physical illness to gain access to drugs, female staff, a safer environment, or a situation in which escape may be possible. In the truly forensic setting (that pertaining to the judicial system) this means feigning or exaggerating disability to delay or avoid prosecution, obviate criminal responsibility, or obtain a lesser sentence. Most commonly, attempts to feign cognitive disability arise in relation to competency to proceed. Although courts have been reticent to allow expert testimony regarding an individual's veracity in personal injury and civil tort actions (*Commonwealth v. Zamarripa*, 1988; *Nicholson v. American National Insurance*, 1998), these instances typically involve the expert providing testimony before a jury. Judges appear to have less concern regarding this issue when it is addressed in a bench proceeding. In fact, it is clear that courts take attempts to circumvent the criminal judicial process quite seriously.

In *United States v. Greer* (1998), the 5th Circuit Court of Appeals affirmed a decision that underscored the seriousness of criminal defendants feigning mental illness. Charles Greer was indicted for kidnapping and various firearms violations. He had a long criminal history and had been committed several times to inpatient psychiatric facilities. He had been found incompetent to proceed on previous occasions and was initially found incompetent to proceed in relation to the state aspect of these charges. He then underwent an inpatient competency evaluation performed by Richard

Frederick, PhD. Dr. Frederick later testified that Greer was not only competent to stand trial but that he was also feigning psychotic illness. After hearing testimony from a defense expert, the court found defendant Greer competent to stand trial.

Greer's bizarre behavior during trial preparation prompted his attorney to file another motion to determine competency. He was examined by a psychiatrist from the Texas Department of Criminal Justice, who opined that he was incompetent. The government did not contest the opinion, and Greer was found incompetent. He was then referred for inpatient mental health treatment to restore his competency. After 2 months of observation, mental health staff were unable to find any active psychotic process or serious mental disease. Mary Alice Conroy, PhD, later testified that in her view he had been malingering. The court found that Greer was feigning mental illness and that he was competent to stand trial. However, on the first day of trial, he took his clothes off and attempted to flush them down the holding cell toilet. He also spit up between 10 and 16 half-dollar-size splotches of blood and was taken to a local hospital. The jail's director of infirmary services testified that an abrasion found in Greer's mouth had caused the bleeding and that such abrasions were commonly caused by self-inflicted scratches. She also testified that it appeared Greer was gagging himself rather than vomiting blood. Without the jury present, the U.S. District Judge told Greer that he believed Greer was a malingerer. He also told Greer that if he acted up or tried to disrupt the trial while in the courtroom, he would be removed from the courtroom and the trial would proceed in his absence. Greer exhibited additional disruptive behaviors during the trial and was subsequently removed. The jury convicted Greer in his absence. At sentencing, the court granted the government's argument that Greer's sentence be enhanced for obstructing justice, since he had feigned mental illness prior to and during trial. He received a 210-month sentence with the enhancement, whereas he would have received a 185-month sentence without it. The court's decision to enhance the man's sentence due to malingering was affirmed on appeal.

Future Directions

Research regarding free-standing and embedded neurocognitive malingering measures has increased in recent years. Although there is a tremendous interest in the subject, most of the work is being done in civil populations. That work needs to continue, particularly in verifying specificity rates among differing neuropathological conditions. However, more work also needs to occur in the criminal setting. There is no guarantee that these measures perform in this setting as they do in the civil (typically personal injury) setting, and the base rates of malingering appear to be higher in the presentence criminal setting than in most civil settings. Additionally, there is increased reason to identify feigned versus genuine mental retardation among criminal defendants in light of the U.S. Supreme Court decision banning execution of mentally retarded criminals (*Atkins v. Virginia*, 2002). Along with that need, there exists a need to validate existing detection methods for use with the MR population. Finally, additional research needs to occur related to sensitivity rates of SVT for remote memory. It is very possible that criterion cutoffs could be established. Denney (1996) and Frederick and Denney (1998) found that most individuals with no memory for events in question fell near the mean, suggesting that simulation designs could be developed to establish criterion-based cutoff scores. Little is known about how standard competency to stand trial measures perform in regard to malingering and exaggeration of neurocognitive dysfunction. Clearly, there is a need for a great deal more research regarding exaggeration of neurocognitive deficits in the criminal forensic setting.

Notes

Opinions expressed in this chapter are those of the authors and do not necessarily represent the position of the Federal Bureau of Prisons or the U.S. Department of Justice.

1. Identified as true positives/(true positives + false negatives).
2. Identified as true negatives/(true negatives + false positives).

References

Adatto, C. P. (1949). Observations on criminal patients during narcoanalysis. *Archives of Neurology and Psychiatry*, *62*, 82–92.

Allen, L. M., Conder, R. L., Green, P., & Cox, D. R. (1997). *CARB '97 manual for the Computerized Assessment of Response Bias*. Durham, NC: CogniSyst, Inc.

Allen, L. M., & Green, P. (1999). *Severe TBI sample performance on CARB and the WMT: Supplement to the CARB '97 and Word Memory test manuals*. Durham, NC: CogniSyst, Inc.

Allen, L. M., Iverson, G. L., & Green, P. (2002). Computerized Assessment of Response Bias in forensic neuropsychology. In J. Hom & R. L. Denney (Eds.), *Detection of response bias in forensic neuropsychology* (pp. 205–225). New York: Haworth Medical Press.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Anastasi, A. (1961). *Psychological testing* (2nd ed.). New York: Macmillan.

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. *Clinical Neuropsychologist*, *21*, 899–916.

Atkins v. Virginia, 536 U. S. 304 (2002).

Baldessarini, R. J., Finkelstein, S., & Arana, G. W. (1983). The predictive power of diagnostic tests and the effects of prevalence of illness. *Archives of General Psychiatry*, *40*, 569–573.

Berry, D. T. R., & Schipper, L. J. (2007). Detection of feigned psychiatric symptoms during forensic neuropsychological examinations. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 226–263). New York: Oxford University Press.

Bianchini, K. J., Curtis, K. L., & Greve, K. W. (2006). Compensation and malingering in traumatic brain injury: A dose–response relationship? *Clinical Neuropsychologist*, *20*, 831–847.

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

Bigler, E. D. (1990). Neuropsychology and malingering: Comment on Faust, Hart, and Guilmette (1988). *Journal of Consulting and Clinical Psychology*, 58, 244–247.

Binder, L. M. (1990). Malingering following minor head trauma. *Clinical Neuropsychologist*, *4*, 25–36.

Binder, L. M. (1993). An abbreviated form of the Portland Digit Recognition Test. *Clinical Neuropsychologist*, *7*, 104–107.

Binder, L. M. (2002). The Portland Digit Recognition Test: A review of validation data and clinical use. In J. Hom & R. L. Denney (Eds.), *Detection of response bias in forensic neuropsychology* (pp. 27–41). New York: Haworth Medical Press.

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

Binks, P. G., Gouvier, W. D., & Waters, W. F. (1997). Malingering detection with the Dot Counting Test. *Archives of Clinical Neuropsychology*, *12*, 41–46.

Boone, K. B. (2007). A reconsideration of the Slick et al. (1999) criteria for malingered neurocognitive dysfunction. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective* (pp. 29–49). New York: Guilford Press.

Boone, K. B., & Lu, P. H. (2003). Noncredible cognitive performance in the context of severe brain injury. *Clinical Neuropsychologist*, *17*, 244–254.

Boone, K. B., Lu, P. H., & Herzberg, D. S. (2002). *The Dot Counting Test manual*. Los Angeles: Western Psychological Services.

Boone, K. B., Salazar, X., Lu, P., Warner-Chacon, K., & Razani, J. (2002). The Rey 15-Item Recognition Trial: A technique to enhance sensitivity of the Rey 15-Item Memorization Test. *Journal of Clinical and Experimental Neuropsychology*, *24*, 561–573.

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

Boyd, A. R., McLearen, A. M., Meyer, R. G., & Denney, R. L. (2007). *Detection of deception*. Sarasota, FL: Professional Resource Press.

Bradford, J. W., & Smith, S. M. (1979). Amnesia and homicide: The *Padola* case and a study of thirty cases. *Bulletin of the American Academy of Psychiatry and Law*, *7*, 219–231.

Brady, J. P., & Lind, D. L. (1961). Experimental analysis of hysterical blindness. *Archives of General Psychiatry*, *4*, 331–339.

Brockhaus, R., & Merten, T. (2005). Neuropsychologische diagnostik suboptimalen leistungsverhaltens mit dem Word Memory Test [Neuropsychological assessment of suboptimal performance: The Word Memory Test]. *Nervenarzt*, *75*, 882–887.

Brockhaus, R., Peker, Ö., & Fritze, E. (2003, July). *Testing effort in Turkish-speaking subjects: Validation of a translation of the Word Memory Test (WMT)*. Paper presented at the joint meeting of the International Neuropsychology Society and Gesellschaft fur Neuropsychologie, Berlin, Germany.

Bush, S. S., Ruff, R. H. Tröster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H., et al. (2005). Symptom validity assessment: Practice issues and medical necessity. *Archives of Clinical Neuropsychology*, *20*, 419–426.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Manual for administration and scoring the Minnesota Multiphasic Personality Inventory–2*. Minneapolis: University of Minnesota Press.

Chavez, M. D., Abrahams, J. P., & Kohlmaier, J. (2007). Malingering on the Social Security Disability Consultative Exam: A new rating scale. *Archives of Clinical Neuropsychology*, *22*, 1–14.

Commonwealth v. Zamarripa, 379 Pa. Super. 208, 549 A.2d 980 (1988).

Conder, R., Allen, L., & Cox, D. (1992). *Manual for the Computerized Assessment of Response Bias*. Durham, NC: CogniSyst, Inc.

Cornell, D. G., & Hawk, G. L. (1989). Clinical presentation of malingerers diagnosed by experienced forensic psychologists. *Law and Human Behavior*, *13*, 357–383.

Daubert v. Merrell Dow Pharmaceuticals, 509 U.S. 579 (1993).

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.

Denney, R. L. (1996). Symptom Validity Testing of remote memory in a criminal forensic setting. *Archives of Clinical Neuropsychology*, *11*, 589–603.

Denney, R. L. (1999). A brief Symptom Validity Testing procedure for Logical Memory of the Wechsler Memory Scale—Revised which can demonstrate verbal memory in the face of claimed disability. *Journal of Forensic Neuropsychology*, *1*(1), 5–26.

Denney, R. L. (2007). Assessment of malingering in criminal forensic neuropsychological settings. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A Neuropsychological perspective* (pp. 428–452). New York: Guilford Press.

Denney, R. L., & Wynkoop, T. F. (2000). Clinical neuropsychology in the criminal forensic setting. *Journal of Head Trauma Rehabilitation*, *15*, 804–828.

Ekman, P. (1992). *Telling lies: Clues to deceit in the marketplace, politics, and marriage*. New York: Norton.

Elwood, R. W. (1993). Psychological tests and clinical discriminations: Beginning to address the base rate problem. *Clinical Psychology Review*, *13*, 409–419.

Etherton, J. L., Bianchini, K. J., Greve, K. W., & Ciota, M. A. (2005). Test of memory malingering performance is unaffected by laboratory-induced pain: Implications for clinical use. *Archives of Clinical Neuropsychology*, *20*, 375–384.

Evans, C. (2006). What violent offenders remember of their crime: Empirical explorations. *Australian and New Zealand Journal of Psychiatry*, *40*(6/7), 508–518.

Faust, D., & Guilmette, T. J. (1990). To say it's not so doesn't prove that it isn't: Research on the detection of malingering (reply to Bigler). *Journal of Consulting and Clinical Psychology*, *58*, 248–250.

Faust, D., Hart, K., Guilmette, T. J., & Arkes, H. R. (1988). Neuropsychologists capacity to detect adolescent malingerers. *Professional Psychology: Research and Practice*, *19*, 508–515.

Frederick, R. I. (2002). A review of Rey's strategies for detecting malingered neuropsychological impairment. In J. Hom & R. L. Denney (Eds.), *Detection of response bias in forensic neuropsychology* (pp. 1–25). New York: Haworth Medical Press.

Frederick, R. I. (2003). *Validity Indicator Profile Manual, revised.* Minneapolis: Pearson Assessments.

Frederick, R. I., & Carter, M. (1993, August). *Detection of malingered amnesia in a competency evaluee*. Paper presented at the Annual Convention of the American Psychological Association, Toronto, Canada.

Frederick, R. I., Carter, M., & Powel, J. (1995). Adapting symptom validity testing to evaluate suspicious complaints of amnesia in medicolegal evaluations. *Bulletin of the American Academy of Psychiatry and the Law*, *23*, 231–237.

Frederick, R. I., Crosby, R. D., & Wynkoop, T. F. (2000). Performance curve classification of invalid responding on the Validity Indicator Profile. *Archives of Clinical Neuropsychology*, *15*, 281–300.

Frederick, R. I., & Denney, R. L. (1998). Minding your "p's and q's" when using forced-choice recognition tests. *Clinical Neuropsychologist*, *12*, 193–205.

Frederick, R. I., Sarfaty, S., Johnston, J. D., & Powel, J. (1994). Validation of a detector of response bias on a forced-choice test of nonverbal ability. *Neuropsychology*, *8*, 118–125.

Frederick, R. I., & Speed, F. M. (2007). On the interpretation of below-chance responding in forced-choice tests. *Assessment*, *14*, 3–11.

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

Gorissen, M., Sanz, J. C., & Schmand, B. (2005). Effort and cognition in schizophrenia patients. *Schizophrenia Research*, *78*, 199–208.

Gouvier, D., Hayes, J. S., & Smiroldo, B. B. (1998). The significance of base rates, test sensitivity, test specificity, and subject's knowledge of symptoms in assessing

TBI sequelae and malingering. In C. R. Reynolds (Ed.), *Detection of malingering during head injury litigation* (pp. 55–79). New York: Plenum Press.

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

Green, P. (2004). *Green's Medical Symptom Validity Test (MSVT) user manual*. Edmonton, Alberta, Canada: Green's Publishing.

Green, P. (2006). *The Nonverbal Medical Symptom Validity Test: MS Windows computer program*. Edmonton, Alberta, Canada: Green's Publishing.

Green, P., & Allen, L. M. (1999). *Performance of neurological patients on the Word Memory Test (WMT) and Computerized Assessment of Response Bias (CARB): Supplement to the Word Memory Test and CARB '97 manuals*. Durham, NC: CogniSyst, Inc.

Green, P., & Iverson, G. L. (2001). Validation of the Computerized Assessment of Response Bias in litigating patients with head injuries. *Clinical Neuropsychologist*, *15*, 492–497.

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

Green, P., Lees-Haley, P. R., & Allen, L. M. (2002). The Word Memory Test and the validity of neuropsychological test scores. In J. Hom & R. L. Denney (Eds.), *Detection of response bias in forensic neuropsychology* (pp. 97–124). New York: Haworth Medical Press.

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.

Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment*, 6, 218–224.

Greve, K., & Bianchini, K. (2006a). Classification accuracy of the Portland Digit Recognition Test in traumatic brain injury: Results of a known-groups analysis. *Clinical Neuropsychologist*, *20*, 816–830.

Greve, K., & Bianchini, K. (2006b). Should the retention trial of the Test of Memory Malingering be optional? *Archives of Clinical Neuropsychology*, *21*, 117–119.

Greve, K., & Bianchini, K. (2007). Detection of cognitive malingering with tests of executive function. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 171–225). New York: Oxford University Press.

Greve, K., Bianchini, K., Mathias, C., Houston, R., & Crouch, J. (2002). Detecting malingered performance with the Wisconsin Card Sorting Test: A preliminary investigation in traumatic brain injury. *Clinical Neuropsychologist*, *16*, 179–191.

Grosz, H. J., & Zimmerman, J. A. (1965). Experimental analysis of hysterical blindness: A follow-up report and new experimental data. *Archives of General Psychiatry*, *13*, 255–260.

Grote, C. L., Kooker, E. K., Garron, D. C., Nyenhuis, D. L., Smith, C. A., & Mattingly, M. L. (2000). Performance of compensation seeking and non-compensation seeking samples on the Victoria Symptom Validity Test: Cross-validation and extension of a standardization study. *Journal of Clinical and Experimental Neuropsychology*, *22*, 709–719.

Guttmacher, M. S. (1955). *Psychiatry and the law*. New York: Grune & Stratton.

Hartman, D. E. (2002). The unexamined lie is a lie worth fibbing: Neuropsychological malingering and the Word Memory Test. *Archives of Clinical Neuropsychology*, *17*, 709–714.

Haughton, P. M., Lewsley, A., Wilson, M., & Williams, R. G. (1979). A forced-choice procedure to detect feigned or exaggerated hearing loss. *British Journal of Audiology*, *13*, 135–138.

Heaton, R. K., Smith, H. H., Lehman, R. A. W., & Vogt, A. T. (1978). Prospects for faking believable deficits on neuropsychological testing. *Journal of Consulting and Clinical Psychology*, *46*, 892–900.

Hiscock, M., & Hiscock, C. K. (1989). Refining the forced-choice method for the detection of malingering. *Journal of Clinical and Experimental Neuropsychology*, *11*, 967–974.

Hopwood, J. S., & Snell, H. K. (1933). Amnesia in relation to crime. *Journal of Mental Science*, *79*, 27–41.

Howe, L. L., Anderson, A. M., Kaufman, D. A., Sachs, B. C., & Loring, D. W. (2007). Characterization of the Medical Symptom Validity Test in evaluation of clinically referred memory disorders clinic patients. *Archives of Clinical Neuropsychology*, *22*, 753–761.

Iverson, G. L. (2001). Can malingering be detected with the Judgment of Line Orientation Test? *Applied Neuropsychology*, *8*, 167–173.

Kay, S. R., Opler, L. A., Fiszbein, A. (1986). *Positive and Negative Syndrome Scale (PANSS) rating manual*. New York: Albert Einstein College of Medicine.

Killgore, W. D., & DellaPietra, L. (2000). Using the WMS-III to detect malingering: Empirical validation of the Rarely Missed Index (RMI). *Journal of Clinical and Experimental Neuropsychology*, *22*, 761–771.

Labarge, A. S., McCaffrey, R. J., & Brown, T. A. (2003). Neuropsychologists' abilities to determine the predictive value of diagnostic tests. *Archives of Clinical Neuropsychology*, *18*, 165–175.

Larrabee, G. J. (1992). On modifying recognition memory tests for detection of malingering. *Neuropsychology*, *6*, 23–27.

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

Larrabee, G. J. (2003b). Detection of symptom exaggeration with the MMPI-2 in litigants with malingered neurocognitive dysfunction. *Clinical Neuropsychologist*, *17*, 54–68.

Larrabee, G. J. (2005). Assessment of malingering. In G. J. Larrabee (Ed.), *Forensic neuropsychology: A scientific approach* (pp. 115–158). New York: Oxford University Press.

Larrabee, G. J., Greiffenstein, M. F., Greve, K. W., & Bianchini, K. J. (2007). Refining diagnostic criteria for malingering. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 334–371). New York: Oxford University Press.

Lee, A., Boone, K. B., Lesser, I., Wohl, M., Wilkins, S., & Parks, C. (2000). Performance of older depressed patients on two cognitive malingering tests: False positive rates for the Rey 15-Item Memorization and Dot Counting Tests. *Clinical Neuropsychologist*, *14*, 303–308.

Lees-Haley, P. R., Iverson, G. L., Lange, R. T., Fox, D. D., & Allen, L. M., III. (2002). Malingering in forensic neuropsychology: *Daubert* and the MMPI-2. *Journal of Forensic Neuropsychology, 3*(1/2), 167–203.

Leitch, A. (1948). Notes on amnesia in crime for the general practitioner. *Medical Press*, *219*, 459–463.

Lewis, J. L., Simcox, A. M., & Berry, D. T. R. (2002). Screening for feigned psychiatric symptoms in a forensic sample by using the MMPI-2 and Structured Inventory of Malingered Symptomatology. *Psychological Assessment*, *14*, 170–176.

Loring, D., Larrabee, G. J., Lee, G. P., & Meador, J. (2007). Victoria Symptom Validity Test performance in a heterogeneous clinical sample. *Clinical Neuropsychologist*, *21*, 522–531.

Loring, D. W., Lee, G. P., & Meador, K. J. (2005). Victoria Symptom Validity Test performance in non-litigating epilepsy surgery candidates. *Journal of Clinical and Experimental Neuropsychology*, *27*, 610–617.

Lynch, B. E., & Bradford, J. M. W. (1980). Amnesia: Its detection by psychophysiological measures. *Bulletin of the American Academy of Psychiatry and the Law*, *8*, 288–297.

Macciocchi, S. N., Seel, R. T., Alderson, A., & Godsall, R. (2006). Victoria Symptom Validity Test performance in acute severe traumatic brain injury: Implications for test interpretation. *Archives of Clinical Neuropsychology*, *21*, 395–404.

Marshall, P. S., & Happe, M. (2007). The performance of individuals with mental retardation on cognitive tests assessing effort and motivation. *Clinical Neuropsychologist*, *21*, 826–840.

Merten, T., Green, P., Henry, M., Blaskewitz, N., & Brockhaus, R. (2005). Analog validation of German-language symptom validity tests and the influence of coaching. *Archives of Clinical Neuropsychology*, *20*, 719–726.

Meyers, J. E., Galinsky, A. M., & Volbrecht, M. (1999). Malingering and mild brain injury: How low is too low? *Applied Neuropsychology*, *6*, 208–216.

Meyers, J. E., & Volbrecht, M. E. (1998). Validation of Reliable Digits for detection of malingering. *Assessment*, *5*, 301–305.

Meyers, J. E., & Volbrecht, M. E. (2003). A validation of multiple malingering detection methods in a large clinical sample. *Archives of Clinical Neuropsychology*, *18*, 261–276.

Miller, R. D. (2003). *People v. Palmer*: Amnesia and competency to proceed revisited. *Journal of Psychiatry and Law*, *31*, 165–185.

Millis, S. R. (2002). Warrington's Recognition Memory Test in the detection of response bias. In J. Hom & R. L. Denney (Eds.), *Detection of response bias in forensic neuropsychology* (pp. 147–166). New York: Haworth Medical Press. [Copublished simultaneously in *Journal of Forensic Neuropsychology*, *2*(3/4), 147–166]

Millis, S. R. (2004). Evaluation of malingered neurocognitive disorders. In M. Rizzo

& P. J. Eslinger (Eds.), *Principles and practice of behavioral neurology and neuropsychology* (pp. 1077–1089). Philadelphia: Saunders.

Millis, S. R., Putnam, S. H., Adams, K. M., & Ricker, J. H. (1995). The California Verbal Learning Test in the detection of incomplete effort in neuropsychological evaluation. *Psychological Assessment*, *7*, 463–471.

Mittenberg, W., Aquila-Puentes, G., Patton, C., Canyock, E. M., & Heilbronner, R. L. (2002). Neuropsychological profiling of symptom exaggeration and malingering. In J. Hom & R. L. Denney (Eds.), *Detection of response bias in forensic neuropsychology* (pp. 227–240). New York: Haworth Medical Press. [Copublished simultaneously in *Journal of Forensic Neuropsychology*, *3*(1/2), 227–240]

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.

Nelson, N. W., Boone, K. B., Dueck, A., Wagener, L., Lu, P., & Grills, C. (2003). Relationship between eight measures of suspect effort. *Clinical Neuropsychologist*, *17*, 263–272.

Nicholson v. American National Insurance, 154 F.3d 875 (8th Cir. 1998).

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 Press.

O'Connell, B. A. (1960). Amnesia and homicide. *British Journal of Delinquency*, *10*, 262–276.

Pankratz, L. (1979). Symptom validity testing and symptom retraining: Procedures for the assessment and treatment of functional sensory deficits. *Journal of Consulting and Clinical Psychology*, *47*, 409–410.

Pankratz, L., & Erickson, R. C. (1990). Two views of malingering. *Clinical Neuropsychologist*, *4*, 379–389.

Pankratz, L., Fausti, S. A., & Peed, S. (1975). A forced-choice technique to evaluate deafness in the hysterical or malingering patient. *Journal of Consulting and Clinical Psychology*, *43*, 421–422.

Parkin, A. J. (1996). Focal retrograde amnesia: A multi-faceted disorder? *Acta Neurologica Belgica*, 96(1), 43–50.

Parwatikar, S. D., Holcomb, W. R., & Menninger, K. A., II. (1985). The detection of malingered amnesia in accused murderers. *Bulletin of the American Academy of Psychiatry and the Law*, *13*, 97–103.

Paul, D. S., Franzen, M. D., Cohen, S. H., & Fremouw, W. (1992). An investigation into the reliability and validity of two tests used in the detection of dissimulation. *International Journal of Clinical Neuropsychology*, *14*, 1–9.

Power, D. J. (1977). Memory, identification and crime. *Medicine, Science, and the Law*, *17*, 132–139.

Price, G. E., & Terhune, W. B. (1919). Feigned amnesia as a defense reaction. *Journal of the American Medical Association*, *72*, 565–567.

Reitan, R. M., & Wolfson, D. (2002). Detection of malingered and invalid test results using the Halstead–Reitan Battery. In J. Hom & R. L. Denney (Eds.), *Detection of negative response bias in forensic neuropsychology* (pp. 275–314). New York: Haworth Medical Press. [Copublished simultaneously in *Journal of Forensic Neuropsychology*, *3*(1/2), 275–314]

Rey, A. (1941). L'Examen psychologie dans las cas d'encephalopathie traumatique [The psychological examination in cases of traumatic encephalopathy]. *Archives de Psychologie*, *28*, 286–340.

Rey, A. (1958). *L'examen clinique en psychologie* [The psychological examination]. Paris: Presses Universitaires de France.

Reznek, L. (2005). The Rey 15-Item Memory Test for malingering: A meta-analysis. *Brain Injury*, *19*, 539–543.

Richman, J., Green, P., Gervais, R., Flaro, L., Merten, T., Brockhaus, R., et al. (2006). Objective tests of symptom exaggeration in independent medical examinations. *Journal of Occupational and Environmental Medicine*, *48*, 303–311.

Rogers, R. (1986). *Conducting insanity evaluations*. New York: Van Nostrand Reinhold.

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

Rosenfeld, B., Sands, S. A., & Van Gorp, W. G. (2000). Have we forgotten the base rate problem?: Methodological issues in the detection of distortion. *Archives of Clinical Neuropsychology*, *15*(4), 349–359.

Ross, S. M. (2000). Profound retrograde amnesia following mild head injury: Organic or functional? *Cortex*, *36*(4), 521–537.

Rubinsky, E. W., & Brandt, J. (1986). Amnesia and criminal law: A clinical overview. *Behavioral Sciences and the Law*, *4*, 27–46.

Schacter, D. L. (1986). Amnesia and crime: How much do we really know? *American Psychologist*, *41*, 286–295.

Schmand, B., Kuipers, T., Van der Gaag, M., Boxveld, J., Bulthuis, F., & Jellema, M. (1994). Cognitive disorders and negative symptoms as correlates of motivation deficit in psychotic patients. *Psychological Medicine*, *24*, 869–884.

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.

Simon, M. J. (1994). The use of the Rey Memory Test to assess malingering in criminal defendants. *Journal of Clinical Psychology*, *50*, 913–917.

Slick, D. J., Hopp, G., Strauss, E., & Thompson, G. B. (1997). *Victoria Symptom Validity Test version 1.0 professional manual*. Odessa, FL: Psychological Assessment Resources.

Slick, D. J., Iverson, G. I., & Green, P. (2000). California Verbal Learning Test indicators of suboptimal performance in a sample of head-injury litigants. *Journal of Clinical and Experimental Neuropsychology*, *22*, 569–579.

Slick, D. J., Sherman, E. M. S., & Iverson, G. L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *Clinical Neuropsychologist*, *13*, 552–553.

Slick, D. J., Tan, J. E., Strauss, E., Mateer, C. A., Harnadek, M., & Sherman, M. S. (2003). Victoria Symptom Validity Test scores of patients with profound memory impairment: Nonlitigant case studies. *Clinical Neuropsychologist*, *16*, 495–505.

Suhr, J. A., & Barrash, J. (2007). Performance on standard attention, memory, and psychomotor speed tasks as indicators of malingering. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits* (pp. 131–170). New York: Oxford University Press.

Sweet, J. J., & King, J. H. (2002). Category Test validity indicators: Overview and practice recommendations. In J. Hom & R. L. Denney (Eds.), *Detection of negative response bias in forensic neuropsychology* (pp. 241–274). New York: Haworth Medical Press. [Copublished simultaneously in *Journal of Forensic Neuropsychology*, *3*(1/2), 241–274]

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

Taylor, P. J., & Kopelman, M. D. (1984). Amnesia for criminal offences. *Psychological Medicine*, *14*, 581–588.

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.

Teichner, G., Waid, L. R., & Buddin, W. H. (2005, October). *The Medical Symptom Validity Test (MSVT): Data from a sample demonstrating a range of neurological and psychiatric disorders*. Poster presented at the National Academy of Neuropsychology Annual Conference, Tampa, FL.

Theodor, L. H., & Mandelcorn, M. S. (1973). Hysterical blindness: A case report and study using a modern psychophysical technique. *Journal of Abnormal Psychology*, *82*, 552–553.

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

Tombaugh, T. N. (2002). The Test of Memory Malingering (TOMM) in forensic psychology. In J. Hom & R. L. Denney (Eds.), *Detection of response bias in forensic neuropsychology* (pp. 69–96). New York: Haworth Medical Press.

Tydecks, S., Merten, T., & Gubay, J. (2006). The Word Memory Test and the One-in-Five-Test in an analog study with Russian speaking participants. *International Journal of Forensic Psychology*, *1*, 29–37.

United States v. Greer, 158 F.3d 228 (5th Cir. 1998).

Vallabhajosula, B., & van Gorp, W. G. (2001). Post-*Daubert* admissibility of scientific evidence on malingering of cognitive deficits. *Journal of the American Academy of Psychiatry and the Law*, *29*, 207–215.

Vickery, C. D., Berry, D. T. R., Inman, T. H., Harris, M. J., & Orey, S. A. (2001). Detection of inadequate effort on neuropsychological testing: A meta-analytic review of selected procedures. *Archives of Clinical Neuropsychology*, *16*, 45–73.

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 Press.

Vilar-López, R., Santiago-Ramajo, S., Gómez-Río, M., Verdejo-García, A., Llamas, J. M., & Pérez-García, M. (2007). Detection of malingering in a Spanish population using three specific malingering tests. *Archives of Clinical Neuropsychology*, *22*, 379–388.

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.

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

Wilson v. United States, 391 F.2d 460 (D.C. Cir. 1968).

Wynkoop, T. F., & Denney, R. L. (1999). Exaggeration of neuropsychological deficit in competency to stand trial. *Journal of Forensic Neuropsychology*, *1*, 29–53.

Wynkoop, T. F., & Denney, R. L. (2001). Test reviews: Computerized Assessment of Response Bias (CARB), Word Memory Test (WMT), and Memory Complaints Inventory (MCI). *Journal of Forensic Neuropsychology*, *2*, 71–77.

Wynkoop, T. F., & Denney, R. L. (2005). Test review: Green's Word Memory Test for Windows. *Journal of Forensic Neuropsychology*, *4*, 101–105.

Yanez, Y., Fremouw, W., Tennant, J., Strunk, J., & Coker, C. (2006). Effects of severe depression on TOMM performance among disability-seeking outpatients. *Archives of Clinical Neuropsychology*, *21*, 161–165.

Zachary, R. A. (1986). *Shipley Institute of Living Scale—Revised manual*. Los Angeles: Western Psychological Services.

Zimmerman, J., & Grosz, H. J. (1966). "Visual" performance of a functionally blind person. *Behaviour Research and Therapy*, *4*, 119–134.

Guilford Publications
72 Spring Street
New York, NY 10012
212-431-9800
800-365-7006
www.guilford.com