

Posttraumatic Stress Disorder: Assessing Response Style and Malingering

Steve Rubenzer

Received: 19 January 2009 / Accepted: 9 April 2009 / Published online: 27 May 2009
© Springer Science + Business Media, LLC 2009

Abstract Posttraumatic stress disorder (PTSD) may form the basis for disability or worker's compensation claims or a personal injury lawsuit. While now achieving widespread acceptance among treating professionals and the public, PTSD is the subject of several controversies and the possibility of faking in a compensation context. There appears to be a dramatic split among mental health professionals who write primarily from a treatment or plaintiff perspective and those who take a more skeptical approach. This article reviews recent developments in the assessment of malingering, including symptom validity measures, and applies them to the assessment of PTSD. Recommendations for current practice are provided.

Keywords Posttraumatic stress disorder, PTSD · Malingering · Faking · Exaggeration · Symptom validity testing, SVT

Posttraumatic stress disorder (PTSD) is one of the most common psychological disorders claimed in a personal injury context and may form the basis for a disability or worker's compensation claim. PTSD is nearly unique among psychiatric disorders, in that the diagnostic criteria specify the presumed etiology: exposure to a significant trauma or stressor. Thus, the very definition of the disorder helps to establish the legal requirement, in a personal injury suit, of proximate cause (Shuman 2003).

The Diagnostic and Statistical Manual of Mental Disorders TR (DSM-IV-TR; American Psychiatric Association 2000) states that malingering should be ruled out in medicolegal

contexts, and the DSM-IV-TR *Differential Diagnosis* section for PTSD specifically includes this possibility. PTSD may be easily faked on self-report measures. Many are face-valid scales with little or no check on overendorsement or fabrication, and, not surprisingly, most subjects are able to portray PTSD when presented with symptom checklists (Burgess and McMillen 2001; Hickling et al. 2002; Lees-Haley 1990). Guriel and Fremouw (2003) reviewed approaches to assessing exaggeration or malingering of PTSD and concluded no extant instrument could be considered fully adequate. The present article will seek to update readers with substantial developments since the Guriel and Fremouw paper, most notably that of cognitive symptom validity testing now so prevalent in neuropsychology.

Rosen (2004a, b, 2006; see also McNally 2003; Rosen and Taylor 2007) argued that the literature on PTSD may be badly compromised by the failure of researchers to assess for malingering among presenting patients, even when the entire sample is compensation seeking. This failure potentially contaminates much of what is known about the disorder. For example, one correlate of PTSD is antisocial behavior (McNally 2003), which may include deception, exploitation, and substance abuse. Antisocial behavior and drug use are sometimes viewed as a *consequence* of PTSD without making any serious attempt to determine if such traits were present *before* the alleged injury. Further, antisocial personality disorder is one of four DSM-IV-TR indicators of potential malingering.

Numerous studies report high levels of psychiatric comorbidity with PTSD (Brady 1997; Brady et al. 2000; Brunello et al. 2001; Kessler et al. 1995; Skodol et al. 1996). There are at least five explanations for this covariance with very different implications: (a) that the comorbidity is "real," [e.g., that depression and other symptoms are frequent, co-occurring responses to trauma], (b) that apparent comorbidity is due to intentional symptom overendorsement, (c) that comorbidity reflects underlying neuroticism or negative

S. Rubenzer (✉)
11914 Astoria, Suite 490,
Houston, TX 77089, USA
e-mail: rubenzer.steve@att.net

affectivity (and thus is real, but artifactual), (d) that apparent comorbidity is the result of acquiescent response style, and (e) that apparent comorbidity is the result of a dramatizing communication style. Thus, the basic issue of comorbidity is a morass. The failure to consider response styles has resulted in several published recommendations that journal editors demand disclosure of the litigation status of study participants and that those with incentives to exaggerate be identified and analyzed separately from those without such motivations (Rosen 2004a, b, 2006; Rosen and Taylor 2007). Without such provisions, mean profiles of supposed genuine PTSD patients or claimants on the Minnesota Multiphasic Personality Inventory 2 (MMPI-2) or Personality Assessment Inventory (PAI) may be uninterpretable.

PTSD appears widely accepted in the early twenty-first century by both professional groups and the lay public, but its core features and its integrity as a discreet diagnostic syndrome have been challenged (McNally 2003, 2004; Rosen 2004a; Taylor and Asmundson 2008). Some studies find that the degree of trauma experienced has little relationship to the severity of symptoms or impairment, while indices of adjustment before the trauma are solid predictors (Bowman and Yehuda 2004). A recent meta-analysis of 77 studies found that previous psychiatric history, childhood abuse, and family psychiatric history were consistently associated with development of PTSD. Less consistent predictors included gender, race, age, education, previous trauma, and general childhood adversity (Brewin et al. 2000). Another review reported lower intelligence, neuroticism, negativistic personality traits, and dissociation surrounding the trauma as predictors of subsequent PTSD diagnosis (McNally 2003).

The focus on personality and historical predictors is a shift from early conceptions of the disorder. Even life-threatening events often do not lead to PTSD: Breslau and Kessler (2001) found 89.6% of Detroit adults reported experiencing at least one traumatic event meeting DSM-IV-TR criterion A, yet only 9.2% developed PTSD. Further, two thirds of those who initially showed symptoms showed improvement or resolution within 3 months (Rosen 2004a). Thus, persisting, chronic PTSD is the exception, not the rule. Yet, the public perception appears quite different. Noting that PTSD has seemingly become the expected response to tragedy, Summerfield (2004) asserted:

An expansive mental health industry has in effect promoted the idea that the trials of life represent noxious influences easily able to penetrate the average citizen, not just to hurt but to disable. This is to endorse a much thinner-skinned version of a person than previous generations would have recognized or respected (p. 234).

Early writers conceptualized PTSD as a disorder produced by a stressor that would cause extreme distress

to virtually anyone (Bowman and Yehuda 2004; Young and Yehuda 2006). However, predictive studies have long demonstrated the role of personality traits, such as neuroticism (Miller 2003), which are long-standing and are associated with a variety of negative emotions and maladaptive coping mechanisms (McCrae and Costa 1990). Emphasizing the contribution of personality factors, Lees-Haley (1997a, b) noted that the most common personal injury claimant MMPI-2 profile is 1/3 and offered this synopsis:

Many of the profiles which appear in this sample are indicative of chronic conditions in addition to any current discomfort....These elevations also suggest poor insight and denial, which may support the thesis that plaintiff exaggeration is better thought of in terms of pathology or rationalization than malingering, which implies conscious intent. Descriptions such as sad, bitter, cynical, miserable, pessimistic, and dysphoric along with lack of energy, concentration problems, and physical problems support the conclusion that these are genuinely disturbed people, regardless of exaggeration.

The modal plaintiff appears to be an unhappy somatizer involved in a social context that encourages rationalization, projection of blame, and complaining (p. 753).

In less inferential terms, a model case of PTSD may be the result of a stressful event on someone already above average in neuroticism, leading to an increase of distress over an already elevated baseline. Even emotionally hardy people might suffer prototypical PTSD symptoms as the result of a severe stressor, while those scoring average in neuroticism would have an intermediate risk of PTSD development.

Many treating clinicians and writers appear overly trusting about the honesty of their PTSD patients or subjects. For example, Daly and Johnson (2002), while reporting that all their subjects were seeking compensation, minimized the possibility of malingering: “The *victims* [italics added] in this study appear to have been genuine, honest people... They were largely a law-abiding group who had previously shown respect for, and trust in, authority” (p. 463). Blanchard and Hickling (1997) did not collect medical records and stated that MMPI-2 findings may falsely label their patients as exaggerating. Ironically, these same two investigators (Hickling et al. 2002) earlier researched their own clinicians’ ability to detect actors trained as simulators and found troubling results. Despite having access to various self-report scales and psychophysiological data, clinicians described as “highly experienced with a trauma population” failed to identify any of the six actors. When told the nature of the study, clinicians identified three of the six actors but misidentified three patients.

While clinicians who work in treatment settings may believe malingering is rare (Blanchard and Hickling 1997), there is little check on this assumption and much disagreement. Drukteinis (2003) and Resnick (2003) asserted that outright malingering is rare, although exaggeration is common. However, Frueh et al. (2005) reported that 69–89% of veterans treated in psychiatric trauma clinic went on to seek disability—an extraordinarily high percentage compared to disability rates among World War II veterans (Goldstein et al. 1987), Vietnam era naval aviator prisoners of war (POWs; Nice et al. 1996), and holocaust survivors (Eaton et al. 1987). Freeman et al. (2008) reported that 53% of veterans presenting for treatment failed the Structured Interview for Reported Symptoms (SIRS; Rogers 1992) while another 24% obtained scores in the *indeterminate* range. Burkett and Whitley (1998) estimated, based on review of military records, that about 75% of Vietnam veterans who receive disability due to PTSD were never exposed to combat. Merten et al. (2006) reported 51.1% of PTSD claimants undergoing independent medical evaluations failed cognitive validity tests. Demakis et al. (2008) reported 29% of PTSD claimants failed at least one cognitive validity measure, while 49% scored above the predetermined cutoff scores on MMPI-2 validity scales (F, F(p), Fake Bad Scale (FBS)) or the negative bias scale of the Detailed Assessment of Posttraumatic Stress (DAPS; Briere 2001).

Neuropsychologists have been particularly active in the assessment of response style, which they extend to ostensible cognitive tasks. Because neuropsychological tests require maximal effort to be valid, researchers have developed tests sensitive to less than adequate effort or intentional failure. The National Academy of Neuropsychologists (NAN) recently issued a position statement declaring *symptom validity testing* (SVT) medically necessary for *all* neuropsychological examinations (Bush et al. 2005). Not infrequently, these examinations involve claimants without brain injury, including conditions such as depression, chronic pain, fibromyalgia, and PTSD. Prominent SVT tests include the Word Memory Test (Green 2005), the Computerized Assessment of Response Bias (Allen et al. 1997), the Victoria Symptom Validity Test (Slick et al. 1997), and the Test of Memory Malingering (TOMM; Tombaugh 1996). Aside from these freestanding tests, indices utilizing scores from within existing cognitive tests, such as the Wechsler Adult Intelligence Scale (WAIS-III; Mittenberg et al. 2001; Greve et al. 2003), and Wechsler Memory Scale-III (see Larrabee 2007a), and for neuropsychological test batteries (Meyers and Volbrecht 2003) have been developed.

As more tests of response style were developed, psychologists proposed more objective ways of interpreting and combining test results and other observations (Greiffenstein

et al. 2004; Pankrantz 1998; Rogers 1997a). Greve and Bianchini (2004a) proposed setting cutoff scores of individual scales to achieve a specificity of 0.95 (false-positive error rate of 5%), allowing sensitivities (ability to detect suboptimal effort) to fall where they may. Often, these are below 0.30. Since most response-style tests have low to moderate sensitivities, many authors (Bush et al. 2005; Iverson and Franzen 1996; Lynch 2004; Meyers and Volbrecht 2003; Nelson et al. 2003) recommend employing multiple validity tests and procedures. Although Otto (2008) cautioned that using multiple highly correlated feigning measures increases the risk of a false-positive error, several neuropsychologists have argued that such tests have very small intercorrelations among clinical groups. More important, empirical evaluations have found better classification figures for failure on two or three validity indicators than for any of the individual tests (Larrabee 2003, 2008; Meyers and Volbrecht 2003; Vickery et al. 2004; Victor et al. 2008).

The most successful effort to synthesize clinical observations of potential malingering has been *the Slick criteria* proposed by Slick et al. (1999), which focuses on detection of feigned cognitive deficits. Briefly summarized, the criteria require the presence of an external incentive and some combination of implausible performance on one or more neuropsychological or self-report tests and inconsistent or implausible presentation. The scoring rules yield groups described as *definite*, *probable*, and *possible* malingering. Below-chance performance on a forced-choice validity test is the one sign that, given only the presence of an external incentive, is sufficient to render a judgment of *definite* exaggeration: no amount of impairment can account for a performance that is significantly below chance. Several response-style studies have used the Slick criteria classifications to create criterion groups, and some forensic examiners reference these classifications in their reports, potentially providing a reliable, validated, and generally accepted classification process. Larrabee et al. (2007) made suggestions for modifications to the original Slick criteria, noting that multiple studies have suggested that failure on any three validity indicators, including either psychiatric or cognitive measures, provides a highly sensitive and specific decision rule. Victor et al. (2008) recently argued for a “two-failure” rule.

Assessment of PTSD in Forensic and Compensation Contexts

There are differing opinions on the appropriate focus of PTSD assessment, but most writers emphasize the need for use of structured interviews or tests and use of collateral information. Aside from these basic guidelines, there are widely differing

recommendations. Vasterling and Bailey (2005) asserted, “Attention to neuropsychological impairment is paramount in clinical settings” (p. 200). Keane et al. (1998) suggested that psychophysiological assessment had strong empirical support and surpassed the classification accuracy of the dexamethasone suppression test, while a subsequent chapter by Keane et al. 2003 was far less bullish. Other psychophysical assessment investigators, such as Pitman and Orr (2003), concluded that such measures “are limited and require careful scrutiny to avoid misuse” (p. 222).¹ Psychometric assessment of level of functioning/behavioral impairment is rarely mentioned.

While psychiatric symptoms predominate over cognitive ones in the DSM-IV-TR diagnostic criteria for PTSD, inability to recall aspects of the trauma is counted toward criterion C, and “concentration problems” can be counted toward criterion D (persistent symptoms of increased arousal). PTSD patients and claimants often report or display memory impairment (Sachinvala et al. 2000; Solomon and Mikulincer 2006). Nonetheless, few chapters on forensic PTSD evaluation address cognitive assessment. Although several authors (Beblo et al. 2005; Brewin et al. 2007; Moore 2009; Vasterling and Bailey 2005) concluded that there is substantial evidence of cognitive impairment in PTSD, Vasterling and Bailey acknowledged that extant studies have failed to assess for optimal effort or feigning. Others (Cromwell et al. 2002; Twamley et al. 2004) observed no significant deficits. Horner et al. (2007) reported that veterans diagnosed with PTSD who passed the TOMM produced lower scores on measures of attention, psychomotor speed, visuoconstruction, memory, and cognitive flexibility than did veterans without psychiatric diagnoses. However, PTSD patients differed from other psychiatric patients only on attention. Demakis et al. (2008) found virtually no correlation between the severity of PTSD symptoms and cognitive impairment on neuropsychological tests. Further, those that passed cognitive validity tests achieved a mean *T*-score of 48.8 on the cognitive battery, suggesting no substantial impairment.

The Word Memory Test (WMT) scoring program (Green 2005) lists scores for various diagnostic groups, including PTSD, distinguishing between those that passed or failed the validity portion of the test. The five normal groups produced scores on the WMT’s functional memory scales from 0.1 to 1.1 SDs above PTSD patients that passed the WMT. However, the normal groups also scored higher by 0.38 SDs on two of the three primary effort indicators. This suggests some patients may have not

performed to the best of their ability, despite passing the WMT.²

Some recent studies examined cognitive deficits in PTSD without assessing effort but yielded interesting findings. Burriss et al. (2007) reported that depressive symptoms mediated the cognitive deficits they observed in PTSD patients (see also Johnsen et al. 2008), although Rohling et al. (2002) and Green (2008) reported that depression does not produce lowered cognitive scores when examinees exert good effort. Johnsen et al. (2008) reported deficits on delayed recall, which was mediated by less efficient learning in the PTSD sample. However, attention and recognition memory was intact. Lastly, two recent large studies (Kremen et al. 2007; Parslow and Jorm 2007) and a review (Moore 2009) have concluded that cognitive deficits in PTSD precede the trauma and, in fact, act as risk factors for development of PTSD symptoms. Parslow and Jorm found that those experiencing PTSD symptoms showed less improvement in an immediate word recall task from the first to second testing date. However, PTSD symptoms of reexperiencing and arousal were predicted by lower word recall, digit span, coding speed, and verbal intelligence assessed 3 years before the trauma. Kremen et al. found that twins with PTSD had Armed Forces Qualification Test scores that averaged 3.87 percentile points lower than their cotwins before their military service. Genetic factors were assessed to account for 5% of the variance in PTSD status through cognitive ability, plus an additional 26% through noncognitive factors.

Many influential PTSD writers devote relatively little attention to the assessment of response style (Blanchard and Hickling 1997; Briere 2002; Keane et al. 2003; Koch et al. 2005; Shuman 2003; Simon 2003; Young et al. 2006). Others devote a chapter in a book (Koch et al. 2005), while some suggest it is a primary issue to be addressed (Druketeinis 2003; Greiffenstein et al. 2004; Gurriel and Fremouw 2003; Knoll and Resnick 2006; McGuire 1999; Resnick 2003, Resnick et al. 2008; Rubenzer 2005, 2006; Young et al. 2007). Apart from agreement on the need for objective measures and, to a lesser extent, use of collateral information, there appears to be little agreement between those who trust the clinical presentation of claimants and those who view assessment of response style as crucial.

Malingering PTSD

There are indications that malingering may be common among PTSD claims. In the *Aleutian Enterprise* sinking,

¹ A recent meta-analysis (Pole 2007) found small to moderate effects of PTSD on psychophysical measures such as resting heart rate and changes in physiological measures to standard and idiographic trauma-related stimuli.

² The WMT effort indicators do not correlate with IQ or age above age 10, and even modest decrements in performance from normative levels are associated with depressed scores on other measures.

86% of survivors reported PTSD symptoms leading to disability, far exceeding the more typical figures of 25 to 40% in similar tragedies. Postlitigation interviews with these claimants found that attorneys coached most litigants (Rosen 1995). Some veterans who claimed PTSD never experienced combat or, in some cases, were never even in the armed services (Burkett and Whitley 1998; Frueh et al. 2000; Monnier et al. 2005). Such findings suggest that forensic examiners should not assume exposure to the claimed traumatic event in the absence of reliable collateral information.

Some signs of PTSD malingering are provided by Resnick (1997, 2008), Hall and Hall (2006), and Greiffenstein et al. (2004). The latter authors used the sum of such indicators to classify subjects and it is possible to calculate an effect size, although the authors did not report one. For men, it was quite large ($d=2.9$), while, for women, it was enormous ($d=7.2$). There were only 15 men in the clinical comparison group so the difference between sexes may not be reliable. Such results are promising, but there is a need for further empirical evaluation of this and similar lists of improbable symptom patterns.

Self-report Inventories and Structured Interviews

Observers have long commented on the tendency of PTSD patients to produce elevated scores on MMPI validity indices (Hyer et al. 1987, 1988). At first, many viewed this as a function of the severity of the disorder and the variety of its symptoms. Over time, however, others noted that the extremely pathological test scores observed were inconsistent with the outpatient status of most PTSD patients and that the disability rates observed far exceeded that seen in previous wars or tragedies.

Most PTSD diagnostic interviews and self-report scales represent straightforward queries about symptoms and allow motivated persons to present themselves as having the requisite symptoms to meet the diagnostic criteria (Burges and McMillan 2001; Lees-Haley and Dunn 1994; Hickling et al. 2002). Many instruments have no means to detect exaggeration or unreliable responding. One structured interview, the *Clinician-Administered PTSD Scale* (CAPS; Blake et al. 1995), has a consistency scale to assess unreliable responding, but the only study that examined its utility found it completely ineffective at identifying exaggeration (Hickling et al. 2002). The Atypical Responding Scale on the Trauma Symptom Inventory (TSI; Briere 1995), a self-report inventory, has produced inconclusive results (discussed below).

The Minnesota Multiphasic Personality Inventory-2 The MMPI-2 (Butcher et al. 1989) has two scales, PS and PK, designed to assess PTSD symptoms. These scales, however, appear highly sensitive to general distress and are not

specific to PTSD (Scheibe et al. 2001), and they are not very sensitive to PTSD as well (Senior and Douglas 2000). More useful are the MMPI-2 validity scales, although most traditional indices (L, F, F-K, K) show mixed evidence of ability to distinguishing malingerers from those with genuine PTSD (Greene 2000; Greiffenstein et al. 2007; Perconte and Goreczny 1990; Rogers et al. 2003). Recommended cutoff scores for the F scale vary dramatically: Lees-Haley (1992) recommended *T*-scores equal or greater than 62, while Elhai et al. (2000) stated that even a *T*-score of 120 should not be considered strong evidence of malingering. Rogers et al. (2003) recommended use of F(p) and Ds for assessment of malingered PTSD, noting both substantial effect sizes and stable cutoff scores across studies, and utilized cutoff scores of ≥ 35 (raw) for Ds and $F(p) \geq 7$ in a subsequent study (Rogers et al. 2008).

The F scale has a number of items that also load on the clinical scales, and elevations among PTSD patients are thus somewhat ambiguous. The F(p) scale was developed to offer a purer measure of intentional exaggeration (Arbisi et al. 1995) and a number of studies found the F(p) scale to be the most effective MMPI-2 scale at detecting malingered PTSD (Arbisi et al. 2006; Bury and Bagby 2002; Elhai et al. 2000, 2001). A meta-analytic review by Rogers et al. (2003) came to the same conclusion. However, several authors (Greiffenstein et al. 2004; Larrabee 2005, 2007b; Nelson et al. 2006) noted that most of these studies had a fatal design flaw: they compared students asked to simulate PTSD with claimants or veterans diagnosed with PTSD, who are at high risk for malingering or exaggeration (Burkett and Whitley 1998; Frueh et al. 2005). The authors did not assess response style among patients or claimants. The effectiveness of F(p) with better-designed studies is far less impressive: Nelson et al. (2006), in a meta-analysis of response-style studies that included the FBS, found an average effect size for F(p) of only 0.35—substantially smaller than most other indices. Resnick et al. (2008) reported more positive findings in a quantitative summary of previous studies, but the basis for selection of studies was not clear; some important ones (e.g., Greiffenstein et al. 2004) were omitted, and many of those included had the design flaw described above. Another recent study (Crawford et al. 2006) found claimants actually scored lower than the clinical comparison group on F(p) and Berry and Schipper (2007) noted that F(p)'s effect size decreased from simulator to known group designs. While three other studies since the Nelson meta-analysis reported positive findings; one (Marshall and Bagby 2006) used the same faulty design as previous studies. Arbisi et al. (2006) used veteran claimants for both honest and simulating groups. Although the authors assured participants that their responses would not affect their claim status, it is doubtful the 25% possibility of a \$25 reward would motivate many

participants to risk a lifetime of monetary and educational benefits. Efendov et al. 2008 compared claimants and recovered PTSD patients asked to simulate current PTSD symptoms. However, the authors only assessed claimants for feigning on formal measures if the staff suspected malingering. Clinicians have not demonstrated the ability to make valid judgments of response style in the absence of formal measures (Hickling et al. 2002; Miller 2005; Rosen and Phillips 2004; but see Victor et al., 2008). In addition, authors' decisions about malingering from testing were likely heavily influenced by F, Fb, and F(p), while FBS scores were not available (Sellbom, personal communication, January 5, 2009). The study thus does not adequately assess for feigning and confounds predictors and the criterion. Lastly, inspection of the item content of F(p) reveals that none of the items are clearly related to PTSD criteria and that endorsement of several items would seem quite undesirable in a personal injury context. As Lees-Haley et al. (1991) hypothesized, the goal of a claimant is to present oneself as an honest stable person who has suffered a terrible injury. A number of items on F(p), when endorsed in the keyed direction, would suggest otherwise. While a raw score of 7 or higher is substantial evidence of exaggeration or invalid responding (Greene 2000, 2008; Rogers et al. 2003), such scores may only occur in cases of confusion or blatant overreporting or falsification. When simulators are more sophisticated, its sensitivity is likely to be low. Consequently, a low F(p) should not be taken as strong evidence of honest responding,

Lees-Haley et al. (1991) developed the FBS to detect exaggeration in a civil context and Pearson Assessments recently added it to its extended score report. The 39-item FBS contains only four items that overlap with DSM-IV PTSD criteria and only six that are scored on the empirically derived PK scale. A number of studies found FBS to be the best or the only MMPI-2-valid response-style scale when psychological injury or disability due to a nonpsychotic disorder is claimed (Crawford et al. 2006; Greiffenstein et al. 2004; Larrabee 2003; Lees-Haley et al. 1991; Lees-Haley 1992), while others found it to be of little use. Most of the latter suffered from the design flaw discussed above for F(p). A recent meta-analysis of methodologically adequate FBS studies found an average effect size of 0.96, which was larger than for other MMPI-2 validity scales (Nelson et al. 2006). Only five effect sizes from three studies were available for PTSD, yielding an average $d=0.99$ (95% CI=0.76–1.21). Because of the limited database, the authors suggested caution in the use of FBS with PTSD patients. Even these limited data were compromised by making comparisons between all three groups in Greiffenstein et al. (2004; N. W. Nelson, personal communication, July, 2008), when the serious injury group was not screened for validity and thus did not meet

inclusion requirements of the meta-analysis. Another included study (Lees-Haley 1992) contrasted pseudo-PTSD patients with other claimants who claimed personal injury but did not elevate the PK or PS scales and so did not appear to be endorsing PTSD symptoms.

Although there has been controversy surrounding the FBS and its value (Arbisi and Butcher 2004; Ben-Porath et al. 2009; Butcher et al. 2003, 2008; Greve and Bianchini 2004b; Lees-Haley and Fox 2004; Williams et al. 2009), recent studies (Ardolf et al. 2007; Bianchini et al. 2008; Demakis et al. 2008; Sellers et al. 2006; Wygant et al. 2007) provide evidence of FBS's validity. Unlike the F family and other traditional response-style scales, FBS typically shows a substantial relationship to performance on cognitive SVTs (Larrabee 2003; Sellers et al. 2006; Wygant et al. 2007; but see Whitney et al. 2008). A strongly supportive study (Crawford et al. 2006) was published after the Nelson et al. analysis, although Rogers (2008a) was harshly critical of its methodology. Two recent reviews of MMPI-2 response-style indicators in compensation contexts supported the use of FBS relative to other MMPI-2 validity indicators (Berry and Schipper 2007; Greiffenstein et al. 2007). Lastly, a recent survey (Sharland and Gfeller 2007) of neuropsychologists, the discipline most active in creating and researching response-style measures over the past decade, found the FBS to be among the top five most frequently used validity indicators even before it was added to the Pearson scoring reports.

The FBS is not without potential problems. Widely different cutoff scores are maximally effective across studies, and two groups appear prone to elevated scores: females with a prior psychiatric history and patients with severe objectively manifested physical distress, such as drug withdrawal. Greiffenstein et al. (2007) reported that data aggregated across studies with known groups found a cutoff score of 25 or greater associated with a specificity of 0.95, whereas cutoff scores of 28 or greater had a specificity of 0.988. While the FBS has appeared more sensitive than other MMPI-2 scales in several studies, when cutoff scores are set to minimize false-positive errors, sensitivity naturally suffers as a result. Greve et al. (2006) found that FBS, at a cutoff score of greater than 27 set to achieve a specificity of 0.95, had sensitivities from 0.23 to 0.71 in groups with increasingly probable exaggeration (from *suspect malingering* to *definite malingering*). However, 8.9% and 11.1% of *incentive only* and *no incentive subjects* scored above the cutoff—about twice the number expected. While the most sensitive of the validity scales assessed, FBS also had the most false positives. Ben-Porath et al. (2009) acknowledged that litigation and legitimate injury or somatoform disorder can create significantly elevated FBS scores (raw scores 23–28), and such scores do not distinguish feigning from nonfeigning groups.

However, scores above 28 were rarely found (false-positive rates from .01 to 0.03) in nonlitigating cases.

Other established MMPI-2 scales potentially useful in assessment of exaggerated PTSD are Ds, Ds-r, and Ego Strength (ES; Lees-Haley 1997a, b; Rogers et al. 2003). Meyers et al. (2002) created a validity index that combines seven response-style indicators (e.g., F, F(p), F-K, FBS, O-S, Ds-r, ES), and, although it outperformed its constituent parts in the original study, it has not exceeded other validity indices in cross-validation attempts by others (Bianchini et al. 2008; Greve et al. 2006).

Greene (2008) published clinical norms for the MMPI-2's major response-style indicators based on Caldwell's (2003, unpublished raw data) data set of over 50,000 protocols from various settings. Scores at the 95th percentile should have a false-positive rate of no more than 5%, whereas those at the 99th percentile should have false-positive rates of 1% (see Table 1). Three additional factors warrant attention. Rogers (1997a) reported estimated malingering rates of 7.4% and 7.8% in clinical samples, and some patient groups in the Caldwell data set probably had substantial potential for financial motivation. Thus, the percentile scores given are likely conservative estimates of false-positive rates among honest responders. This is supported by the very high *T*-scores required to reach the 98th or

99th percentile, particularly for F(p), since it overlaps minimally with actual psychopathology. Most notable is that the error rates for a given cutoff score on FBS are considerably higher than reported by Greiffenstein et al. (2007). This may be because the norms of Greiffenstein et al. used subjects without compensation motivation rather than unscreened clinical samples with unknown proportions of response bias.

Elhai et al. (2002) created the F_{ptsd} to detect feigned or exaggerated PTSD. Although it showed incremental validity over other response-style scales in the original study and was sensitive to simulation in subsequent studies, these studies did not replicate the original finding of incremental validity (Elhai et al. 2004; Marshall and Bagby 2006; Whitney et al. 2008). Frueh et al. (2005) reported that F_{ptsd} failed to discriminate groups differing in level of actual combat exposure despite claiming PTSD. The F_{ptsd} 's method of construction (empirical identification of items that distinguished veterans asked to simulate PTSD with actual veterans claiming to suffer PTSD but not assessed for malingering) is problematic, and use of this score is not recommended, particularly with civilians.

Several specialized new scales and indices have also appeared. Gervais and colleagues empirically keyed items with SVT failure as the criterion. Referred to as the *Response Bias Scale* (RBS), there has been one published replication of the original scale (Nelson et al. 2007a) and a shortened revised scale developed through multiple regression (Gervais et al. 2007). A substantial correlation with failure on multiple SVTs has been observed ($r=0.49$; Gervais et al. 2007), potentially providing corroboration of SVT results through a separate modality. A recent study found it more sensitive to memory complaints than other MMPI-2 response bias scales, yet uncorrelated with actual memory performance on the California Verbal Learning Test (Gervais et al. 2008). Henry et al. (2006) developed a 15-item scale designed to be highly sensitive and specific to somatic malingering—the exaggeration or feigning of somatic problems frequently found among personal injury claimants (Lees-Haley 1997a, b) and former Pacific theater POWs (Goldstein et al. 1987). Whitney et al. (2008) found the RBS and Henry-Heilbronner Index (HHI) the best MMPI-2 scales for predicting TOMM failure and found FBS relatively ineffective, in contrast to other studies.

Eakin et al. (2006) reported that the MMPI-2's most established validity scales (F, Fb, F(p), Ds) were significantly elevated by PTSD symptoms in college students (Cohen's d s=0.71–1.40) but also showed substantial effect sizes between feigning and PTSD groups (d s=0.84–1.04). Although the students were not screened for exaggeration or external motivation, they are a more credible group than those with obvious compensation motives. On the other hand, they probably represent the lower end of symptom

Table 1 MMPI-2 validity scores for clinical settings adapted from Greene (2008)

Scale	Percentile			
	90	95	98	99
F	15 ^a	19 ^b	25 ^c	29 ^d
F-K	5	10	17	21
Fb	12 ^e	17 ^f	2 ^g	24 ^h
F(p)	4 ⁱ	5 ^j	7 ^k	9 ^l
FBS	28	31	34	35
Ds	26	30	35	38

T-scores for Fb were calculated from means and SDs given in Friedman et al. (2001)

^a *T*-score equivalent=82 (males), 89 (females)

^b *T*-score equivalent=95 (males), 103 (females)

^c *T*-score equivalent=113 (males), 120+ (females)

^d *T*-score equivalent=120+ (males), 120++ (females)

^e *T*-score equivalent=91.6 (males), 89.0 (females)

^f *T*-score equivalent=112 (males), 108.4 (females)

^g *T*-score equivalent=128.4 (males), 123.9 (females)

^h *T*-score equivalent=140.7 (males), 135.5 (females)

ⁱ *T*-score equivalent=70.6 (males), 73.6 (females)

^j *T*-score equivalent=77.8 (males), 81.5 (females)

^k *T*-score equivalent=92.3 (males), 97.4 (females)

^l *T*-score equivalent=106.8 (males), 105.3 (females)

severity of those that meet PTSD criteria. The authors examined optimal cutoff scores and reported diagnostic statistics for each, shown in Table 2. They are presented for expositional purposes only since the PTSD subjects were not actual patients or claimants, and specificities are much lower than desirable. Nonetheless, they are congruent with those of Lees-Haley et al. (1991) and Lees-Haley (1992).

Given the dearth of well-designed studies and widely varying cutoff scores recommended across studies, no firm empirically based recommendation can be made for cutoff scores specifically for PTSD for most validity scales, other than to set them for high specificity as suggested by the Caldwell norms. Table 3 displays mean scores on MMPI-2 validity indices across several studies for PTSD subjects and feigners. It reports only studies that used a PTSD sample at relatively low risk for feigning (Elhai et al. 2001; Eakin et al. 2006) or used potent validity indicators (Lees-Haley 1992; Greiffenstein et al. 2004). Nonetheless, it is apparent that PTSD subjects and simulators in Elhai et al. (2001) produced much higher scores than either student subjects diagnosed with PTSD (but not seeking treatment) or claimants in the Greiffenstein et al. study. A recent study (Geraerts et al. 2006a) found that childhood sexual abuse survivors recruited from the community did not show exaggeration on two symptom validity measures, so the Elhai et al. patient data may reflect self-selection of a more symptomatic sample or one with potential secondary gain. The high score of simulators may be due the instructional set, which did not include warning of validity scales and detection strategies.

The MMPI-2 RF (RF; Ben-Porath and Tellegan 2008) incorporates the MMPI-2's recently introduced Restructured Clinical Scales while dropping most of the familiar clinical scales. It includes parallels to the traditional validity scales and the FBS, as well as a completely new index, the Infrequent Somatic Response scale (Fs). Validity scales on

the MMPI-2 RF no longer suffer from significant item overlap with each other, and the original F scale has been refined so that all items are truly infrequent responses in the normative population (Y. Ben-Porath, personal communication, December 15, 2008). Reduced versions of the FBS, RBS, and HHI will all be scorable on the RF, although the official venter does not score the latter two. No effort was made to retain items for the Ds/Ds-r scale because of its high overlap with other validity scales (Ben-Porath, December 14, personal communication) and the four controversial items on the original F(p) scale, which overlapped with L, have been dropped for the new F(p)-r. Nonetheless, all revised validity scales correlate in excess of 0.90 with the originals, most in the high 0.90s, with norms derived from the MMPI-2 database (Y. Ben-Porath, personal communication, December 15, 2008). Wygant et al. (2008) presented evidence that the new scales correlate with cognitive SVTs and with response-style measures on Multifactor Health Inventory *Extreme Physical Symptoms* scale and the Pain Coping Inventory *Symptom Magnification* scale. Several researchers have presented data that the F(p)-r substantially outperforms the original F(p) scale (Pivovarova and Frederick 2009; Toomey et al. 2009; Tyner 2009). In sum, the MMPI-2 RF validity scales represent a potential advancement and a promising new scale for compensation settings (Fs), but the primary source of validity evidence is extrapolation from the original scales. While the conceptual leap is not large, at present, there is no peer-reviewed data on their comparability.

The Personality Assessment Inventory The PAI (Morey 1991) has three major faking bad indices, the Negative Impression scale (NIM), the Malingering Index (MI; a constellation of eight unusual profile features associated with feigning), and the Rogers Discriminant Function (RDF). Morey (1996) recommended a cutoff score of $T \geq 77$ for NIM and that an MI of 3 suggests concern about exaggeration, while scores of 5 or more are rarely found in legitimate clinical samples. Sellbom and Bagby (2008) exhaustively reviewed PAI validity scale studies, noting 15 publications with both simulation and known group designs. Although described as the best-researched scales among multiscale inventories, the authors deemed all indices best used as screening measures. The authors suggested that, in forensic contexts, the standard cutoff score for NIM ($T \geq 77$) should give suspicion of malingering, while $NIM \geq 110$ or $MI \geq 5$ is very likely due to malingering but will miss most feigners due to low sensitivity. Sellbom and Bagby concluded that the three PAI exaggeration indices had excellent discrimination in simulation groups. NIM and MI also performed well in three known group studies, but the RDF was completely ineffective in the one known group study that examined it (see Table 4).

Table 2 Diagnostic statistics for optimal cutoff scores in Eakin et al. (2006)

Index	Sens	Spec	PPP	NPP	OCA
F>11 ^a	0.71	0.70	0.74	0.67	0.71
Fb>7 ^b	0.79	0.65	0.74	0.71	0.73
F(p)>4 ^c	0.52	0.87	0.83	0.59	0.67
Ds>25 ^d	0.66	0.74	0.76	0.63	0.69

base rate=0.56

Sens sensitivity, Spec specificity, PPP positive predictive power, NPP negative predictive power, OCA overall classification accuracy

^a Equivalent to a *T*-score of 70 for males, 75 for females

^b Equivalent to a *T*-score of 71.1 for males, 69.6 for females

^c Equivalent to a *T*-score of 70.6 for males, 73.6 for females

^d Equivalent to a *T*-score of 78.0 for males, 75.6 for females

Table 3 MMPI-2 validity scores for PTSD subjects and feigners in known groups or in studies seemingly at low risk for exaggeration

Study	PTSD			Feigners			
	Elhai et al. (2001) ^a	Greiffenstein et al. (2004)	Eakin et al. (2006)	Elhai et al. (2001)	Lees-Haley (1992)	Greiffenstein et al. (2004)	Eakin et al. (2006)
Study type	Sim	BS	Sim	Sim	KG ^b	BS	Sim
F	89.9 (21.9)	69.4 (20.4)	72.5 (19.7)	111.6 (17.7)	76.9 (17.9)	75.3 (19.3)	91.1 (22.4)
F-K	5.8 (89.0)	-3.8 (11.2)	-	24.2 (16.1)	2.9 (8.6)	-3.3 (12.3)	-
Fb	-	-	70.2 (24.3)	-	-	-	95.5 (25.0)
F(p)	3.1 ^c	2.3 (2.2)	2.7 ^d	7.7 ^e	-	2.8 (2.2)	5.0 ^f
FBS	21.2 M ^g 23.3 F	20.2 (5.9)	-	23.6 M ^g 26.8 F	27.2 (5.2)	29.1 (5.0)	-
Ds ^h	82.2	-	72.2 (16.0)	100.0 (18.3)	-	-	86.6 (18.2)
ES	-	-	-	-	19.4 ⁱ (10.6)	-	-

Lees-Haley did not have a well-defined PTSD group. Authors vary in reporting of *T*-scores or raw scores for some scales, such as F(p), FBS, and Ds. In order to facilitate comparison, I have converted these into the form most frequently discussed in the literature. Original data in *T*-score form are provided below

BS bootstrap design, KG known group, Sim simulation study

^a This study employed adult sexual abuse patients, not screened for external incentives

^b Subjects in this study did not meet PTSD criterion A

^c *T*-score=63.78 (SD=19.6)

^d *T*-score=60.1 (SD=14.4)

^e *T*-score=97.4 (SD=26.6)

^f *T*-score=77.8 (SD=21.3)

^g These raw scores for men and women, respectively, were estimated from the average *T*-score reported in Elhai et al. Original *T*-scores were 81.59 (SD=14.46) for student feigners and 74.68 (SD=15.25) for PTSD subjects

^h Raw score values for this row are 27.2, 21.9, 36.6, and 29.6, respectively

ⁱ The score in this cell is a *T*-score

Guriel and Fremouw's (2003) review noted only two studies examining malingering PTSD on the PAI (Calhoun et al. 2000; Liljequist et al. 1998) and concluded that empirical data showed that the PAI's response-style indicators did not equal the discrimination of corresponding MMPI-2 scales. However, in both of these studies, the patient groups were veterans unscreened for possible feigning. In Liljequist et al., this author calculated a Cohen's *d* of 1.05 between the simulator and PTSD groups for NIM, indicating relatively good discrimination despite the potentially contaminated patient group. This is not true of Calhoun et al., however, and the high and anomalous false-positive rates for various indices are likely due to substantial feigning rates in

Table 4 Effect sizes for PAI validity indices in Sellbom and Bagby (2008)

	Simulation		Known group	
	<i>d</i>	<i>N</i> studies	<i>d</i>	<i>N</i> studies
NIM	1.30	6	1.47	3
MI	1.21	5	1.00	3
RDF	1.84	5	-0.09	1

the PTSD groups (see Table 5).³ Another study (Scragg et al. 2000) was better designed and apparently overlooked by the Guriel and Fremouw review. It utilized a simulation design with nonstudents and clinical patients not in current litigation. Of 25 employees asked to feign PTSD, only 11 produced a prototypical PTSD profile on the PAI (defined as having a diagnosis of PTSD offered by the interpretive software), and six of these subjects had scores of greater than 85 on the NIM, whereas no subject in the clinical comparison group did so. RDF and MI were also effective. The three measures produced moderate sensitivities of 0.45–0.63, with very high specificities (0.94–1.00). Guriel-Tennant and Fremouw (2006) found that NIM and MI were fairly sensitive to feigning of PTSD by naïve subjects but much less so for subjects coached for only 15 min. Those with presumably honest PTSD symptoms were no more effective at feigning than those without. No actual PTSD subjects were included, and there was a minimal award offered for successful feigning. A dissertation (Eakin 2004) found that

³ I include all available studies in this table, as one of two simulation studies that used unscreened veterans as their genuine PTSD sample reported statistics comparable to the better-controlled studies.

Table 5 Effect size and diagnostic statistics for PTSD feigning on PAI validity indicators

Scale or index	Study	Simulator group	Comparison group	No compensation/ screened?	Cautioned to avoid detection?	Sens	FPR	LR	Cohen's <i>d</i>
NIM	Liljequist et al. (1998)	Students	Veterans diagnosed with PTSD and AD	N, HR	N	–	–	–	1.05
	Eakin (2004)	Students with recent trauma, neg on PCL (naïve)	Students with PTSD ^a	N, LR	Y	–	–	–	1.51
		Students with recent trauma, neg on PCL (coached)	Students with PTSD ^a	N, LR	Y	–	–	–	1.01
NIM > 70	Eakin et al. (2006)	Students with recent trauma, neg on PCL	Students with PTSD ^a	N, LR	Y	–	–	–	0.19
	Guriel-Tennant and Fremouw (2006)	Students with trauma hx, not PTSD (naïve)	–	N, LR	N	0.70	–	–	–
		Students with trauma hx, no PTSD (Coached)	–	N, LR	N	0.43	–	–	–
NIM ≥ 73	Calhoun et al. (2000)	Students	Inpatient veterans	N, HR	Y	0.83	0.65	1.3	–
NIM ≥ 84	Calhoun et al. (2000)	Students	Inpatient veterans	N, HR	Y	0.74	0.57	1.3	–
NIM > 84	Scragg et al. (2000)	Employees	PTSD patients ^b	Y	Y	0.55	0.00	Undef	–
NIM ≥ 92	Calhoun et al. (2000)	Students	Inpatient veterans	N, HR	Y	0.61	0.35	1.7	–
MI	Eakin (2004)	Students with trauma hx, no PTSD (naïve)	Students with PTSD ^a	N, LR	Y	–	–	–	1.51
		Students with trauma Hx (coached)	Students with PTSD ^a	N, LR	Y	–	–	–	1.01
	Eakin et al. (2006)	Students with recent trauma, neg on PCL	Students with PTSD ^a	N, LR	Y	–	–	–	0.64
MI ≥ 3	Scragg et al. (2000)	Employees	PTSD patients	Y	P ^b	0.45	0.06	7.5	–
	Liljequist et al. (1998)	Students	Veterans diagnosed with PTSD and AD	N, HR	N	0.59	0.08	7.4	–
MI ≥ 4	Liljequist et al. (1998)	Students	Veterans diagnosed with PTSD and AD	N, HR	N	0.45	0.00	Undef	–
MI ≥ 5	Calhoun et al. (2000)	Students	Inpatient veterans	N, HR	Y	0.61	0.30 ^c	2.0	–
RDF	Eakin (2004)	Students with trauma Hx (Naïve)	Students with PTSD ^a	N	Y	–	–	–	0.63
		Students with trauma Hx (Coached)	Students with PTSD ^a	N	Y	–	–	–	0.69
	Eakin et al. (2006)	Students with recent trauma, neg on PCL	Students with PTSD ^a	N	Y	–	–	–	0.36
RDF > 0.06	Scragg et al. (2000)	Employees	PTSD patients ^b	Y	Y	0.70	0.05	14.0	–

LR low risk of exaggeration in PTSD group, HR high risk, PCL PTSD checklist, AD alcohol dependence

^a These subjects scored above the designated cutoff score on the PCL and CAPS

^b Patients were not in litigation

^c This anomalous finding is probably due to substantial feigning in the PTSD group

the PAI validity scales and several clinical scales effectively discriminated students with PTSD symptoms from those who were trauma-exposed but without PTSD symptoms. However, a subsequent publication by Eakin et al. (2006) found the PAI validity scales and indices strongly affected by actual PTSD, with no significant differences between actual PTSD students and feigners. In sum, the evidence for the PAI validity indicators is mixed, even in better-designed studies, and actual PTSD patients are likely to produce

elevations on them. Table 6 summarizes average scores observed in presumed genuine PTSD and feigning groups in better-designed studies. The primary advantage of the PAI is that the MI (and potentially the RDF) adds a detection strategy not found on other instruments.

The Millon Multiaxial Clinical Inventory III (Millon et al. 1997) Authors have long questioned the validity scales of the Millon Multiaxial Clinical Inventory III (MCMI-III)

Table 6 Means and SDs of apparently genuine PTSD subjects and simulators on the PAI in better-designed studies

Study	Presumed genuine PTSD			Simulators					
	Scragg et al. (2000)	Eakin (2004) ^a	Eakin et al. (2006) ^a	Naive			Coached		
	Scragg et al. (2000)	Eakin (2004) ^a	Eakin et al. (2006) ^a	Scragg et al. (2000)	Eakin (2004)	Guriel-Tennant and Fremouw (2006)	Eakin (2004)	Guriel-Tennant and Fremouw (2006)	Eakin et al. (2006)
NIM	61.4 (11.9)	58.6 (14.4)	61.5 (12.8)	82.2 (21.4)	84.4 (19.5)	87.3	76.9 (21.1)	66.4	63.9 (12.8)
MI	–	54.4 (8.9)	59.0 (15.3)	–	82.8 (25.0)	82.4	70.7 (19.4)	63.3	59.6 (12.0)
RDF	–	52.6 (8.9)	52.7 (11.1)	–	61.3 (17.3)	–	60.3 (13.0)	–	56.7 (11.5)

All studies in this table are simulation studies, so that scores for feigners do not reflect actual claimants. All were warned about avoiding detection or presence of validity scales. Values for Guriel-Tennant and Fremouw (2006) are collapsed across trauma-positive and trauma-negative groups

^a Subjects in these studies were students diagnosed with PTSD, with an unknown number in treatment. Scores reported may be lower than for populations in treatment or legitimate claimants with PTSD

with PTSD claimants (Lees-Haley 1992). Berry and Schipper (2007) concluded, “The MCMI-III does not appear to be a viable choice for detecting feigned psychiatric symptoms” (p. 257), citing problems with the initial construction, only two published studies (Daubert and Metzler 2000; Schoenberg et al. 2003), modest sensitivity and specificity, and lack of known group studies. Sellbom and Bagby (2008) were even more emphatic, stating, “Under no circumstances should practitioners use this instrument in forensic evaluations to determine response style” (p. 205).

The Structured Interview for Reported Symptoms Many have considered the SIRS (Rogers 1992) as the gold standard for assessing malingered psychopathology, but at the time of the Guriel and Fremouw review, there were no published studies utilizing actual PTSD patients or claimants. The only relevant study (Rogers et al. 1992) involved a number of PTSD simulators and found that, although some scales (*blatant symptoms*, *subtle symptoms*, *selectivity of symptoms*, *severity of symptoms*) effectively distinguished feigners from a mixed group of inpatients, several did not. Since this paper, there has been a dissertation (Eakin 2004) and two publications (Freeman et al. 2008; Rogers et al. 2008) that examined the SIRS with compensation-seeking samples.

Eakin (2004) improved on many previous studies by requiring that all subjects, both simulators and controls, actually experienced one or more trauma. Actual PTSD and trauma-only (non-PTSD) groups were established by clinical interviews and the CAPS. Subjects that did not report PTSD symptoms were assigned to either naïve or coached groups. Naïve test takers were warned about validity scales on the Personality Assessment Inventory and received both this test and the SIRS with no further instruction. Coached fakers received a lecture and a list of PTSD symptoms, were warned

of the PAI’s validity scales, and given a chance to ask questions. All simulators were eligible for a cash bonus for successful feigning. Presumed genuine PTSD students showed somewhat higher scores on *subtle symptoms* and *selectivity of symptoms* scales than control subjects, but not on the other SIRS primary scales. Naïve feigners scored significantly higher than PTSD subjects on all eight primary SIRS scales and four supplemental ones, while coached feigners did so on all of the primary scales but *improbable and absurd symptoms* and *symptom combinations*. The average score for naïve feigners was in the *definite feigning* range for *selectivity of symptoms* and *severity of symptoms*, while for coached feigners these scores fell in the *probable feigning* range. For naïve feigners, the average score for *subtle symptoms* fell in the *probable feigning* range. The average score for genuine PTSD subjects approached the *probable feigning* range on *selectivity of symptoms*, while scores for *subtle symptoms* and *severity of symptoms* were in the *indeterminate* range. The author did not apply the standard SIRS decision rules or report diagnostic statistics or effect sizes. This author calculated effect sizes for individual scales and reports them in Table 7, along with findings from Rogers et al. (2008; discussed below).

Freeman et al. (2008) administered the SIRS to 74 chronic PTSD patients but did not provide any external criterion with which to validate SIRS classifications. While the authors reported total SIRS scores for *clearly exaggerated* and *apparently honest* groups based on the SIRS’s primary two decision rules, they calculated the total score incorrectly, summing the eight primary scales (p. 376).

Rogers et al. (2008) carried out the most extensive study to date on the use of the SIRS with personal injury and disability claimants using a bootstrapping design. Most of the 569 subjects carried multiple diagnoses, but only 7.3%

Table 7 Mean scores for PTSD and simulator groups and Cohen's *d* for the SIRS' eight primary scales

SIRS scale	Group means and SDs				Cohen's <i>d</i> (feigning vs. honest responders)			
	PTSD ^a	Simulator groups			Eakin (2004)		Rogers et al. (2008)	
	Eakin (2004)	Rogers (1992), naïve	Eakin (2004), naïve	Eakin (2004), coached	Naive	Coached	Feigned mental disorder	Feigned cognitive disorder
Rare symptoms	0.7 (1.3)	5.47 ^c	4.4 ^b (4.2)	3.4 ^b (3.0)	1.19	1.17	1.47	0.79
Symptom combinations	0.7 (1.1)	6.00 ^b	3.9 ^b (4.0)	2.3 (1.8)	1.09	1.07	1.82	0.98
Improbable and absurd symptoms	0.2 (0.6)	3.43 ^b	2.4 (2.2)	1.0 (1.3)	0.92	0.79	1.51	1.06
Blatant symptoms	2.0 (2.2)	14.07 ^c	9.3 ^b (7.6)	6.6 ^b (4.5)	1.34	1.30	2.75	1.31
Subtle symptoms	9.9 ^b (6.5)	18.07 ^c	17.9 ^c (6.0)	15.7 ^c (6.7)	1.28	0.82	1.86	1.34
Selectivity of symptoms	16.7 ^b (10.0)	18.00 ^c	33.6 ^d (11.3)	28.5 ^c (11.5)	1.58	1.70	2.03	1.26
Severity of symptoms	4.9 ^b (8.3)	14.13 ^c	19.3 ^d (15.2)	19.3 ^d (15.2)	1.18	1.09	2.40	1.41
Reported vs. observed symptoms	2.2 (1.6)	4.93 ^b	6.1 ^b (3.8)	5.9 ^b (3.0)	1.34	1.54	1.70	0.92
Total	–	–	–	–	–	–	2.42	1.31

Scores for subjects of Rogers et al. (2008) were not provided because most were not diagnosed with PTSD

^a Presumed genuine in apparent absence of external incentive and student status

^b Indeterminate range on SIRS

^c Probable feigning range

^d Definite feigning range

^e Score is borderline between *indeterminate* and *probable feigning*

carried a diagnosis of PTSD. Subjects were classified by MMPI-2 validity scales (F(p) and Ds) or cognitive symptom validity test measures (Victoria Symptom Validity Scale, Test of Memory Malinger, Letter Memory Test) into genuine, indeterminate, and feigning groups. Presumed genuine groups (psychiatric or cognitively intact) had no indications, even borderline ones, of possible exaggeration in each respective domain. The eight primary SIRS scales correlated highly with group assignment for feigning mental disorder and performed surprisingly well at predicting group membership regarding cognitive effort. Very large effect sizes (*M* Cohen's *d*=1.94) were observed between subjects judged to be feigning mental disorder and those who were classified as valid responders on both psychiatric and cognitive SVT measures. Table 7 shows effect sizes observed for each scale for the two feigning groups relative to the claimants who passed all psychiatric and cognitive validity indicators.

The authors performed analyses to show that different diagnoses (depression, anxiety, PTSD) showed minimal differences on the SIRS scales. However, none of the feigning mental illness claimants were diagnosed with PTSD, possibly because clinicians were not blind to the results of the SIRS or MMPI-2. However, test scores alone assigned subjects to criterion groups.

Because the authors deemed bootstrapping studies as inappropriate for setting cutoff scores, they did not report

diagnostic statistics. Despite the high Cohen's *ds* reported, the average scores for those judged to be definitely feigning do not quite result in a judgment of feigning on the SIRS's primary decision rules: no mean scores are in the *definite feigning* range and scores for only two (*subtle symptoms* and *severity of symptoms*) fall in the *probable feigning* range (although the score for *selectivity of symptoms* very nearly does). Together, these observations suggest the sensitivity of the SIRS's first two decisions rules is 0.50 or less (since the mean of third score is slightly below the cutoff for *probable feigning*).⁴ However, the SIRS has a tertiary rule, to be invoked when the first two conditions are not met (Rogers et al. 1992): select items are summed and, if the total exceeds 76, feigning is suggested (this score is referred to as "total" in Table 7). In the Rogers et al. (2008) data, the average total SIRS score of 99.41 (SD=23.81) well exceeded the recommended cutoff score, and based on normal curve percentages suggests a sensitivity for this rule of about 0.84 and a false-positive rate of about 9.7%. Use of the other two SIRS

⁴ Although Rogers described his study as a bootstrapping design, the measures used are well validated; conservative cutoff scores were used, and only extreme groups were compared. This is a strong design comparable to, for example, classification based on the SIRS, which Rogers has cited as an example of a known group design.

rules would likely result in lower false positives at the cost of considerably less sensitivity.

Two methodological features may have bolstered the SIRS's d values over those for other instruments. First, Eakin (2004) warned subjects about validity scales on the PAI but not the SIRS. Secondly, the use by Rogers et al. (2008) of an indeterminate group allowed the creation of purer criterion groups than in other studies, and this necessarily excluded a substantial portion of the sample. The result is higher d values than would be observed if the sample were cut in two, as in most studies. This is not a research flaw (Rogers, 1997b, 2008a), but some correction for the portion of the sample excluded is probably needed to ensure comparability across studies.

Screening Measures The author's position is that screening measures are inappropriate when the differentiation of legitimate from feigned presentations is a primary goal, although they may be very helpful in other settings.⁵ Since Smith (2008) recently provided comprehensive reviews of two of the best contenders, the *Structured Inventory of Malingered Symptoms* (SIMS; Widows and Smith 2005) and *Miller Forensic Assessment of Symptoms Test* (M-FAST, Miller 2001), these two instruments are covered but briefly. The SIMS is a brief (75-item) test that assesses exaggerated endorsement of symptoms in a variety of areas, including depression, psychosis, and mental deficiency. There are several studies that support its validity and substantial discriminating power (Cohen's $d=1.13$ – 3.52 ; Smith 2008) and one demonstrating resistance to coaching (Jelicic et al. 2007), but none with PTSD patients. There is also concern that published cutoff scores produce excessive false-positive results among clinical control subjects: 60% in Poythress et al. (2001), 39% in Lewis et al. (2002), and 33–40% in a thesis cited by Smith (2008).

The M-FAST, a brief structured interview used for screening purposes, has shown good discrimination in several populations (Smith 2008), including PTSD. Guriel et al. (2004) found that the standard M-FAST cutoff score of ≥ 6 accurately detected 68% of student simulators, that coaching simulators about PTSD symptoms actually increased the detection rate (to 87%), and that coaching about escaping detection did not lower sensitivity (0.69). Used together, the M-FAST and TSI accurately detected 90% of simulators. Guriel-Tennant and Fremouw (2006) found that trauma history did not aid PTSD simulation, and

in contrast to the earlier study, coaching decreased detection rates from 0.84 to 0.52. In the only study that examined discrimination of feigners from valid PTSD patients (non-compensation context, passed SIRS and MMPI-2 validity indicators), the M-FAST's standard cutoff score of six generated a sensitivity of 0.63 and specificity of 0.85 separating student simulators (Guy et al. 2006).

The M-FAST is a well-researched and useful instrument, but because it is marketed as a screening test and because the SIRS is longer better-established and the M-FAST has substantial false-negative errors relative to the SIRS (Miller 2001), it would be hard to justify its use over the SIRS, even to rule out feigning in an otherwise highly credible disability case.

Specialized Self-report Measures The TSI (Briere 1995) assesses various symptoms of PTSD and includes three validity scales. The most important of these is the Atypical Responding scale (ATR), which contains items similar to those found on the MMPI-2 scale 8 and F or F(p). Briere (1995) recommended a cutoff score of $T > 65$ for suspected exaggeration, with scores above 90 deemed invalid. While seven studies (Edens et al. 1998; Efendov et al. 2008; Elhai et al. 2005; Guriel et al. 2004; Guriel-Tennant and Fremouw 2006; Rosen et al. 2006; Porter et al. 2007) have examined the TSI's ability to detect feigning of PTSD, there are no true known group studies and all existing studies are seriously limited. Where high effect sizes occurred, they are due to contrasting students responding honestly with simulators (see "Appendix"). When studies compared simulators with patient or presumed honest claimant groups, effect sizes were quite small. The ATR may be modestly effective, but the available research provides scant scientific evidence and little guidance in establishing an appropriate cutoff score or estimating diagnostic statistics for forensic application. The TSI has a sibling in the DAPS (Briere 2001), but its validity scales have so far only been indirectly examined (see Demakis et al. 2008).

Memory and concentration complaints are common among many psychiatric and neuropsychiatric populations, including those claiming PTSD (Sachinvala et al. 2000; Solomon and Mikulincer 2006; Resnick et al. 2008). The Memory Complaints Inventory (Green 2004) is a brief inventory of self-reported memory problems. Domains assessed include self-assessments of autobiographical memory, verbal memory, numerical memory, visuospatial memory, pain interfering with memory, memory interfering with work, and amnesia for antisocial behavior or complex actions. Norms are available for various psychiatric groups, broken down by whether or not the subjects passed the Word Memory Test. Green has argued persuasively for separating scores of subjects who are responding honestly and those that independent evidence shows are not.

⁵ It is unfortunate that the SIMS and M-FAST authors described their measures as screening tests, as this may artificially constrain their use: with higher cutoff scores, both may well provide positive predictive power/specificity statistics comparable to other response-style instruments.

Unfortunately, the program does not report MCI scores for PTSD patients or claimants. High scores on the MCI may occur due to either exaggeration or depression, are not related to performance on neuropsychological tests, but are predictive of SVT failure (Green 2008). As depression often coexists with PTSD and can lead to elevated MCI scores, claimants' scores should be compared with those of depressed patients in the MCI software database.

Cognitive Symptom Validity Tests

Symptom validity tests used in neuropsychology rely on a variety of strategies to detect poor effort or intentional failure. One might ask if the MMPI-2 has effective validity scales, why does one need other tests? The answer is, of course, that "effective" is a matter of degree. While the MMPI-2's validity scales rack up impressive numbers in general psychiatric samples when simulators fake psychosis, discriminating the sort of distress typically seen in a disability or law suit is more challenging. Nearly all studies report a Cohen's d of 0.8–1.2 for this task: a substantial effect size but one that separates the groups by only a single SD. For the SIRS, effects sizes *may* be larger, but in order to keep false-positive errors low, cutoff scores are set high at the expense of sensitivity, so the SIRS will also miss a substantial portion of feigners. In contrast, Green (2005) reported sensitivity of 0.97 or better in simulator groups and nearly perfect specificity with nondemented populations. Such figures suggest a Cohen's d in the area of 4.0. Comparing PTSD subjects in the WMT scoring program database that passed or failed the WMT yields Cohen's d s of 1.60–3.94 for the three major effort indices. Although these figures are contaminated with the criterion, they also would tend to underestimate performance of the test as a whole, which relies jointly on the three measures. Empirically, Larrabee (2003) found that the combination of SVT failure and scoring above the cutoff on FBS was the most potent combination of response-style indicators examined. Failure on a performance-based validity test can corroborate feigning in a modality distinct from self-report, in a mode that probably requires intentional poor performance (consciously not attending to the task, purposely answering incorrectly), and weighs against interpreting elevations on self-report validity scales as benign overreporting. Failure on both types of tests is considerably stronger evidence of feigning than evidence based on self-report alone, and performance on cognitive SVTs can shed light on moderately elevated self-report validity indices.

Below-chance performance has traditionally been the gold standard for attribution of malingering but typically results in identification of less than 10% or less of feigners (Greve et al. 2009). The TOMM allows scoring for both

below-chance performance and norm-referenced criteria. TOMM performance is not significantly affected by anxiety (Ashendorf et al. 2004; O'Bryant et al. 2007), depression (Ashendorf et al. 2004; Iverson et al. 2007; Rees et al. 2001; Rohling et al. 2002; Yanez et al. 2006), psychosis (Duncan 2005), and pain (Etherton et al. 2005; Iverson et al. 2007), but it is affected by dementia (Teichner and Wagner 2004) and perhaps mild retardation (Hurley and Deal 2006; but see Simon 2007). There is no means to distinguish feigning from true cognitive impairment from TOMM test data alone.

The WMT (Green 2005) includes several memory tasks and three primary embedded validity indicators. It has been the subject of fewer formal investigations in psychiatric samples than the TOMM but has been extensively researched by a number of independent researchers (Bauer et al. 2007; Brockhaus and Merten 2004; Gervais et al. 2004; Greve et al. 2008; Morel 2008; Sullivan et al. 2007). The author has presented considerable data that the effort indicators are relatively unaffected by most neurological or psychological conditions short of dementia, and even very impaired groups, such as mentally retarded adults and fetal alcohol syndrome children, are usually able to pass them (Green 2005). The author recommends subjects have at least a third-grade reading level. Even highly sophisticated simulators such as psychologists and neurologists were unable to produce impaired but credible profiles nor were patients who previously passed the WMT and given an incentive for failing believably during a second administration (Green 2005). Because the WMT includes multiple actual measures of memory performance and these have both a theoretical and empirical hierarchy of performance, WMT profile information provides an internal check of validity, a feature lauded by Rogers (2008b). Lastly and what truly distinguishes the WMT is its sensitivity. While many SVTs have sensitivities between 0.30 and 0.60, the WMT achieves nearly 100% sensitivity among simulator groups (Green 2005) and has generated failure rates nearly three times as high as the TOMM in populations of suspected malingerers in some disability contexts (Demakis et al. 2008; Gervais et al. 2004; see also Bauer et al. 2007 and Green 2007). Importantly, these failures are corroborated by failure on other effort tests, external data, or retesting with incentives to improve performance (Green 2007). Other established cognitive SVTs have sensitivities of only from 0.29 (TOMM) to 0.62 (Medical Symptom Validity Test; Green 2004) relative to the WMT (Green 2007). Hartman (2002) proposed eight criteria for SVTs used with head injury patients and asserted that, among current SVTs, only the WMT met all of them. However, some of these (e.g., has been evaluated, known error rate) are not currently met in the published literature for PTSD examinees. There is less published research on the WMT

with psychiatric groups than the TOMM, although several groups are included among the normative groups in the software program. In a study of schizophrenics, 72% failed the WMT (Gorissen et al. 2005). However, these subjects are not necessarily false positives, as schizophrenics are notorious for their lack of motivation, and in the Gorissen study, WMT failure correlated with independent assessments of negative symptoms.

Greve et al. (2008) reported that their study was the first to test the WMT in a known group design, although the design better fits Rogers' (1997b, 2008c) *bootstrapping* category. The authors reported that, although the WMT produced a higher sensitivity than the TOMM, it also produced a large false-positive rate (24%) among subjects that included moderate head injuries. Contrary to previous claims, the two instruments showed comparable discriminating power. However, P. Green (personal communication, March 24, 2009) criticized this study for using validity measures that are inferior to the WMT in creating the criterion groups, failure to rule out external incentives, equating effort testing with malingering, and for not using the WMT's profile validity checks.

Goldberg et al. (2007) surveyed the existing literature for the effects of psychiatric illness on all SVTs. They concluded that there is no evidence that depression, regardless of type or severity, affected any of 12 different validity tests and indicators. They noted that the effects of some conditions, such as bipolar illness, personality disorder, and PTSD, have not been investigated and that data regarding anxiety and somatoform conditions are limited. Despite the lack of direct examination, there is no reason to believe that PTSD, without significant neurological involvement (e.g., moderate head injury or worse), should present difficulty passing the TOMM or WMT if adequate effort is applied since (a) even groups with moderately severe cognitive impairment are able to do so and, (b) when cognitive deficits are observed in PTSD, they tend to be small and do not extend to recognition memory.

While the TOMM is well researched, insensitive to mood, pain, and psychiatric disorder, and widely accepted among ABBP forensic diplomates (Lally 2003), a recent survey suggests a rapid rise in the WMT's use and acceptance among neuropsychologists (Sharland and Gfeller 2007). The WMT's author has also developed both a simpler version called the Medical Symptom Validity Test (MSVT; Green 2004) and a nonverbal SVT (Nonverbal Medical Symptom Validity Test; Green 2007), both of which share the WMT's positive features. Neither appears to match the WMT's exquisite sensitivity although they do exceed the TOMM (Green 2006, 2007). The MSVT has several publications to its credit (Blaskewitz et al. 2008; Carone 2008; Howe and Loring 2008; Stevens et al. 2008) and offers an additional benefit over potential competitors: its subtests closely parallel the WMT's but are considerably

easier. Performance levels on the MSVT that are lower than comparable scales on the WMT would be highly suspicious. The MSVT lacks the sensitivity and extensive research base of the WMT and because of its relative simplicity, might be more transparent to brighter subjects. While the MSVT does not currently rival the WMT for choice of a single cognitive SVT with adults of near-average intelligence, it would be a good choice for a second one or for cognitively impaired subjects.

Because memory complaints or deficits are not core PTSD symptoms, some claimants may not exaggerate or intentionally fail an SVT based on memory. Thus, passing a cognitive SVT, in the absence of memory complaints, is ambiguous. However, failure is not, and if a subject presents with feigned memory problems, the WMT is quite likely to detect it.

If the examiner employs intelligence testing or neuropsychological batteries, many potential indices of effort are available. A full discussion of these techniques is beyond the scope of this paper, and Larrabee (2007a) and Boone (2007a) provide recent and comprehensive treatment of these topics. Over the past decade, levels of performance on many neuropsychological tests associated with failed SVTs have been documented and can be used in assessing the credibility of the test profile. The Meyers Neuropsychological Battery has ten built-in validity indicators. Other multifaceted tests, like the WAIS and WMS, have had such indices developed by researchers outside of the publishing company. However, these have often been developed and validated on rather specific groups (i.e., mild head injury claimants), using discriminant analysis or other methods that may result in overfitting or misfitting for different conditions, and none have been specifically validated for PTSD. Several simple indices are likely to be more robust across conditions and have been well validated with multiple conditions and have shown resistance to depressed mood or potentially distracting stimuli such as pain but not mental retardation or dementia (Babikian and Boone 2007). These include the age-corrected Digit Span scale score, the difference of the Vocabulary and Digit Span age-corrected scale scores ($V_{ss} - DS_{ss}$), and Reliable Digit Span (the number of digits repeated forward and backward, scored correct for both trials, summed). Table 8 provides diagnostic statistics associated with the various indices and cutoff scores. The authors noted that the various digit span indicators correlate only modestly with other validity indicators, and they are not appropriate for the developmentally disabled or examinees with dementia.

Other Tests of Potential Use

Morel (1998) developed the *Morel Emotional Numbing Test* (MENT; see also Morel and Shepherd 2008a) to distinguish

Table 8 Digit-span-based validity indicators, cutoff scores, and diagnostic statistics from Babikian and Boone's review

Index	Cutoff score	Sensitivity	Specificity
Age-corrected scale score	≤4	0.19–0.32	>0.90
	≤5	0.47	>0.90 ^a
Voc _{ss} – D _{ss}	≥2	0.72	0.79 ^b
	≥5	0.50 ^c	>0.92 ^b
Reliable digit span	≤7	0.50–0.70	0.56–0.81 ^d

^a Except for stroke victims

^b Will be higher among those with moderate to severe head injuries

^c In subjects producing a full-scale IQ of 85 or greater

^d Higher error rates observed with more severe head injuries

real from feigned PTSD and focuses on a core PTSD symptom. Norms are available for legitimate PTSD *claimants*, other psychiatric groups, and for patients identified as probably exaggerating. No subject in the legitimate patient or claimant groups failed the MENT, as opposed to 80% of the noncredible group. A recent study (Morel 2007) found adequate sensitivity (0.64) and excellent specificity (1.00) among claimants that failed the WMT. The MENT was handicapped in this analysis, as only a small portion of this sample was diagnosed or claimed PTSD. Readers should be aware that the MENT used by European researchers (Geraerts et al. 2006a, b) uses color pictures and different faces than the original and is not comparable. Messer and Fremouw (2007) created the MENT-R (revised) when the authors had trouble locating the original instrument, and the MENT-R does not supplant the original. It differs from the original MENT in a number of ways, including number of pictures and number of emotions modeled. Because other researchers used their own variations of the MENT, these publications do not provide direct evidence for the MENT or its published cutoff scores, although they do support the MENT's general rationale. A meta-analysis has been published (Morel and Shepherd 2008b), albeit as a letter to the editor, and the analysis included studies that used the variants of the MENT. Morel and Marshman (2008) argued that the MENT meets all of the criteria articulated by Hartman (2002) for appraising SVTs, including the *Daubert* criteria, although this conclusion may be premature.

True replication by independent researchers is needed but forthcoming: Merten et al. (2009, manuscript submitted for publication) reported that the MENT correlated 0.66 with the WMT among a group of consecutive PTSD claimants and that 41% of this group failed the MENT, compared to 51% failure rate on the WMT. Merten et al. (2009, manuscript submitted for publication) reported that informing simulating subjects about PTSD symptoms actually increased the MENT's sensitivity from 0.70 to

0.95, best of all the measures examined, while warning subjects about the presence of validity scales reduced sensitivity to 0.65–0.75, which was slightly lower than figures for the WMT. Diederick and Merten (2009, manuscript submitted for publication) reported that the MENT demonstrated a sensitivity of 0.41 for analog patients simulating either PTSD or depression. Although they also reported a specificity of 1.00, the control subjects were not actual patients or claimants.

The MENT has unique strengths that warrant possible inclusion in a forensic PTSD assessment: it has demonstrated excellent specificity and good sensitivity even among PTSD plaintiffs, whereas many validity studies utilize PTSD *patients* as the control group. It is a welcome diversification of modalities beyond memory testing and self-report, and it addresses one of the core identified features of PTSD—emotional numbing. A possible limitation of the test is that some subjects may not find the MENT's presented rationale convincing: Messer and Fremouw reported that subjects assigned a mean believability rating of 2.9 (SD=1.06) on a 0–4 scale, where 4 was “very believable,” to their version of the MENT. Coached malingerers assigned a lower rating (2.5, SD=0.9). The study by Merten et al. also suggests that the MENT may be vulnerable to general coaching. Thus, as with other measures discussed, passing the MENT does not rule out exaggeration.

Discussion

PTSD presents challenges to those who would assess claimants and disability applicants. Previous clinical lore has held that extreme elevations on MMPI validity scales are normal for such groups, confounding efforts to assess malingering or exaggeration. The present analysis of validity scores in known group and nonlitigation samples finds consistent, moderate elevations on most validity scales across different measures, with these being higher in treatment-seeking than student samples and higher still in compensation-seeking groups.

State-of-the-art assessment of feigning requires multiple modalities. None of the available methods has a truly satisfactory database for assessment of PTSD, and judgments must rely in part on studies that utilized other patient groups. The SIRS shows good validity of most of its primary scales and covers a broad, but not comprehensive, range of detection strategies. It also allows formal assessment of observed behaviors. Firm decision rules specific to PTSD are not available, but the standard criteria appear appropriate and conservative. The MMPI-2 supplies a number of potentially useful scales, including F, Fb, F(p), FBS, Ds, ES, and the newer Response Bias Scale of Gervais and colleagues, some of which have no direct

parallel on the SIRS. Decision rules for the F family of scales (other than Fp) have varied widely across studies, with authors of known group or strong simulation designs recommending much lower scores than others. The PAI has mixed empirical support for the assessment of malingered PTSD, hamstrung by strong coaching manipulations in several studies and the anomalous findings of Eakin et al. (2006). While the non-PTSD literature supports the validity of the PAI's fake bad indicators, the data for setting cutoffs for PTSD feigning are sparse. No studies have examined the MCMI-III response-style indicators with trauma patients or claimants, and two recent negative reviews weigh strongly against their use in response-style assessment.

Few publications have addressed the use of cognitive symptom validity assessment outside of neuropsychology, which provides a powerful and complementary assessment of response style. Rosen and Powel (2003) was the first published use of a cognitive SVT with a PTSD claimant, while two large recent studies have done so in compensation-seeking samples (Demakis et al. 2008; Rogers et al. 2008). Although most SVTs are focused toward cognitive symptoms, the MENT is not. Because general memory impairment is not a core feature of PTSD, the meaning of a passed cognitive SVT is somewhat ambiguous, even when the SVT has high sensitivity among those simulating cognitive impairment. Some test takers who simulate in other ways may not do so on such a task.

Should SVTs be given when memory or cognitive complaints are not reported during an unstructured interview or part of the claim? Rosen and Powel (2003) and Otto (2008) opined that they should not. Otto asserted it would be inappropriate to administer such measures if the examinee is not claiming cognitive impairment because passing the SVT would not be evidence of honest responding. The present author's view is that the report of memory complaints is heavily dependent on the methods used to inquire about them. An open-ended question about psychiatric or mental symptoms will yield fewer positive responses than will a similar inquiry about memory or concentration problems. A structured inventory that asks about dozens of perceived memory problems will likely produce still more, while a performance test will likely identify some examinees who did not report memory problems on self-report measures. Given the ubiquity of memory complaints in clinical patients (and even the nonclinical public) and the fact that concentration and memory problems are cited in DSM-IV as symptoms of PTSD, I recommend that the Memory Complaints Inventory be given routinely. Whether or not memory problems are reported, in the absence of psychosis, serious cognitive impairment, or internal or external distraction, SVT failure (assuming the test has high specificity) strongly suggests that the test taker did not perform to the best of his or her ability.

Some may consider the use of multiple validity measures, particularly in the cognitive domain, as a fishing expedition. Others may argue that, even if someone fails a cognitive SVT, this has limited implications for a diagnosis of PTSD, which entails primarily psychiatric symptoms. However, DSM-IV-TR states, "Malingering should be ruled out in those situations in which financial remuneration, benefit eligibility, and forensic determination play a role" (p. 427), and defines malingering as the intentional production of false or grossly exaggerated symptoms. There is no prescription that feigned symptoms be limited to one distinct disorder, which would produce absurd outcomes. While performance on cognitive and psychopathology validity tests may constitute separate factors in disability contexts, the factors may show substantial correlations (Nelson et al. 2007b). Further, many researchers have found substantial correlations between self-report scales and SVT performance, including FBS (Larrabee 2003; Nelson et al. 2007a, b; Sellers et al., 2006; Vagnini 2003; Wygant et al. 2007), RBS (Gervais et al. 2007; Nelson et al. 2007a, b; Whitney et al. 2008), F and F(p) (Dearth et al. 2005; Whitney et al. 2008), Ds2 (Dearth et al.), and SIRS scales (Rogers et al. 2008).

Psychologists are obligated to provide examinees with basic information about the evaluation, including the purpose of the evaluation, who is requesting it, the confidentiality or lack thereof, and a general description of the procedures to be used (American Psychological Association 2002). Should examinees be warned that the validity of their responses and presentation will be assessed? Youngjohn et al. (1999) argued that such warnings do not deter benign exaggeration and make feigners more savvy, and this position was supported by Suhr and Gunstad (2007), even though their warning was nonspecific and the validity test was embedded in a neuropsychological battery. Boone (2007a) argued that medical professionals do not offer such warnings and Wetter and Corrigan (1995) asserted that such warnings violate test security. Warning subjects about MMPI-2 validity scales also lowers their effectiveness (Rogers et al. 1993). Slick et al. (2004) found neuropsychologists widely split on the issue, with 54% never warning examinees about the presence of effort or feigning tests, while 37.5% reported they always do. Considering these various factors, Boone (2007a) suggested that a statement in the informed consent emphasize the importance of giving one's best effort at all times and warning that exaggeration "may make the results more problematic to interpret."

In contrast to the forensic literature, much of the clinical literature is sympathetic to the point of advocacy, and very few researchers have attempted a rigorous assessment of external motivation or independent assessment of response style in patient or claimant groups. To the contrary, authors casually labeled compensation-seeking subjects as "genuine PTSD" or "bona fide PTSD," even in studies that purport to

investigate malingering. Gervais et al. (2001) reported that SVT failure is about as frequent in persons who have already obtained disability status as those applying for it. Given this observation, such persons should not be classified as noncompensation subjects but ideally should be analyzed separately from both compensation-seeking and non-compensation-seeking groups. Lastly, financial awards do not exhaust the potential motivations for malingering, which may include continued prescriptions for controlled substances, reduced demand for employment or other duties, or sympathy from others.

Another methodological issue in need of redress is determining the actual effort expended by subjects asked to simulate. While plaintiffs may be eligible for a lifetime of income or a multimillion-dollar award, the typical subject in a simulation study typically has less than a one in four chance at a \$30 prize. Researchers need to be more creative in establishing motivational schemes that more closely approximate real life, which may include negative as well as positive consequences depending on performance (see Rogers 2008a). Regardless of the motivation scheme used, manipulation checks, such as a simple rating scale (see Fig. 1), are needed. Similarly, 15 minutes of coaching and preparation probably does not match the investment of a claimant who has a lifetime of income to gain and access to the internet, local libraries, and friends for months preceding the evaluation. Known group or high-quality bootstrapping designs circumvent these problems, and the addition of cognitive SVTs and the MENT can greatly assist in creating valid criterion groups, as can checklists such as used by Greiffenstein et al. (2004). Rogers (1997b, 2008a) made numerous thoughtful recommendations regarding research design, and few studies to date have come close to meeting them. Methodologically deficient studies will not yield valid or useful data, and I urge journal editors to reject future submissions that do not meet the majority of concerns raised by Rogers.

Considering the substantial base rates of exaggeration in compensation-seeking populations and the DSM-IV-TR's and the National Association of Neuropsychologists' statements on assessment of malingering or response bias, I assert that thorough assessment of presentation validity is the first and most important task of every personal injury or disability examination. Franklin and Thompson (2005) argued that assuming a high base rate biases the examination

and argued against the proposition I have just proposed. However, not to consider response bias, when base rates approach 50%, amounts bias in the other direction. Exaggeration and feigning is a topic examiners must routinely confront and do so competently. A respectable work or personal history or friendly demeanor is not evidence of valid responding. Nearly half of college students surveyed reported they would be willing to feign mental illness to escape criminal responsibility or to recover money in a lawsuit (Iverson 1996). Ford (2005) gave numerous examples of physicians malingering by proxy for their patients' benefit, while Stone and Boone (2007) cited examples of *saints* malingering illness of various sorts. The only condition that would justify the omission of fake bad instruments and scales in a compensation setting would be an absence of symptoms and claimed disability, and this should prompt use of fake good indicators such as the MMPI-2 K scale.

Recommendations for a Response-Style Battery

Having surveyed the available literature, I conclude, like Gurriel and Fremouw (2003), that there is yet no gold standard for assessing feigning of PTSD. While several instruments and indices have strong validity evidence across various psychiatric or neurological conditions, few have more than two well-designed supportive studies specific to PTSD. Nonetheless, this author recommends the SIRS, MMPI-2, Memory Complaints Inventory, the original MENT, and Word Memory Test. Because cutoff scores are now set to achieve high specificity at the cost of sensitivity and the centrality of psychiatric symptoms, use of all three self-report measures is encouraged, despite some duplication in detection strategies.

The SIRS has been studied in both simulator and known group designs and has been shown resistant to coaching. Two well-designed studies (Eakin 2004; Rogers et al. 2008) have produced positive results in assessment of personal injury claimants, although latter contained few subjects with a PTSD diagnosis. The SIRS utilizes eight different strategies of response bias detection and is the most commonly used measure of malingering in forensic evaluations (Archer et al. 2006). Although sensitivity for the two primary decision rules is limited, specificity appears strong and the SIRS scoring protocol allows for an *indeterminate* classification when there are indications of feigning short of the criteria. Practitioners should also score the SIRS total score as detailed in the manual for a more sensitive indicator but should be mindful that it may produce about a 10% false-positive rate.

I also recommend the MMPI-2. The F(p) scale provides a good assessment of rare symptoms, although this strategy is somewhat redundant with the SIRS scales. The FBS,

How hard did you try to follow the instructions you were given in this study?

0	1	2	3	4
I made no attempt at all	I didn't put a lot of thought or effort into it.	I thought about how to best answer on many items	I made a good effort, but not my very best	I did the best I could; I tried really hard

Fig. 1 A brief hypothetical scale to assess motivation in subjects asked to simulate

RBS, and ES scales add additional detection strategies. Validity indicators that function well for feigned psychosis and in criminal populations can be insensitive in personal injury contexts. While strong elevation of any “fake bad” validity scale (>98th percentile on Caldwell’s norms) is evidence of exaggeration or feigning (in the absence of inconsistent responding, confusion, or poor reading ability), the F family of scales will likely have poor sensitivity in more sophisticated subjects and F and Fb may be elevated in severely impaired individuals. FBS performs well in personal injury contexts when methodologically adequate studies are considered, while other scales such as Ds and ES have demonstrated validity, but typically at lower levels than FBS. Some neuropsychologists are already routinely using newer scales like RBS and the HHI, which are generating new research rapidly. While arguably not ready for forensic use yet, when each scale has three methodologically sound published studies by different researchers showing consistent results (Lilienfeld et al. 2000), they will be. Given the current pace of research in this field, this may happen by this article’s publication date. Until then, they might provide supportive evidence for more established scales (e.g., supporting interpretation of exaggeration from an elevated FBS or F(p) scale). Scales such as FBS and RBS are distinct from traditional validity scales, as demonstrated by a recent factor analysis among personal injury claimants (Nelson et al. 2007b). They are more related to malingering of somatic problems and failure on cognitive SVTs than scales such as F and F(p). Lastly, the MMPI-2-RF’s new and refined validity scales will make it a strong alternative to its stable mate if independent research confirms its success at identifying response styles.

The Memory Complaints Inventory offers a way to assess reported memory and concentration symptoms in a standardized manner. Until norms are available for PTSD examinees, I recommend that examiners refer to the norms for major depression, both those that passed and failed the WMT, as well as the respective norms for anxiety disorders.

The MENT has unfortunately mutated into three separate tests. It has demonstrated insensitivity to even the most impairing psychiatric disorders, and its effective cutoff scores have remained consistent across studies. Several currently submitted papers will remove its primary limitation at the time of this writing: the absence of published replication of the original instrument by outside researchers.

The Word Memory Test should also be included in assessment of PTSD response validity. There is no reason to believe that failure on a well-validated SVT would be due to anything other than poor effort in the absence of serious neurological involvement. Even if cognitive problems exist and were previously undetected, profile analysis features of the WMT allow distinction between simulators and those with genuine impairment, and it appears resistant

to faking even by bright and very well-educated subjects. While there have been no published reports of the WMT’s effectiveness with PTSD claimants, it has a broad and impressive literature with various medical and psychiatric groups that undoubtedly have greater cognitive impairment than PTSD patients. Whether or not the subject presents with memory complaints or poor performance on memory tasks, failing the WMT is strong evidence of poor effort and probable intentional failure in the absence of inability to attend to the task, retardation, or dementia. Although the WMT has very high sensitivity to those that simulate cognitive problems, not all malingering PTSD claimants will do so. While use of multiple cognitive SVTs is often recommended, adding the TOMM to the WMT in two databases of 1,315 disability claimants yielded only eight subjects that failed the TOMM but not the WMT. In contrast, 330 failed the WMT but not the TOMM (Green 2007). There are little data available whether the addition of additional SVTs can add to the accuracy of the WMT’s classifications.

While there are a number of useful instruments available, there is room for new approaches. For example, no self-report inventory or structured interview has a validity scale comprising plausible but inaccurate PTSD-like symptoms. There are many details about PTSD symptoms, such as time of onset, duration during a single symptom episode, amount of reported distress, and interference with work or play, that might distinguish actual from malingered PTSD. Checklists of features proposed by Resnick (2003) or used by Greiffenstein et al. (2004) deserve further formal study: the days of unvalidated validity indicators should have receded into pre-*Daubert* history. While collateral sources can provide valuable information, they are potentially subject to deception by the claimant and may have a vested interest in the claim. Although there are a very few inventories of psychiatric symptoms designed to be completed by family members (e.g., the Katz Adjustment Scale; Katz and Warren 1999), they do not contain validity scales. Lastly, psychophysiological assessment may yet have a valuable role to play. While current studies indicate that only about 60% of PTSD patients show the expected psychophysical reactions, this may well be the result of contamination of subject pools by simulating patients.

Integrating findings from multiple instruments and domains can present conceptual challenges. When a claimant either passes or fails all validity measures, there is little question of the interpretation. But what if only cognitive SVTs are failed, and self-report indices are either borderline or within normal limits? Such a pattern could well occur in a prepared or coached examinee. Failure on the TOMM or WMT is strong evidence of invalid presentation and cannot responsibly be ignored. Conversely, failure on self-report and structured interview measures without accompanying SVT failure should not cause question of the former data, given that the

latter may be perceived as irrelevant. As an initial step in integrating various test data, I present some of the more promising response-style indicators across instruments in Table 9, along with suggested cutoff scores, associated false-positive rates, and a preliminary weighting system.

A sizable portion of claimants may receive coaching from their attorneys or engage in their own research prior to the examination (Lees-Haley 1997b). Suhr and Gunstad (2007) found 75% of trial attorneys spent between 25 and 60 min preparing clients for a psychological evaluation and that information about popular validity tests is increasingly available on the Internet. Forty-four percent of students in one study who were asked to simulate brain damage spent more than an hour researching their role, including accessing the Internet, reading books or articles, and talking to friends and even to physicians and psychologists. Thus, examiners need to be cognizant of possibility that one or more instruments in their arsenal may be compromised and may choose to use or have available one or two relatively unknown techniques.

The instruments in this review assess feigning as manifested by exaggeration or fabrication of symptoms. However,

malingering can occur by misattribution of symptoms to a different event or by denial of preexisting psychiatric problems. Collateral sources and review of complete medical and psychiatric records will usually be necessary to address these issues. In addition, as with most forensic interviewing, examiners should begin with open-ended questions, progressing to more direct inquires and structured assessment instruments. While many feigning subjects can easily portray PTSD on rating scales, most have difficulty doing so during open-ended questioning (Burges and McMillan 2001). Interviews should be recorded on audio or audio–video media, consistent with recommendations for preserving forensic data (Committee on Ethical Guidelines for Forensic Psychologists 1991). Formal testing, however, should not be recorded to preserve test integrity (NAN 2003).

Lastly, examiners must decide whether, in the face of exaggeration, they can reliably distinguish malingering from factitious motivations or conscious from unconscious motivation. Hamilton et al. (2008) and P. Green (personal communication, May 22, 2009) have argued that somatoform patients who fail validity tests are doing so consciously and

Table 9 Selected psychometric indicators of feigning

Indicator	Weight	Error rate	Primary source(s)
F(p) 4–6	1	0.04–0.10	Greene (2008)
F(p) ≥ 7	2	0.02	Greene (2008); Rogers et al. (2003)
FBS = 27–28	1	0.05	Greiffenstein et al. (2004)
FBS > 28	2	0.01–0.03	Ben-Porath et al. (2009)
Ds > 30	1	0.05	Greene (2008)
Ds > 35	2	0.02	Greene (2008); Rogers (2003)
SIRS one scale, <i>definite feigning</i> range	1	? ^a	Rogers (1992)
SIRS, three scales in probable feigning range	2	0.005	Rogers (1992)
SIRS Tot ≥ 76	2	0.00	Rogers (1992)
PAI NIM > 84	1 ^b	0.00	Scragg et al. (2000)
PAI MI ≥ 3	1	0.06–0.08	Scragg et al. (2000), Liljequist et al. (1998)
PAI MI ≥ 5	2	0.00	Scragg et al. (2000); Morey (1991)
WMT IR or DR = 83–89 ^{c, d}	1	0.05	Green (2005)
WMT failure ^d	2	? ^e	Green (2005)
Any SVT significantly below chance (any trial or combination of trials)	5	0.05 ^f	Pankratz et al. (1975)
TOMM failure on Trial 2 or retention trial ^d	2	0.00 ^g	Ashendorf et al. (2004)
MENT failure	2	0.00	Morel (1998, 2008)
Digit span ACSS ≤ 5 or V _{ACSS} –DS _{ACSS} ≥ 5 ^d	1	0.10 ^h	Babikian and Boone (2007)

^a Unable to discern from manual, so rule is assigned only one point

^b This score is given only one point, despite the high specificity, because it is based on only one study

^c Range indicated is two SDs or more below the norm for 25 PTSD patients who passed the WMT

^d Where no evidence of dementia, retardation, or moderate to severe head injury

^e No specific data available. Error rate should be miniscule given high specificity among cases of neurological insult

^f Although below-chance performance is often based on the 0.05 alpha level, this is not comparable to the other error rates in this table

^g No patient with severe anxiety or depression scored below the cutoff on the TOMM

^h Figure refers to highest error rate for the two indicators

intentionally. The concept of somatization is a carryover from psychodynamic views that postulated unconscious motivations. A much more parsimonious explanation is that failure on effort tests, in the context of litigation, is a deliberate attempt to maintain a role perceived as more rewarding than that of functional and working adult. Boone (2007b) argued that, although such subjects are probably aware of their behavior (failing effort tests), they may be unaware of their motivation for doing so. In the face of such difficulties, the examiner may forgo trying to detangle consciousness or motivation and simply note the presence of exaggeration and its impact on the reliability of other examination data.

Conclusion

Assessment of exaggeration or feigning of PTSD presents substantial challenges, but psychologists have a number of tools that are well validated in multiple clinical populations, with growing support in PTSD-specific populations. While self-report and structured interview approaches will likely remain vital, the addition of collateral data, cognitive SVTs, and instruments like the MENT can improve professionals' ability to not only detect feigning but to efficiently identify truly impaired patients and claimants.

Appendix

Summary of Trauma Symptom Inventory studies

Scale or Index	Study	Simulators (unless indicated)	Comparison	Validity Control for PTSD?	Instructed to avoid detection? ^a	Sens	FP	LR	Cohen's <i>d</i>
ATR>58 ^b	Elhai et al. (2005)	Students	PTSD diagnosed outpatients	N	N	0.70	0.57	1.2	0.48
ATR≥61	Edens et al. (1998)	Students	Students	N, LR	Y	0.78	0.08	9.8 ^c	-
ATR≥61	Edens et al. (1998)	20 Students with trauma and one TSI elevation (feign)	20 Students with trauma and one TSI elevation	N, LR	Y	-	0.15	-	-
ATR>61 ^b	Elhai et al. (2005)	Students	PTSD - diagnosed outpatients	N	N	0.65	0.45	1.4	-
ATR≥61	Guriel-Tennant and Fremouw (2006)	Students with recent trauma, neg on PCL (naïve)	Students with recent trauma, pos on PCL and CAPS	N, LR	Y	0.82	-	-	-
		Students with recent trauma, neg on PCL (coached)	Students with recent trauma, pos on PCL and CAPS	N, LR	Y	0.32	-	-	-
ATR≥61	Rosen (2006)	Students	Students	N, LR	N	0.68	0.20	3.4 ^c	1.00 ^c
ATR>61	Guriel et al. (2004)	Students (uncoached)	Students	N, LR	Y	0.84	-	-	1.60 ^c
		Students (coached symptoms)	Students	N, LR	Y	0.83	-	-	1.46 ^c
		Students (coached Sxs and strategies)	Students	N, LR	Y	0.81	-	-	1.28 ^c
ATR>61 ^d	Edens et al. (1998) reanalysis by Rosen (2006)	Students with two elevations on TSI	Students with two elevations on TSI	N, LR	Y	0.86	0.42	2.0	-
ATR≥61	Efendov et al. (2008)	Remitted trauma, valid MMPI and TSI (uncoached)	PTSD claimants, assessed if suspected	P, HR ^c	Y	0.72	0.51	1.4	0.74
ATR≥61	Efendov et al. (2008)	Remitted trauma, valid MMPI and TSI (coached)	PTSD claimants, assessed if suspected	P, HR ^c	Y	0.55	0.51	1.1	0.14
ATR>61	Edens et al. (1998)		Various clinical groups	?	?	-	0.23 ^f	-	-
ATR>63	Rosen (2006)	Students with two elevations on TSI	Students with two elevations on TSI	NA	N	0.65	0.36	1.8	-
ATR>63 ^b	Elhai et al. (2005)	Students	PTSD diagnosed outpatients	N	N	0.52	0.30	1.7	-
ATR>64	Edens et al. (1998)		Various clinical groups	?	?	-	0.17 ^f	-	-

ATR>70	Edens et al. (1998) reanalysis by Rosen et al. (2005)	Students with two elevations on TSI	Students with two elevations on TSI	Y	Y	0.74	0.21	3.5	-
ATR≥90	Edens et al. (1998)	Students	Students	Y	Y	0.27	0.01	48.0	-
ATR≥90	Efendov et al. (2008)	Remitted trauma, valid MMPI and TSI (Uncoached)	PTSD claimants, assessed if suspected	P, HR ^e	Y	0.34	0.87	0.4	-
ATR≥90	Efendov et al. (2008)	Remitted trauma, valid MMPI and TSI (Coached)	PTSD claimants, assessed if suspected	P, HR ^e	Y	0.16	0.87	0.2	-
ATR>90	Elhai et al. (2005)	Students	PTSD diagnosed outpatients	N	Y	0.22	0.40	0.6	-
ATR>90 ^g	Porter (2007)	Students with trauma since age 16	Students with trauma since age 16	NA	N	0.19	0.10	1.9	0.61

HR = high risk; LR = low risk; P = partial.

^aIncludes warnings about validity scales.

^bArticle does not specify whether cutoff score was inclusive of feigning or not. Identified cutoff rule presumed from use of > 90 in other part of article.

^cEffect size figure is computed from comparing an honest responding, non trauma-exposed group to a simulating group.

^dOriginal cutoff score used by Eden et al. was ≥ 61, not >61.

^eFigures shown derived from weighted average of various clinical groups (n=659).

^fCutoff score actually used is not specified, but is presumed from context and figures for sensitivity and specificity.

References

- Allen, L. M., Conder, R. L., Green, P., & Cox, D. R. (1997). *CARB '97: Manual for the computerized assessment of response bias*. Durham: CogniSyst.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington DC: American Psychiatric Association. Text revision.
- American Psychological Association (2002). Ethical principles of psychologists and code of conduct. <http://www.apa.org/ethics/code2002.html>.
- Arbisi, P. A., & Ben-Porath, Y. (1995). An MMPI-2 infrequent response scale for use with psychopathological populations: the infrequency psychopathology scale (F(p)). *Psychological Assessment*, 7, 424–431.
- Arbisi, P. A., & Butcher, J. N. (2004). Failure of the FBS to predict malingering of somatic symptoms: response to critiques by Greve and Bianchini and Lees Haley and Fox. *Archives of Clinical Neuropsychology*, 19, 341–345.
- Arbisi, P. A., Ben-Porath, Y. A., & McNulty, J. (2006). The ability of MMPI-2 to detect feigned PTSD within the context of compensation-seeking. *Psychological Services*, 3, 249–261.
- Archer, R. P., Buffington-Vollum, J. K., Stredny, R. V., & Handel, R. W. (2006). A survey of psychological test use patterns among forensic psychologists. *Journal of Personality Assessment*, 87, 84–94.
- Ardolf, B. R., Denney, R. L., & Houston, C. M. (2007). Base rates of negative response bias and malingered neurocognitive dysfunction among criminal defendants referred for neuropsychological evaluation. *The Clinical Neuropsychologist*, 21, 899–916.
- Ashendorf, L., Constantinou, M., & McCaffrey, R. J. (2004). The effect of depression and anxiety on the TOMM in community-dwelling older adults. *Archives of Clinical Neuropsychology*, 19 (1), 125–130.
- Babikian, T., & Boone, K. B. (2007). Intelligence tests as measures of effort. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective*, pp. 103–127. New York: Guilford.
- Bauer, L., O'Bryant, S. E., Lynch, J. K., McCaffrey, R. J., & Fisher, J. M. (2007). Examining the test of memory malingering trial 1 and word memory test immediate recognition as screening tools for insufficient effort. *Assessment*, 14(3), 215–222. doi:10.1177/1073191106297617.
- Beblo, T., Saavedra, A. S., Mensebach, C., & Driessen, M. (2005). Cognitive and psychophysiological functioning in borderline personality and posttraumatic stress disorder. In T. A. Corales (Ed.), *Focus on posttraumatic stress disorder research*, pp. 51–70. New York: Nova Science.
- Ben-Porath, Y., & Tellegan, A. (2008). *MMPI-2 RF™ (Minnesota Multiphasic Personality Inventory-2 Restructured Form™)*. Minneapolis: University of Minnesota Press, Pearson Assessment Systems.
- Ben-Porath, Y. S., Greve, K. W., Bianchini, K. J., & Kaufmann, P. M. (2009). The MMPI-2 symptom validity scale (FBS) is an empirically-validated measure of over-reporting in personal injury litigants and claimants: Reply to Butcher et al. (2008). *Psychological Injury and Law*, 2(1), 62–85.
- 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. 226–265. New York: Oxford University Press.
- Bianchini, K. J., Etherton, J. L., Greve, K. W., Heinly, M. T., & Meyers, J. E. (2008). Classification accuracy of MMPI-2 validity scales in the detection of pain-related malingering: A known-groups study. *Assessment*, 15(4), 435–449.
- Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Gusman, F. D., Dennis, S., et al. (1995). The development of a clinician-administered PTSD scale. *Journal of Traumatic Stress*, 8(1), 75–90.
- Blanchard, E. B., & Hickling, E. J. (1997). *After the crash: Assessment and treatment of motor vehicle accident survivors*. Washington DC: American Psychological Association.
- Blaskewitz, N., Merten, T., & Kathmann, N. (2008). Performance of children on symptom validity tests: TOMM, MSVT & FIT. *Archives of Clinical Neuropsychology*, 23(4), 379–391.

- Boone, K. B. (ed). (2007a). *Assessment of feigned cognitive impairment: A neuropsychological perspective*. New York: Guilford.
- Boone, K. B. (2007b). A reconsideration of the Slick et al. criteria for malingered neurocognitive dysfunction. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective*, pp. 29–49. New York: Guilford.
- Bowman, M. L., & Yahuda, R. (2004). Risk factor and the adversity-stress model. In G. L. Rosen (Ed.), *Posttraumatic stress disorder: Issues and controversies*, pp. 15–38. Hoboken: Wiley.
- Brady, K. (1997). Posttraumatic stress disorder and comorbidity: Recognizing the many faces of PTSD. *Journal of Clinical Psychiatry*, 58(Suppl 9), 12–15.
- Brady, K., Killeen, T., Brewerton, T., & Lucerini, S. (2000). Comorbidity of psychiatric disorders and posttraumatic stress disorder. *Journal of Clinical Psychiatry*, 61(Suppl. 7), 22–32.
- Breslau, N., & Kessler, R. C. (2001). The stressor criterion in DSM-IV post-traumatic stress disorder: An empirical investigation. *Biological Psychiatry*, 50, 699–704.
- Brewin, C. R., Andrews, B., & Valentine, J. D. (2000). Meta-analysis of risk factors for posttraumatic stress disorder in trauma-exposed adults. *Journal of Consulting and Clinical Psychology*, 68, 748–766.
- Brewin, C. R., Kellner, J. S., Vasterling, J. J., & Field, A. P. (2007). Memory for emotionally neutral information in posttraumatic stress disorder: A meta-analytic investigation. *Journal of Abnormal Psychology*, 116, 448–463.
- Briere, J. (1995). *Trauma Symptom Inventory professional manual*. Odessa: Psychological Assessment Resources.
- Briere, J. (2001). *Detailed assessment of posttraumatic stress professional manual*. Odessa: Psychological Assessment Resources.
- Briere, J. (2002). *Psychological assessment of post-traumatic stress disorder*. Washington: APA.
- Brunello, N., Davidson, J., Deahl, M., Kessler, R. C., Mendlewicz, J., Racagni, G., Shalev, A. Y., & Zohar, J. (2001). Posttraumatic stress disorder: Diagnosis and epidemiology, comorbidity and social consequences, biology and treatment. *Neuropsychobiology*, 43(3), 150–162.
- Brockhaus, R., & Merten, T. (2004). Neuropsychological assessment of suboptimal performance: The Word Memory Test. *Nervenarzt*, 75 (9), 882–887. <http://www.ncbi.nlm.nih.gov/pubmed/15372160>.
- Burges, C., & McMillan, T. M. (2001). The ability of naive participants to report symptoms of post-traumatic stress disorder. *British Journal of Clinical Psychology*, 40, 209–214.
- Burkett, B. G., & Whitley, C. (1998). *Stolen valor: How the Vietnam generation robbed of its heroes and its history*. Dallas: Verity.
- Butcher, J. N., Arbisi, P. A., Atlis, M. M., & McNulty, J. L. (2003). The construct validity of the Lees-Haley Fake Bad Scale: Does this scale measure somatic malingering and feigned emotional distress? *Archives of Clinical Neuropsychology*, 18, 473–485.
- Burriss, L., Ayers, E., Ginsberg, J., Powell, D. A., & Buchanan, S. L. (2007). Learning and memory impairment in PTSD: relationship to depression. *Depression and Anxiety*, 25(2), 149–157.
- Bury, A. S., & Bagby, R. M. (2002). The detection of feigned uncoached and coached posttraumatic stress disorder with the MMP1-2 in a sample of workplace accident victims. *Psychological Assessment*, 14, 474–484.
- Bush, S. S., Ruff, R. M., Tröster, A. I., Barth, J. T., Koffler, S. P., Pliskin, N. H., et al. (2005). Symptom validity assessment: Practice issues and medical necessity. NAN Policy & Planning Committee. *Archives of Clinical Neuropsychology*, 20, 419–426.
- Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *MMPI-2: Manual for administration and scoring*. Minneapolis: University of Minnesota Press.
- Butcher, J. N., Gass, C. S., Cumella, E., Kally, Z., & Williams, C. L. (2008). Potential for bias in MMPI-2 assessments using the Fake Bad Scale (FBS). *Psychological Injury and Law*, 1(3), 147–211.
- Calhoun, P. S., Earnest, K. S., & Tucker, D. D. (2000). Feigning combat-related posttraumatic stress disorder on the Personality Assessment Inventory. *Journal of Personality Assessment*, 75(2), 338–350.
- Carone, D. A. (2008). Children with moderate/severe brain damage/dysfunction outperform adults with mild-to-no brain damage on the Medical Symptom Validity Test. *Brain Injury*, 22(12), 960–971.
- Committee on Ethical Guidelines for Forensic Psychologists. (1991). Specialty guidelines for forensic psychologists. *Law and Human Behavior*, 15(6), 655–665.
- Crawford, E. F., Greene, R. L., Dupart, T. M., Bongar, B., & Childs, H. (2006). MMPI-2 assessment of malingered emotional distress related to a workplace injury: A mixed group validation. *Journal of Personality Assessment*, 86(2), 217–221.
- Cromwell, T. A., Kieffer, K. M., Siders, C. A., & Vanderploeg, R. D. (2002). Neuropsychological findings in combat-related posttraumatic stress disorder. *The Clinical Neuropsychologist*, 16, 310–321.
- Daly, O. E., & Johnston, T. G. (2002). The Derryhirk Inn Incident: The psychological sequelae. *Journal of Traumatic Stress*, 15, 461–464.
- Daubert, S. D., & Metzler, A. E. (2000). The detection of fake-bad and fake-good responding on the Millon Clinical Multiaxial Inventory III. *Psychological Assessment*, 12, 418–424.
- Dearth, C. S., Berry, D. T. R., Vickery, C. D., Vagnini, V. L., Baser, R. E., Orey, S. A., & Cragar, D. E. (2005). Detection of feigned head injury symptoms on the MMPI-2 in head injured patients and community controls. *Archives of Clinical Neuropsychology*, 20(1), 95–110.
- Demakis, G. J., Gervais, R. O., & Rohling, M. L. (2008). The effect of failure on cognitive and psychological symptom validity tests in litigants with symptoms of post-traumatic stress disorder. *The Clinical Neuropsychologist*, 22(5), 879–895.
- Drukteinis, A. M. (2003). Disability determination in PTSD litigation. In R. I. Simon (Ed.), *Posttraumatic stress disorder in litigation: Guidelines for forensic assessment* (2nd ed.), pp. 141–161. Washington: American Psychiatric Press.
- Duncan, A. (2005). The impact of cognitive and psychiatric impairment of psychotic disorders on the Test of Memory Malingering (TOMM). *Assessment*, 12(2), 123–129.
- Eakin, D. E. (2004). *Detection of feigned posttraumatic stress disorder: A multimodal assessment strategy*. Unpublished doctoral dissertation, Auburn University, Auburn, AL.
- Eakin, D. E., Weathers, F. W., Benson, T. B., Anderson, C. F., & Funderburk, B. (2006). Detection of feigned posttraumatic stress disorder: A comparison of the MMPI-2 and PAI. *Journal of Psychopathology and Behavioral Assessment*, 28(3), 145–155.
- Eaton, W. W., Sigal, J. J., & Weinfield, M. (1987). Impairment in holocaust survivors after 33 years: Data from an unbiased community sample. *American Journal of Psychiatry*, 139, 773–777.
- Edens, J. F., Otto, R. R., & Dwyer, T. J. (1998). Susceptibility of the Trauma Symptom Inventory to malingering. *Journal of Personality Assessment*, 71, 379–392.
- Efendov, A. A., Sellbom, M., & Bagby, R. M. (2008). The utility and comparative incremental validity of the MMPI-2 and Trauma Symptom Inventory validity scales in the detection of feigned PTSD. *Psychological Assessment*, 20(4), 317–326.
- Elhai, J. D., Gold, P. B., Frueh, B. C., & Gold, S. N. (2000). Cross-validation of the MMPI-2 in detecting malingered posttraumatic stress disorder. *Journal of Personality Assessment*, 75(3), 449–446.
- Elhai, J., Gold, S., Sellers, A., & Dorfman, W. (2001). The detection of malingered posttraumatic stress disorder with MMPI-2 fake bad indices. *Assessment*, 8(2), 221–236.
- Elhai, J. D., Ruggiero, K., Frueh, B., Beckham, J. C., Gold, P. B., & Feldman, M. E. (2002). The infrequency-posttraumatic stress disorder scale (F(p) tsd) for the MMPI-2: Development and

- initial validation with veterans presenting with combat-related PTSD. *Journal of Personality Assessment*, 79(3), 531–549.
- Elhai, J., Naifeh, J., Zucker, I., Gold, S. E., Deitsch, S. E., & Frueh, B. C. (2004). Discriminating malingering from genuine civilian posttraumatic stress disorder: A validation of three MMPI-2 infrequency scales (F, F(p), and F(p) tsd). *Assessment*, 11(2), 139–144.
- Elhai, J. D., Gray, M. J., Naifeh, J. A., Butcher, J. J., Davis, J. L., Falsetti, S. A., et al. (2005). Utility of the Trauma Symptom Inventory's Atypical Response Scale in detecting malingering post-traumatic stress disorder. *Assessment*, 12(2), 210–219.
- 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.
- Ford, C. V. (2005). *Lies! Lies!! Lies!!!: The psychology of deceit*. Washington: American Psychiatric Press.
- Franklin, C. L., & Thompson, K. E. (2005). Response style and posttraumatic stress disorder (PTSD): A review. *Journal of Trauma & Dissociation*, 6(3), 105–123.
- Freeman, T., Powell, M., & Kimbrell, T. (2008). Measuring symptom exaggeration in veterans with chronic posttraumatic stress disorder. *Psychiatry Research*, 158, 374–380.
- Friedman, A. E., Lewak, R., Nichols, D. S., & Webb, J. T. (2001). *Psychological assessment with the MMPI-2*. Mahwah: Lawrence Erlbaum.
- Frueh, B. C., Hamner, M., Cahill, S., Gold, P., & Hamlin, K. (2000). Apparent symptom overreporting among combat veterans evaluated for PTSD. *Clinical Psychology Review*, 20(7), 853–885.
- Frueh, B. C., Elhai, J. D., Grubaugh, A. L., Monnier, J., Kashdan, T. B., Sauvageot, J. A., et al. (2005). Documented combat exposure of US veterans seeking treatment for combat-related post-traumatic stress disorder. *British Journal of Psychiatry*, 186, 467–472. discussion 473–475.
- Geraerts, E., Jelicic, M., & Merckelbach, H. (2006a). Symptom overreporting and recovered memories of childhood sexual abuse. *Law and Human Behavior*, 30(5), 621–630.
- Geraerts, E., Kozaric-Kovacic, D., Merckelbach, H., Peraica, T., Jelicic, M., & Candel, I. (2006b). Detecting deception of war-related posttraumatic stress disorder. *Journal of Forensic Psychiatry & Psychology*, 20, 278–285.
- Gervais, R. O., Russell, A., Green, P., Allen, L. M., Ferrai, R., & Pieschl, S. (2001). Effort testing in patients with fibromyalgia and disability incentives. *Journal of Rheumatology*, 28(8), 1892–1899.
- Gervais, R. O., Rohling, M. L., Green, P., & Ford, W. (2004). A comparison of WMT, CARB, and TOMM failure rates in non-head injury disability claimants. *Archives of Clinical Neuropsychology*, 19, 475–487.
- Gervais, R. O., Ben-Porath, Y. S., Wygant, D. B., & Green, P. (2007). Development and validation of a Response Bias Scale (RBS) for the MMPI-2. *Assessment*, 14(2), 196–208.
- Gervais, R. O., Ben-Porath, Y. S., Wygant, D. B., & Green, P. (2008). Differential sensitivity of the Response Bias Scale (RBS) and MMPI-2 validity scales to memory complaints. *The Clinical Neuropsychologist*, 22(6), 1061–1079. doi:10.1080/13854040701756930.
- Goldberg, H. E., Back-Madruga, C., & Boone, K. B. (2007). The impact of psychiatric disorders on cognitive symptom validity tests. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective*, pp. 281–309. New York: Guilford.
- Goldstein, G., van Kammern, W., Shelly, C., Miller, D. J., & van Kammern, D. P. (1987). Survivors of imprisonment in the Pacific theatre during World War II. *American Journal of Psychiatry*, 144, 1210–1213.
- Gorissen, M., Sanz de la Torre, J. C., & Schmand, B. (2005). Effort and memory in schizophrenia patients. *Schizophrenia Research*, 7(2-3), 199–208.
- Greene, R. L. (2000). *The MMPI-2: An interpretive manual* (2nd ed.). Boston: Allyn & Bacon.
- Green, P. (2004). *Green's Medical Symptom Validity Test (MSVT) user's manual and program*. Edmonton: Green's.
- Green, P. (2005). *Green's Word Memory Test user's manual*. Edmonton: Green's.
- Green, P. (2006). *Green's Non-Verbal Medical Symptom Validity Test (NV-MSVT) user's manual*. Edmonton: Green's.
- Green, P. (2007). Spoiled for choice: Making comparisons between forced-choice effort tests. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective*, pp. 50–77. New York: Guilford.
- Green, P. (2008). Questioning common assumptions about depression. In J. Morgan & J. Sweet (Eds.), *Neuropsychology of malingering casebook*, pp. 132–144. New York: AACN and Psychology Press.
- Greene, R. L. (2008). Malingering and defensiveness on the MMPI-2. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed.), pp. 159–181. New York: Guilford.
- Greiffenstein, M. F., Baker, W. J., Axelrod, B., Peck, E. A., & Gervais, R. (2004). The Fake Bad Scale and MMPI-2 F-family in detection of implausible psychological trauma claims. *The Clinical Neuropsychologist*, 18, 573–590.
- Greiffenstein, M. F., Fox, D., & Fox, P. R. (2007). The MMPI-2 Fake Bad Scale in detection of noncredible brain injury claims. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective*, pp. 210–238. New York: Guilford.
- Greve, K. W., & Bianchini, J. (2004a). Setting empirical cut-offs on psychometric indicators of negative response bias: A methodological commentary with recommendations. *Archives of Clinical Neuropsychology*, 19, 533–541.
- Greve, K. W., & Bianchini, J. (2004b). Response to Butcher et al. (2003): The construct validity of the Lees-Haley Fake Bad Scale. *Archives of Clinical Neuropsychology*, 19(3), 337–339.
- Greve, K., Bianchini, K. J., Mathias, C. W., Houston, R. J., & Couch, R. J. (2003). Detecting malingering performance on the Wechsler Adult Intelligence Scale: Validation of Mittenberg's approach in traumatic brain injury. *Archives of Clinical Neuropsychology*, 18, 245–260.
- Greve, K., Bianchini, K., Love, J., Brennan, A., & Heinly, M. (2006). Sensitivity and specificity of MMPI-2 validity scales and indicators to malingering neurocognitive dysfunction in traumatic brain injury. *The Clinical Neuropsychologist*, 20(3), 491–512.
- Greve, K. W., Ord, J., Curtis, K. L., Bianchini, K. J., & Brennan, A. (2008). Detecting malingering in traumatic brain injury and chronic pain: A comparison of three forced-choice symptom validity tests. *The Clinical Neuropsychologist*, 22(5), 896–918.
- Greve, K. W., Binder, L. M., & Bianchini, K. J. (2009). Rates of below-chance performance in forced-choice symptom validity tests. *The Clinical Neuropsychologist*, 23(3), 534–544.
- Guriel, J., & Fremouw, W. (2003). Assessing malingering posttraumatic stress disorder: A critical review. *Clinical Psychology Review*, 23, 881–904.
- Guriel, J., Yanez, Y. T., Fremouw, W., Shreve-Neiger, A., Ware, L., Filcheck, H., et al. (2004). Impact of coaching on malingering posttraumatic stress symptoms on the M-FAST and the TSI. *Journal of Forensic Psychology Practice*, 4(2), 37–56.
- Guriel-Tennant, J., & Fremouw, W. (2006). Impact of trauma history and coaching on malingering of posttraumatic stress disorder using the PAI, TSI, and M-FAST. *Journal of Forensic Psychiatry & Psychology*, 17(4), 577–592.
- Guy, L. S., Kwartner, P. P., & Miller, H. A. (2006). Investigating the M-FAST: psychometric properties and utility to detect diagnostic specific malingering. *Behavioral Sciences & the Law*, 24(5), 687–702.

- Hall, R. C. W., & Hall, R. C. W. (2006). Malingering of PTSD: Forensic and diagnostic considerations, characteristics of malingerers and clinical presentations. *General Hospital Psychiatry, 28*, 525–535.
- Hamilton, J. C., Feldman, M. D., & Cunnien, A. J. (2008). Factitious disorder in medical and psychiatric practices. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed.), pp. 128–144. New York: Guilford.
- Hartman, D. E. (2002). The unexamined lie is a lie worth fibbing: Neuropsychological malingering and the Word Memory Test. *Archives of Clinical Neuropsychology, 17*, 709–734.
- Henry, G. I., Heilbronner, R., Mittenberg, W., & Enders, C. (2006). The Henry–Heilbronner Index: A 15-item empirically derived MMPI-2 subscale for identifying probable malingering in personal injury litigants and disability claimants. *The Clinical Neuropsychologist, 20*(4), 786–797.
- Hickling, E. J., Blanchard, E. B., Mundy, E., & Galovski, T. E. (2002). Detection of malingered MVA related posttraumatic stress disorder: An investigation of the ability to detect professional actors by experienced clinicians psychological tests and psychophysiological assessment. *Journal of Forensic Psychology Practice, 2*, 33–54.
- Homer, M. D., Mintzer, J., Turner, T. H., Edmiston, K., Mintzer, O., & Johnson, R. H. (2007). Cognitive functioning in a clinical sample of patients with Post-Traumatic Stress Disorder. Paper presented at the annual meeting of the National Academy of Neuropsychology, Scottsdale, AZ, November.
- Howe, L. S. L., & Loring, D. W. (2008). Classification accuracy and predictive ability of the Medical Symptom Validity Test's dementia profile and general memory impairment profile. *The Clinical Neuropsychologist, 23*, 329–342.
- Hurley, K. E., & Deal, W. P. (2006). Instruments measuring malingering used with individuals who have mental retardation: Potential problems and issues. *Mental Retardation, 44*(2), 112–119.
- Hyer, L., Fallon, J., Harrison, W., & Boudewyns, P. (1987). MMPI overreporting by Vietnam combat veterans. *Journal of Clinical Psychology, 43*(1), 79–83.
- Hyer, L., Boudewyns, P., Harrison, W., O'Leary, W. C., Bruno, R. D., Saucer, R. T., et al. (1988). Vietnam veterans: Overreporting versus acceptable reporting of symptoms. *Journal of Personality Assessment, 52*(3), 475–486.
- Iverson, G. L. (1996). A comment on the willingness of people to malingering following motor vehicle or work-related injuries. *Journal of Cognitive Rehabilitation, May/June*, 10–15.
- Iverson, G. L., & Franzen, M. D. (1996). Using multiple memory procedures to detect simulated malingering. *Journal of Clinical and Experimental Neuropsychology, 18*(1), 36–51.
- Iverson, G. L., Le Page, J., Koehler, B. E., & Shojania, K. (2007). Test of Memory Malingering (TOMM) scores are not affected by chronic pain or depression in patients with fibromyalgia. *The Clinical Neuropsychologist, 21*(3), 532–546.
- Jelicic, M., Merckelbach, H., Candel, I., & Geraerts, E. (2007). Detection of feigned cognitive dysfunction using special malingering tests: a simulation study in naïve and coached malingerers. *International Journal of Neuroscience, 117*(8), 1185–1192.
- Johnsen, G. E., Kanagaratnama, P., & Asbjørnsena, A. E. (2008). Memory impairments in posttraumatic stress disorder are related to depression. *Journal of Anxiety Disorders, 22*(3), 464–474.
- Katz, M. M., & Warren, W. L. (1999). *Katz Adjustment Scales relative report form (KAS-R) manual*. Los Angeles: Western Psychological Services.
- Keane, T. M., Kolb, L. C., Kaloupek, D. G., Orr, S. P., Blanchard, E. B., Thomas, R. G., et al. (1998). Utility of psychophysiological measurement in the diagnosis of posttraumatic stress disorder: results from a Department of Veterans Affairs Cooperative Study. *Journal of Consulting and Clinical Psychology, 66*(6), 914–923.
- Keane, T. M., Buckley, T. C., & Miller, W. (2003). Forensic psychological assessment in PTSD. In R. I. Simon (Ed.), *Posttraumatic stress disorder in litigation: Guidelines for forensic assessment* (2nd ed.), pp. 119–140. Washington: American Psychiatric Publishing.
- Kessler, R. C., Sonnega, A., Bromet, E., Hughes, M., & Nelson, C. B. (1995). Posttraumatic stress disorder in the National Comorbidity Survey. *Archives of General Psychiatry, 52*, 1048–1060.
- Knoll, J., & Resnick, P. J. (2006). The detection of malingered post-traumatic stress disorder. *Psychiatric Clinics of North America, 29*(3), 629–647.
- Koch, W., O'Neill, M., & Douglas, K. (2005). Empirical limits for the forensic assessment of PTSD litigants. *Law and Human Behavior, 29*(1), 121–149.
- Kremen, W. S., Koenen, K. C., Boake, C., Purcell, S., Eisen, S. A., Franz, C. E., et al. (2007). Pretrauma cognitive ability and risk for posttraumatic stress disorder: A twin study. *Archives of General Psychiatry, 64*, 361–368.
- Lally, S. J. (2003). What tests are acceptable for use in forensic evaluations? A survey of experts. *Professional Psychology: Research and Practice, 34*(5), 491–498.
- Larrabee, G. J. (2003). Detection of malingering using atypical performance patterns on standard neuropsychological tests. *The Clinical Neuropsychologist, 17*, 410–425.
- 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. (2007a). Identification of malingering by pattern analysis on neuropsychological tests. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits*, pp. 80–99. New York: Oxford.
- Larrabee, G. J. (2007b). Evaluation of exaggerated health and injury symptomatology. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits*, pp. 264–286. New York: Oxford University Press.
- Larrabee, G. J. (2008). Aggregation across multiple indicators improves the detection of malingering: Relationship to likelihood ratios. *The Clinical Neuropsychologist, 22*, 666–679.
- Larrabee, G. J., Greiffenstein, M., Greve, K., & Bianchini, K. W. (2007). Refining diagnostic criteria for malingering. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits*, pp. 334–371. New York: Oxford University Press.
- Lees-Haley, P. R. (1990). Malingering mental disorder on the Impact of Events Scale (IES): Toxic exposure and cancer phobia. *Journal of Traumatic Stress, 3*, 315–321.
- Lees-Haley, P. R. (1992). Efficacy of MMPI-2 validity scales and MCMI-II modifier scales for detecting spurious PTSD claims: F, F-K, Fake Bad Scale, ego strength, subtle-obvious subscales, DIS, and DEB. *Journal of Clinical Psychology, 48*(5), 681–689.
- Lees-Haley, P. R. (1997a). MMPI-2 base rates for 492 personal injury plaintiff: Implications and challenges for forensic assessment. *Journal of Clinical Psychology, 53*, 745–755.
- Lees-Haley, P. R. (1997b). Attorneys influence expert evidence in forensic psychological and neuropsychological cases. *Assessment, 4*, 321–324.
- Lees-Haley, P. R., & Dunn, J. T. (1994). The ability of naive subjects to report symptoms of mild brain injury, post-traumatic stress disorder, major depression, and generalized anxiety disorder. *Journal of Clinical Psychology, 50*(2), 252–256.
- Lees-Haley, P. R., & Fox, D. D. (2004). Commentary on Butcher, Arbisi, and McNulty (2003) on the Fake Bad Scale. *Archives of Clinical Neuropsychology, 19*(3), 333–336.
- Lees-Haley, P. R., English, L. T., & Glenn, W. J. (1991). A fake bad scale on the MMPI-2 for personal injury claimants. *Psychological Reports, 68*, 203–210.

- 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*(2), 170–176.
- Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2000). The scientific status of projective techniques. *Psychological Science in the Public Interest, 1*, 27–66.
- Liljequist, L., Kinder, B. N., & Schinka, J. A. (1998). An investigation of malingering posttraumatic stress disorder on the Personality Assessment Inventory. *Journal of Personality Assessment, 71*(3), 322–336.
- Lynch, W. J. (2004). Determination of effort level, exaggeration, and malingering in neurocognitive assessment. *Journal of Head Trauma Rehabilitation, 19*, 277–283.
- Marshall, M. B., & Bagby, R. M. (2006). The incremental validity and clinical utility of the MMPI-2 Infrequency Posttraumatic Stress Disorder Scale. *Assessment, 13*(4), 417–429. doi:10.1177/1073191106290842.
- McCrae, R. R., & Costa, P. T., Jr. (1990). *Personality in adulthood*. New York: Guilford.
- McGuire, B. E. (1999). The assessment of malingering in traumatic stress claimants. *Psychiatry, Psychology, and the Law, 6*(2), 163–173.
- McNally, R. J. (2003). Progress and controversy in the study of posttraumatic stress disorder. *Annual Review of Psychology, 54*, 229–252.
- McNally, R. J. (2004). Conceptual problems with the DSM-IV criteria for posttraumatic stress disorder. In G. L. Rosen (Ed.), *Posttraumatic stress disorder: Issues and controversies*, pp. 1–14. Hoboken: Wiley.
- Merten, T., Friedel, E., & Stevens, A. (2006). Insufficient cooperativeness in forensic neuropsychiatric assessment: Prevalence estimates of negative response bias. *Versicherungsmedizin, 58*(1), 19–21.
- Messer, J. M., & Fremouw, W. J. (2007). Detecting malingered posttraumatic stress disorder using the Morel Emotional Numbing Test-Revised (MENT-R) and the Miller Forensic Assessment of Symptoms Test (M-FAST). *Journal of Forensic Psychology Practice, 7*(3), 33–57.
- Meyers, J. W., & Volbrecht, M. W. (2003). A validation of multiple malingering detection methods in a large clinical sample. *Archives of Clinical Neuropsychology, 18*, 261–276.
- Meyers, J. W., Millis, S. R., & Volkert, K. (2002). A validity index for the MMPI-2. *Archives of Clinical Neuropsychology, 17*(2), 157–169.
- Miller, H. A. (2001). *M-FAST: Miller Forensic Assessment of Symptoms Test professional manual*. Odessa: Psychological Assessment Resources.
- Miller, M. W. (2003). Personality and the etiology and expression of PTSD: A three-factor perspective. *Clinical Psychology: Science and Practice, 10*, 373–393.
- Miller, H. A. (2005). The Miller-Forensic Assessment of Symptoms Test (M-FAST): Test generalizability and utility across race, literacy, and clinical opinion. *Journal of Criminal Justice and Behavior, 32*(6), 591–611.
- Millon, T., Davis, R., & Millon, C. (1997). *The Millon Clinical Multiaxial Inventory-III manual* (2nd ed.). Minneapolis: National Computer Systems.
- Mittenberg, W., Theroux, S., Aguila-Puentes, G., Bianchini, K., Greve, K., & Rayls, K. (2001). Identification of malingered head injury on the Wechsler Intelligence Scale, 3rd ed. *The Clinical Neuropsychologist, 15*, 440–445.
- Monnier, J., Kashdan, T. B., Sauvageot, J. A., Hamner, M. B., Burkett, B. G., George, W., et al. (2005). Documented combat exposure of US veterans seeking treatment for combat-related post-traumatic stress disorder. *British Journal of Psychiatry, 186*, 467–472.
- Moore, S. A. (2009). Cognitive abnormalities in posttraumatic stress disorder. *Current Opinion in Psychiatry, 22*(1), 19–24.
- Morel, K. B. (1998). Development and preliminary validation of a forced-choice test of response bias for posttraumatic stress disorder. *Journal of Personality Assessment, 70*(2), 299–314.
- Morel, K. B. (2007). Comparison of the Morel Emotional Numbing Test for Posttraumatic Stress Disorder to the Word Memory Test in neuropsychological evaluations. *The Clinical Neuropsychologist, 22*(2), 350–362.
- Morel, K. B. (2008). Development of a validity scale for combat-related posttraumatic stress disorder: Evidence from simulated malingerers and actual disability claimants. *Journal of Forensic Psychiatry & Psychology, 19*(1), 52–63.
- Morel, K. R., & Marshman, K. C. (2008). Critiquing symptom validity tests for posttraumatic stress disorder: A modification of Hartman's criteria. *Journal of Anxiety Disorders, 22*(8), 1542–1550.
- Morel, K. R., & Shepherd, B. E. (2008a). Developing a Symptom Validity Test for posttraumatic stress disorder: Application of the binomial distribution. *Journal of Anxiety Disorders, 22*(8), 1297–1302.
- Morel, K. R., & Shepherd, B. E. (2008b). Meta-analysis of the Morel Emotional Numbing Test for PTSD: Comment on Singh, Avasthi, and Grover. *German Journal of Psychiatry, 11*, 128–131.
- Morey, L. (1991). *Personality Assessment Inventory: Professional manual*. Tampa: Psychological Assessment Resources.
- Morey, L. (1996). *An interpretive guide to the Personality Assessment Inventory (PAI)*. Tampa: Psychological Assessment Resources.
- National Academy of Neuropsychology. (2003). Independent and court-ordered forensic neuropsychological examinations: Official statement of the National Academy of Neuropsychology. <http://nanonline.org/NAN/ResearchPublications/PositionPapers/IndependentAndCourtOrderedExams.aspx>. Accessed January 17, 2009.
- Nelson, N. W., Sweet, J. J., & Heilbronner, R. L. (2007a). Examination of the new MMPI-2 Response Bias Scale (Gervais): Relationship with MMPI-2 validity scales. *Journal of Clinical and Experimental Neuropsychology, 29*(1), 67–72.
- Nelson, N. W., Sweet, J. J., Berry, D. T. R., Bryant, F. B., & Granacher, R. P. (2007b). Response validity in forensic neuropsychology: Exploratory factor analytic evidence of distinct cognitive and psychological constructs. *Journal of the International Neuropsychological Society, 13*, 440–449.
- Nelson, N. W., Boone, K., Dueck, A., Wangener, L., Lu, P., & Grillis, C. (2003). Relationships between eight measures of suspect effort. *The Clinical Neuropsychologist, 17*(2), 263–272.
- Nelson, N. W., Sweet, J. J., & Demakis, G. J. (2006). Meta-analysis of the MMPI-2 Fake Bad Scale: Utility in forensic practice. *The Clinical Neuropsychologist, 20*, 39–58.
- Nice, D. S., Garland, C. F., Hilton, S. M., Baggett, J. C., & Mitchell, R. E. (1996). Long-term health outcomes and medical effects of torture among US Navy prisoners of war in Vietnam. *Journal of the American Medical Association, 276*, 375–381.
- O'Bryant, S. E., Finlay, C. G., & O'Jile, J. R. (2007). TOMM Performances and self-reported symptoms of depression and anxiety. *Journal of Psychopathology and Behavioral Assessment, 29*(2), 111–114.
- Otto, R. K. (2008). Challenges and advances in assessment of response style in forensic examination contexts. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed.), pp. 365–375. New York: Guilford.
- Pankratz, L. (1998). *Patients who deceive: Assessment and management of risk in providing health care and financial benefits*. Springfield: Charles C. Thomas.
- Pankratz, L., Fausti, S. A., & Peed, S. (1975). A forced-choice technique to evaluate deafness in a hysterical or malingering patient. *Journal of Consulting and Clinical Psychology, 43*, 421–422.

- Parslow, R. A., & Jorm, A. F. (2007). Pretrauma and posttrauma neurocognitive functioning and PTSD symptoms in a community sample of young adults. *American Journal of Psychiatry*, *164*, 509–515.
- Perconte, S., & Goreczny, A. (1990). Failure to detect fabricated posttraumatic stress disorder with the use of the MMPI in a clinical population. *American Journal of Psychiatry*, *147*(8), 1057–1060.
- Pitman, R. K., & Orr, S. P. (2003). Forensic laboratory testing for PTSD. In R. I. Simon (Ed.), *Posttraumatic stress disorder in litigation: Guidelines for forensic assessment*. (2nd ed., pp. 207–223). Washington: American Psychiatric Publishing, Inc.
- Pivovarov, E., & Frederick, R. (2009). Can the MMPI-2 improve interpretation of invalid profiles on the VIP? In R. Frederick (Chair), *Validating tests without criterion groups*. Symposium conducted at the American Psychology Law Society meeting, San Antonio, TX.
- Pole, N. (2007). The psychophysiology of post traumatic stress disorder: A meta-analysis. *Psychological Bulletin*, *133*(5), 725–746.
- Porter, S., Peace, K. A., & Emmett, K. A. (2007). You protest too much methinks: Investigating the features of truthful and fabricated reports of traumatic experiences. *Canadian Journal of Behavioural Science*, *39*(2), 79–91.
- Poythress, N. G., Edens, J. E., & Watkins, M. M. (2001). The relationship between psychopathic personality features and malingering symptoms of major mental illness. *Law and Human Behavior*, *25*(6), 567–581.
- Rees, L. M., Tombaugh, T. N., & Boulay, L. (2001). Depression and the Test of Memory Malingering. *Archives of Clinical Neuropsychology*, *16*(5), 501–506.
- Resnick, P. (1997). Malingering of posttraumatic disorders. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed.), pp. 130–152. New York: Guilford.
- Resnick, P. (2003). Guidelines for the evaluation of malingering in posttraumatic stress disorder. In R. I. Simon (Ed.), *Posttraumatic stress disorder in litigation: Guidelines for forensic assessment* (2nd ed.), pp. 187–205. Washington: American Psychiatric.
- Resnick, P., West, S., & Payne, J. W. (2008). Malingering of posttraumatic disorders. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed.), pp. 109–127. New York: Guilford.
- Rogers, R. (1992). *Structured interview of reported symptoms*. Odessa: Psychological Assessment Resources.
- Rogers, R. (1997a). Introduction. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed.), pp. 1–22. New York: Guilford.
- Rogers, R. (1997b). Researching dissimulation. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed.), pp. 398–426. New York: Guilford.
- Rogers, R. (2008a). Researching response styles. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed.), pp. 411–434. New York: Guilford.
- Rogers, R. (2008b). Detection strategies for malingering and defensiveness. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed.), pp. 14–35. New York: Guilford.
- Rogers, R. (2008c). An introduction to response styles. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed.), pp. 3–13. New York: Guilford.
- Rogers, R., Bagby, R. M., & Dickens, S. E. (1992). *Structured Interview of Reported Symptoms (SIRS) and professional manual*. Odessa: Psychological Assessment Resources.
- Rogers, R., Kropp, P. R., Bagby, R. M., & Dickens, S. E. (1992). Faking specific disorders: A study of the Structured Interview of Reported Symptoms (SIRS). *Journal of Clinical Psychology*, *48* (5), 643–648.
- Rogers, R., Bagby, R. M., & Chakraborty, D. (1993). Feigning schizophrenic disorders on the MMPI-2: Detection of coached simulators. *Journal of Personality Assessment*, *60*(2), 215–226.
- Rogers, R., Sewell, K. W., Martin, M. A., & Vitacco, M. J. (2003). Detection of feigned mental disorders: A meta-analysis of malingering on the MMPI-2. *Assessment*, *10*, 160–177.
- Rogers, R., Payne, J. W., Berry, D. T. R., & Granacher, R. P., Jr. (2008). Use of the SIRS in compensation cases: An examination of its validity and generalizability. *Law and Human Behavior*, . doi:10.1007/s10979-008-9145-9.
- Rohling, M. L., Green, P., Allen, L., & Iverson, G. L. (2002). Depressive symptoms and neurocognitive test scores in patients passing symptom validity tests. *Archives of Clinical Neuropsychology*, *17*(3), 205–222.
- Rosen, G. M. (1995). The Aleutian Enterprise sinking and posttraumatic stress disorder: Misdiagnosis in clinical and forensic settings. *Professional Psychology: Research and Practice*, *26*, 82–87.
- Rosen, G. M. (2004a). Malingering and the PTSD data base. In G. Rosen (Ed.), *Posttraumatic stress disorder: Issues and controversies*, pp. 85–99. Hoboken: Wiley.
- Rosen, G. M. (2004b). Litigation and reported rates of posttraumatic stress disorder. *Personality and Individual Differences*, *36*, 1291–1294.
- Rosen, G. M. (2006). DSM's cautionary guideline to rule out malingering can protect the PTSD database. *Journal of Anxiety Disorders*, *20*(4), 530–535.
- Rosen, G. M., & Phillips, W. R. (2004). A cautionary lesson from simulated patients. *Journal of the American Academy of Psychiatry and the Law*, *32*, 132–133.
- Rosen, G. M., & Powel, J. H. (2003). Use of a symptom validity test in the forensic assessment of posttraumatic stress disorder. *Journal of Anxiety Disorders*, *17*, 361–367.
- Rosen, G. M., & Taylor, S. (2007). Pseudo-PTSD. *Journal of Anxiety Disorders*, *21*, 201–210.
- Rosen, G. M., Sawchuk, C. N., Atkins, D. C., Brown, M., Price, J. R., & Lees-Haley, P. R. (2005). Risk of false positives when identifying malingered profiles using the Trauma Symptom Inventory. *Journal of Personality Assessment*, *86*(3), 329–333.
- Rosen, G. M., Sawchuk, C., Atkins, D. C., Brown, M., Price, I. R., & Lees-Haley, P. R. (2006). The risk of false-positives when using ATR cut-scores to detect malingered posttraumatic reactions on the Trauma Symptom Inventory (TSI). *Journal of Personality Assessment*, *86*(3), 329–333.
- Rubenzon, S. J. (2005). Malingering psychiatric disorders and cognitive impairment in personal injury settings. *For the Defense*, *47*(4), 18–25. 67.
- Rubenzon, S. J. (2006). Malingering of psychiatric problems, brain damage, chronic pain, and controversial syndromes in a personal injury setting. *FDCC Quarterly*, *56*(4), 499–523.
- Sachinvala, N., von-Scotti, H., McGuire, M., Fairbanks, L., Bakst, K., McGuire, M., et al. (2000). Memory, attention, function, and mood among patients with chronic posttraumatic stress disorder. *Journal of Nervous and Mental Disease*, *188*, 818–823.
- Scheibe, S., Bagby, R., Miller, L., & Dorian, B. (2001). Assessing posttraumatic stress disorder with the MMPI-2 in a sample of workplace accident victims. *Psychological Assessment*, *13*(3), 369–374.
- Schoenberg, M. R., Dorr, D., & Morgan, C. D. (2003). The ability of the Millon Clinical Multiaxial Inventory, third edition, to detect malingering. *Psychological Assessment*, *15*, 198–204.
- Scragg, P., Bor, R., & Mendham, M. C. (2000). Assessment feigning post-traumatic stress disorder on the PAI. *Clinical Psychology and Psychotherapy*, *7*, 155–160.

- Sellbom, M., & Bagby, R. M. (2008). Response styles in multiscale inventories. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed.), pp. 182–206. New York: Guilford.
- Sellers, S., Mitchell, K., Byrne, M. K., & Golus, P. (2006). The detection of malingered psychopathology and cognitive deficits: Employing the Fake Bad Scale and the Raven's Standard Progressive Matrices. *Psychiatry, Psychology & Law*, 13(1), 91–99.
- Senior, G., & Douglas, L. (2000). Personal injury base rates and interpretative guidelines for the MMPI-2. Special Topic Presentation at the 20th Annual Conference of the National Academy of Neuropsychology. Orlando, Florida, USA, 15–18 November.
- Sharland, M. J., & Gfeller, J. D. (2007). A survey of neuropsychologists' beliefs and practices with respect to the assessment of effort. *Archives of Clinical Neuropsychology*, 22, 213–223.
- Shuman, D. W. (2003). Persistent reexperiences in psychiatry and law: Current and future trends for the role of PTSD in litigation. In R. I. Simon (Ed.), *Posttraumatic stress disorder in litigation: Guidelines for forensic assessment* (2nd ed.), pp. 1–18. Washington: American Psychiatric.
- Simon, R. I. (2003). Forensic psychiatric assessment of PTSD claimants. In R. I. Simon (Ed.), *Posttraumatic stress disorder in litigation: Guidelines for forensic assessment* (2nd ed.), pp. 41–90. Washington: American Psychiatric.
- Simon, M. J. (2007). Performance of mentally retarded forensic patients on the Test of Memory Malingering. *Journal of Clinical Psychology*, 63(4), 339–344.
- Skodol, A., Schwartz, S., Dohrenwend, B., Levav, L., Shrout, P. E., & Reiff, M. (1996). PTSD symptoms and comorbid mental disorders in Israeli war veterans. *British Journal of Psychiatry*, 169(6), 717–725.
- Slick, D. L., Hopp, G., Strauss, E., & Thompson, G. (1997). *Victoria Symptom Validity Test: Professional manual*. Odessa: Psychological Assessment Resources.
- Slick, D. J., Sherman, E. M. S., & Iverson, G. L. (1999). Diagnostic criteria for malingered neurocognitive dysfunction: Proposed standards for clinical practice and research. *The Clinical Neuropsychologist*, 13(40), 545–561.
- Slick, D. L., Tan, J. E., Strauss, E., & Hultsch, D. F. (2004). Detecting malingering: A survey of experts' practices. *Archives of Clinical Neuropsychology*, 19, 465–473.
- Smith, G. P. (2008). Brief screening measures for detection of feigned psychopathology. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (3rd ed.), pp. 323–342. New York: Guilford.
- Solomon, Z., & Mikulincer, M. (2006). Trajectories of PTSD: A 20-year longitudinal study. *American Journal of Psychiatry*, 163, 659–666.
- Stevens, A., Friedel, E., Mehren, G., & Merten, T. (2008). Malingering and uncooperativeness in psychiatric and psychological assessment: Prevalence and effects in a German sample of claimants. *Psychiatry Research*, 157(1–3), 191–200.
- Stone, D. C., & Boone, K. B. (2007). Feigning of physical, psychiatric, and cognitive symptoms: Examples from history, the arts, and animal behavior. In K. B. Boone (Ed.), *Assessment of feigned cognitive impairment: A neuropsychological perspective*, pp. 3–12. New York: Guilford.
- Suhr, J. A., & Gunstad, J. (2007). Coaching and malingering: A review. In G. J. Larrabee (Ed.), *Assessment of malingered neuropsychological deficits*, pp. 287–311. New York: Oxford University Press.
- Suhr, J. A., & Gunstadt, J. (2000). The effects of coaching on the sensitivity and specificity of malingering measures. *Archives of Clinical Neuropsychology*, 15, 415–424.
- Sullivan, B. K., May, K., & Galbally, L. (2007). Symptom exaggeration by college adults in attention-deficit hyperactivity disorder and learning disorder assessments. *Applied Neuropsychology*, 14(3), 189–207.
- Summerfield, D. (2004). Cross-cultural perspectives on the medicalization of human suffering. In G. Rosen (Ed.), *Posttraumatic stress disorder: Issues and controversies*. (pp. 233–246). John Wiley & Sons.
- Taylor, S., & Asmundson, G. J. C. (2008). Posttraumatic stress disorder: Current concepts and controversies. *Psychological Injury and Law*, 1(1), 59–74. doi:1007/s12207-008-9009-0.
- 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(3), 455–464.
- Tombaugh, T. N. (1996). *The Test of Memory Malingering (TOMM)*. Toronto: Multi-Health Systems.
- Toomey, J., Sellbom, M., & Wygant, D. (2009). The utility of MMPI-2-RF validity scales (F-r and Fp-r) in identifying criminal defendants who feign psychopathology on the SIRS. In R. Frederick (Chair), *Validating tests without criterion groups*. Symposium conducted at the American Psychology Law Society meeting, San Antonio, TX.
- Twamley, E. W., Shami, S., & Stein, M. B. (2004). Neuropsychological function in college students with and without posttraumatic stress disorder. *Psychiatry Research*, 126, 265–274.
- Tyner, E. A. (2009). Validating the SIRS with MGVI using the MMPI-2 validity scales as estimators of base rate. In R. Frederick (Chair), *Validating tests without criterion groups*. Symposium conducted at the American Psychology Law Society meeting, San Antonio, TX.
- Vasterling, J. J., & Bailey, K. (2005). Neuropsychological finding in adults with PTSD. In J. J. Vasterling & C. R. Brewin (Eds.), *Neuropsychology of PTSD: Biological, cognitive, and clinical perspectives*, pp. 178–207. New York: Guilford.
- Vickery, C. D., Berry, T. R., Dearth, C. S., Vagnini, V. L., Baser, R. E., Baser, R. E., et al. (2004). Head injury and the ability to feign neuropsychological deficits. *Archives of Clinical Neuropsychology*, 19, 37–48.
- Victor, T. L., Boone, K. B., Serpa, J. G., Buehler, J., & Ziegler, E. A. (2008). Interpreting the meaning of multiple symptom validity test failure. *The Clinical Neuropsychologist*, 25, 1–17.
- Wetter, M. W., & Corrigan, S. K. (1995). Providing information to clients about psychological test: A survey of attorney's and law students attitudes. *Professional Psychology: Research and Practice*, 26, 474–477.
- Whitney, K. A., Davis, J. J., Shepard, P. H., & Herman, S. M. (2008). Utility of the Response Bias Scale (RBS) and other MMPI-2 validity scales in predicting TOMM performance. *Archives of Clinical Neuropsychology*, 23, 777–786.
- Widows, M., & Smith, G. P. (2005). *Structured Inventory of Malingered Symptomatology (SIMS) and professional manual*. Odessa: Psychological Assessment Resources.
- Williams, C. L., Butcher, J. N., Gass, C. S., Cumella, E., & Kally, Z. (2009). Inaccuracies about the MMPI-2 Fake Bad Scale in the Reply by Ben-Porath, Greve, Bianchini, & Kaufman. *Psychological Injury & Law* (in press).
- Wygant, D. B., Sellbom, M., Ben-Porath, Y. S., Stafford, K. P., Freeman, D. B., & Heilbronner, R. L. (2007). The relation between symptom validity testing and MMPI-2 scores as a function of forensic evaluation context. *Archives of Clinical Neuropsychology*, 22, 489–499.
- Wygant, D. B., Gervais, R. O., & Ben-Porath, Y. S. (2008). The MMPI-2 Restructured Form (MMPI-2-RF) validity scales: Association with various response bias criteria. Paper presented at the 28th Annual Meeting of the National Academy of Neuropsychology, New York, NY.

- Yanez, Y. T., Fremouw, W., Tennant, J., Strunk, J., & Coker, K. (2006). Effects of severe depression on TOMM performance among disability-seeking outpatients. *Archives of Clinical Neuropsychology, 21*(2), 161–165.
- Young, G., & Yehuda, Y. (2006). Understanding PTSD: Implications for court. In G. Young, A. W. Kane, & K. Nicholson (Eds.). *Psychological knowledge in court: PTSD, pain, and TBI*. New York: Springer Publishing Co.
- Young, G., Kane, A. W., & Nicholson, K. (eds). (2006). *Psychological knowledge in court: PTSD, pain, and TBI*. New York: Springer.
- Young, G., Kane, A. W., & Nicholson, K. (2007). *Causality of psychological injury: Presenting evidence in court*. New York: Springer Publishing Co.
- Youngjohn, J. R., Lees-Haley, P. R., & Binder, L. M. (1999). Comment: Warning malingerers produces more sophisticated malingering. *Archives of Clinical Neuropsychology, 14*, 511–515.