

The Fake Bad Scale (FBS)

Paul R. Lees-Haley, Ph.D., ABPP

The Fake Bad Scale (FBS) was designed to assist in the detection of personal injury malingering with the MMPI-2 (Lees-Haley, English, & Glenn, 1991). The expression “fake bad” was used because historically this phrase had been applied to invalid response sets on the MMPI. In hindsight, the FBS appears to be better thought of as an indication of an invalid response style, for example, in somatic injury cases, including brain injury claims, without necessarily assuming we always know why the patient is exaggerating or presenting an invalid self portrayal. Although originally intended for emotional distress injuries, the practice from which the scale was originally derived included many individuals with neurological, chronic pain, orthopedic and other physical injury claims, and a steadily growing body of peer-reviewed empirical research has indicated that the FBS is useful in the detection of somatic malingering, e.g., cases reporting physical injuries or pain, including neuropsychological cases (e.g., see Greiffenstein, Baker, Gola, Donders, & Miller, 2002; and Larrabee, 1996, 1997, 1998). Larrabee (1997) suggested that “somatic malingering should be considered whenever elevations on scales 1 and 3 exceed T80, accompanied by a significant elevation on the FBS) (p. 203).

Answer Key:

Following are the MMPI-2 items and scored direction of answering for the Fake Bad Scale (FBS):

Add one point if marked True: 11, 18, 28, 30, 31, 39, 40, 44, 59, 111, 252, 274, 325, 339, 464, 469, 505, 506

Add one point if marked False: 12, 41, 57, 58, 81, 110, 117, 152, 164, 176, 224, 227, 248, 249, 250, 255, 264, 284, 362, 373, 374, 419, 433, 496, 561

This information is for researchers and others interested in the Fake Bad Scale (FBS). We have received and continue to receive inquiries from European countries, Australia, New Zealand, South Africa, Canada, and other countries and from scholars and clinicians in many areas of the United States. If you are performing any new empirical studies using the FBS, please email (preferably, to expedite copying and pasting quotations) or mail a copy of your article to paul@lees-haley.com or Paul R. Lees-Haley, Ph.D., 3021 Panorama Drive, Huntsville, Alabama 35801 Fax 256-551-1036, Telephone 256-551-1024. (For that matter, if you are aware of others' work not cited herein and don't mind sending a note, please do that too, regardless of the findings.) Thank you!

Posthuma & Harper (1998) reported that the FBS “already has demonstrated efficacy in discriminating personal injury litigants who are malingering from other groups” and concluded that “Hence, in reports or expert testimony, forensic psychologists who use the FBS as one of a number of validity measures have a reasonable scientific basis for doing so” (p. 440).

Greiffenstein, Baker, Gola, Donders, and Miller (2002) studied 159 atypical minor head-injury litigants and 68 patients with documented moderate to severe closed head injuries, plus they made analyses using data from 50 moderate to severe closed head injury patients reported in Miller and Donders (2001). Greiffenstein et al. (2002) concluded, “When dealing with minor head-injury claimants seen more than a year after injury, we recommend using Lees-Haley’s original cutting score of 20.” However, recognizing that moderate to severe closed head injury claimants (GCS < 12) may produce more false positives with that cutting score, they advised, “As recommended by Miller and Donders (2001), a more conservative cutting score of 24 or greater should be used with the more severely injured” (p. 1598). Greiffenstein, Baker, Gola, Donders, and Miller (2002) also concluded “Our results indicate that the FBS appears to be superior to the standard MMPI infrequency scales in differentiation of atypical versus better-documented neurological injury when litigation status is held constant” (p. 1598). They added, “The FBS appears to be a valid measure of some form of spurious symptom production, at least in the context of litigated minor head injury. Not only did the FBS appear superior, but the traditional MMPI validity scales were insensitive to large group differences in symptom-history conformity... Our main conclusion is that the FBS scale is a promising tool in the evaluation of validity concerns in litigated neurological claims” (p. 1598).

Larrabee (1996, 1997, & 1998) found that the FBS was superior to “traditional MMPI/MMPI-2 malingering scales, such as F, in 12 litigants, with no electrophysiologic or neuroradiologic evidence of brain damage, who had objective evidence of malingering on neuropsychological tests (e.g., worse-than-chance performance)” (Larrabee, 1997, p. 203). Larrabee (1998) noted that the F Scale has only one item in common with the scales that are the primary ones associated with the somatic symptomatic complaints found most often in neuropsychological settings (Scales 1 and 3).

Millis, Putnam, & Adams (1995) compared 20 mild head injury litigants with external incentives to present impairments associated with their injuries to 20 outpatients with documented moderate and severe traumatic brain injuries. They found that “The FBS had the best diagnostic efficiency of the MMPI-2 scales selected for this study” (p. 3). They noted, “Our findings provide some qualified support for the future research with the FBS, although caution is in order” (p. 3). Posthuma & Harper (1998) studied N = 95 personal injury litigants and concluded “as indicated by our investigation and other recent research, the FBS raw score is likely to play a significant role in assessing malingering” (p. 442). Slick, Hopp, Strauss, & Spellacy (1996) found that Fake Bad Scale scores were correlated both with response time scores and number of items correct on the Victoria Symptom Validity Test (VSVT) for both easy and hard items.

In another study applying the FBS to neuropsychological cases, Putnam, Millis & Adams (1998) found that a cutoff score of 22 resulted in a 92.68% correct classification rate. In a study oriented toward potential utility for detecting invalid emotional distress claims, Lees-Haley (1992) examined data from 119 personal injury claimants. In this emotional distress sample, a cutoff of >23 for men and >25 for women correctly classified 75% of male and 74% of female spurious PTSD claimants and 96% of male and 92% of female controls. In this study pseudo-PTSD patients were defined as those who (1) claimed to be suffering a psychological injury (2) that was so severe that it was disabling (3) due to an experience that was entirely implausible as a candidate for PTSD criterion A in DSM-III-R and (4) scored T = 65 or higher on both PK and PS, the post-traumatic stress disorder subscales of the MMPI-2. The majority of personal injury plaintiffs do not respond at a level indicating invalid response styles on the FBS. In a study of a sample of 492 personal injury plaintiffs, the mean FBS score for plaintiffs was 20.8, the median score was 21, and 90% of plaintiffs scored less than 28 (Lees-Haley, 1997).

Larrabee (2000) studied 33 litigants claiming brain injury who had no medical, neurological or neurodiagnostic evidence of brain damage but who failed the Portland Digit Recognition Test (PDRT). This sample included 20 females and 13 males, mean age 42.3 (SD = 10.0), mean education 12.2 (SD = 2.3). Larrabee found that the FBS was more sensitive than the F Scale in detecting probable malingerers, consistent with previous findings by Millis et al. (1995), Putnam et al. (1998), and Larrabee (1998). Larrabee noted that future research should focus on identifying optimal cutoff scores to maximize sensitivity and specificity.

Martens, Donders, and Millis (2001) noted, “Several studies have demonstrated that the FBS has considerable sensitivity to exaggerated emotional or somatic distress in claimants of personal injury, and that it outperforms traditional MMPI-2 validity scales such as the F scale in this regard” (p. 3). They found that “individuals who were pursuing litigation were far more likely than persons who were not seeking financial compensation to have an FBS score suggestive of IRS [invalid response set]” (p. 8). More specifically, they found that “Individuals with mild THI [traumatic head injury] were more than five times as likely to have elevated FBS scores on the MMPI-2 as individuals with moderate-severe THI, and participants who were

pursuing litigation were also more than five times as likely to meet FBS criteria for IRS than participants who were not in litigation” (p. 12). After noting that invalid response sets are not always synonymous with malingering, they concluded “that the possibility of IRS cannot be ignored in the evaluation of patients with THI, and that the FBS index for the MMPI–2 has great potential in this regard (even though the associated term of ‘fake’ may be a misnomer in some cases)” (p. 15).

Greiffenstein, Baker, Axelrod, Peck and Gervais (in press, 2004) used empirical data to evaluate issues raised by Butcher, Arbisi, Atlis and McNulty (2003) in their critique of the FBS. Greiffenstein et al. noted that “Butcher et al. (2003) used a differential prevalence design to argue the FBS results in a high positive error rate. As pointed out by Larrabee (2003[c]), Butcher et al. did not offer any informal or formal external criterion for feigned mental illness, e.g., no subdivision into genuine and implausible groups. Instead, they uncritically accepted all profiles sent from across the nation as produced by “genuine” patients in the absence of any clinical history or contextual information. Hence, it is not clear how one can calculate false positive rates in the absence of dichotomous criteria or detailed clinical history. As argued by Rogers (1997), a differential prevalence methodology is a “poor substitute to the known-groups comparison” (pg. 17). The present known-groups design offering a reasonable external criterion for implausible symptoms shows the FBS is superior to traditional MMPI validity scales in a personal injury setting.”

Greiffenstein, Baker, Axelrod, Peck and Gervais (in press, 2004) concluded, “In conclusion, the present study adds further support to the notion the FBS is more sensitive to exaggerated psychological damages claims in an applied forensic setting than the MMPI-2 F-family. The FBS is a reliable tool in examining misrepresentation of nonpsychotic emotional problems (exaggerated pain, neurological and post-traumatic anxiety) in persons seeking compensation. Reasons for the effectiveness of the FBS in these settings may lie in development based on endorsements by actual participants in a forensic setting (Lanyon, 2001; Rogers, 1997).”

Larrabee (2003a) studied 26 persons identified as meeting criteria for definite malingered neurocognitive dysfunction (MND) and 29 persons with moderate or severe closed head injury (CHI). Larrabee found that the FBS was the most sensitive MMPI-2 scale for discriminating the malingerers from the head-injured person. Larrabee wrote, “The current results demonstrate two major findings: (1) the sensitivity of the FBS to malingering in neuropsychological settings is superior to any other MMPI-2 validity or standard clinical scale and (2) F, Fb, and F(p) are generally insensitive to malingering of neuropsychological symptoms” (2003a, p. 62). Larrabee pointed out that only the FBS had concurrent validity support by being significantly correlated with the PDRT. He noted, “At high levels of endorsement, FBS scores are associated with 100% positive predictive value (i.e., only malingerers score in this range)” (2003a, p. 65). He also noted other studies such as Meyers, Millis, and Volkert (2002), who found that no nonlitigating chronic pain patient scored higher than 29 on the FBS, and Ross, Millis, Krukowski, Putnam, and Adams (2004), who found that no CHI subject scored over 26.

Iverson et al. (2002) investigated the specificity of the FBS for identifying negative response bias in personal injury claimants. They studied inmate volunteers from a federal prison, medical outpatients, and inpatient substance abuse unit patients. They found that the original cutoff scores correctly identified the majority of inmates instructed to malingering psychopathology but recommended a revised cutoff to avoid unacceptably high rates of false positive classifications. Their cutoff scores produced false positives in the range of 8-24%.

Tsushima and Tsushima (2001) studied 120 patients involved in personal injury litigation, 208 clinical patients, and 43 normal participants. They compared the FBS with the F, Fb, Infrequency Psychopathology scale and Ds-2. They found that only the FBS significantly differentiated the litigating and clinical patients, and that both litigating and clinical patients scored higher than normals on the FBS. They concluded, "The content of the FBS, with several items from the Hypochondriasis and Hysteria scales, appears to enhance the FBS' ability to detect the somatic overreporting often observed with personal injury claimants. The authors suggest that the FBS may be a useful index of symptom magnification when employed within a comprehensive assessment of malingering in personal injury plaintiffs" (Tsushima & Tsushima, 2001, p. 205).

Meyers, Millis, and Volkert (2002) found that a combination of measures was more effective than any single measure for detecting malingering. They concluded that the FBS and other measures contributed to the effectiveness of their procedure.

Miller and Donders (2001) used the MMPI-2 in a study of 150 patients with traumatic brain injury and found, as have other authors, that individuals with mild traumatic brain injury showed paradoxically greater symptomatology than patients with moderate-to-severe traumatic brain injury. They concluded "Furthermore, specific actuarial criteria for possible symptom magnification (Fake Bad Scale) were met about twice as often in patients with mild THI [traumatic head injury] who were seeking financial compensation for alleged acquired dysfunction than in patients with mild THI without such external contingencies. It is concluded that the evaluation of persistent subjective complaints after THI should consider injury severity in concert with psychological and financial/motivational factors" (Miller & Donders, 2001, p. 297).

Ross, Millis, Krukowski, Putnam, and Adams (2004) used a receiver operating characteristic (ROC) curve analysis of the FBS, allowing them to determine the cutoff scores that maximized diagnostic efficiency. They investigated the utility of the FBS for detection of incomplete effort in mild head injury (MHI). They found that a cutoff score of 21 had a sensitivity of 90%, specificity of 90%, and overall correct classification rate of 90%. They noted that "traditional indices of faking bad on the MMPI-2, the F and F-K indices, fared relatively poorly by comparison and added no predictive power over the FBS" (page citation pending availability of hard copy publication of in press article). They added, "Finally, multivariate analyses revealed that although the FBS shares a number of items with Hs and Hy scales, the

FBS carried the majority of variance in predicting incomplete effort in our MHI sample” and reported, “Overall, these findings indicate that the FBS has high sensitivity and specificity in identifying incomplete effort in mild head injury.”

Ross et al. concluded that their results support the use of the FBS in the detection of incomplete effort in head injury, remarking, “In sum, these findings add considerable support to the use of the FBS in compensation-seeking patients with MHI. Indeed, the FBS appears to provide rather unique—and powerful—predictive power in identifying likely malingering in MHI, over and above traditional MMPI-2 validity indices and relevant clinical scales. Further, we found that a cutoff score only slightly higher than that originally reported by Lees-Haley et al. (1991) resulted in maximum sensitivity and specificity for this scale” (page to be cited when hard copy publication available). Ross et al. also concluded that the FBS is more indicative of invalid responding than somatoform disorder, and that it appears to detect somatic rather than psychiatric malingering.

Larrabee (2003b) compared traditional MMPI-2 validity scales, the FBS, and the Arbisi and Ben Porath Infrequency Psychopathology Scale (F(p)) in a study of 33 personal injury plaintiffs who had failed forced choice symptom validity testing and other measures of effort in patterns consistent with the Slick, Sherman, and Iverson (1999) criteria for definite and probable malingered neurocognitive deficits. He found that the FBS was more sensitive to symptom exaggeration than F, Fb, and F(p). Larrabee also found that the definite and probable malingerers produced elevated scores on MMPI-2 scales 1, 3 and 7 that were higher than those of non-litigating severe closed head injury patients and others such as, multiple sclerosis, spinal cord injury, chronic pain, and depression patients. Larrabee suggested that MMPI-2 profiles characteristic of physical injury malingering differ from those malingered psychopathology.

Larrabee also noted, “Meyers et al. (2002) found that only 16 of 100 non-litigating chronic pain patients had an FBS of 25-29 (84% specificity), and none had an FBS of 30 or higher. By contrast, 27 of 100 litigating chronic pain patients had an FBS of 25 to 29, and 15 had an FBS of 30 or more” (Larrabee, 2003b, p. 682).

Larrabee (2003d), studied atypical performance patterns on standard neuropsychological tests administered to 26 persons identified as meeting criteria for definite malingered neurocognitive deficit and 31 persons with moderate or severe closed head injury. Larrabee found that the FBS was the most sensitive MMPI-2 scale in discriminating the malingerers from the head-injured persons. In this two-part study involving initial and cross validation, he found that a cutoff score of >21 had a combined hit rate of 83.6%. This was comparable to the optimal cutoff score of >21 Larrabee had found in another study of exaggerated MMPI-2 symptom reports in personal injury litigants with malingered neurocognitive dysfunction (Larrabee, 2003a), and similar to the finding by Ross, Millis, Krukowski, Putnam, and Adams (2004), who found that scores of 21 or higher had a combined hit rate of 90%, and “nearly identical sensitivity and specificity, in discriminating 59 probable malingerers alleging mild CHI, from 59 non-litigating patients with moderate or severe CHI. Ross et al. found an optimal combined hit

rate of 91.5%, at a cut score of 23 or higher (sensitivity = .881, specificity = .949), with no non-litigating moderate/severe CHI subject producing an FBS greater than 26” (Larrabee, 2003a, p. 64).

Larrabee (2003a) further reported, “The combined error rate of .16 associated with an FBS cut score of 22 or higher in the present study, indicates this cut score will be effective (i.e. superior to base rate prediction alone) with base rates of malingering > 16% and < 84% (Gouvier, 1999). Per Ross et al., who obtained a lower combined error rate of .10, an FBS cut score of 22 or more will be effective when the base rate of malingering is > 10% and < 90% (Gouvier, 1999).” [Referring to Ross, Millis, Krukowski, Putnam, and Adams (2004) and Gouvier, 1999.]

Larrabee (2003c) stated, “An FBS score of 22 or more has good sensitivity and specificity in discriminating litigants with definite MND [malingered neurocognitive dysfunction] (Larrabee, 2003) or probable MND (Ross, Millis, Krukowski, Putnam, & Adams, [2004]) from patients sustaining moderate or severe closed head injury” (p. 396). In a study of pain questionnaires as potential tools for detecting exaggerated pain, Larrabee concluded that until further cross validation is available, pain questionnaires should be used in conjunction with other evidence of symptom exaggeration and poor effort, including measures such as the FBS, PDRT, and WMT.

Greiffenstein, Baker and Gola (1996) concluded that the FBS appears predominantly to measure spurious physical disability claims. This is consistent with Larrabee’s findings in head injury cases and Elhai et al.’s (2000, 2001) findings regarding malingering Posttraumatic Stress Disorder. It is also consistent with the conclusion of Ross, Millis, Krukowski, Putnam, and Adams (2004) that the FBS appears to detect somatic rather than psychiatric malingering.

Rogers, Sewell, and Ustad (1995) included the FBS in an analog malingering study using psychiatric outpatients in a within-subjects design. As Iverson and Binder (2000) pointed out, this study found low sensitivity of the FBS to the extreme exaggeration response set (mean F scale score was 109.6) of chronically mentally ill outpatients and a problematic rate of false positive classifications (19%) in the psychiatric patient sample. Iverson and Binder also concluded, “The FBS appears to be a promising MMPI-2 scale for identifying individuals who may be exaggerating symptoms. Additional research on the sensitivity, specificity, and predictive power of the scale will facilitate clinical use” (Iverson & Binder, 2000, p. *).

Bianchini, Houston, Greve, Irvin, Black, Swift, and Tamimie (2003), used the Slick, Sherman and Iverson criteria for identifying malingering to study four cases in the context of toxic tort litigation. They found elevated FBS scores in their four malingerers in neuropsychological toxic exposure cases. This is consistent with the view that the FBS is applicable in toxic cases but of course these were case reports rather than a controlled comparison study.

Grillo, Brown, Hilsabeck, Price, and Lees-Haley (1994) examined test results from 90 personal injury claimants to explore the relationship between personality disorders (Dependent, Histrionic, Compulsive, Schizoid, Schizotypal, Paranoid, Narcissistic, Borderline, Antisocial, Avoidant, and Passive-Aggressive) as assessed by the MCMI-II and response style measured by MMPI-2 validity scales (F, K, L, F-K, O-S, Es, and FBS). With the exception of the Dependent and Narcissistic scales, all personality disorder scales were found to have a significant relationship with validity indicators in the direction of faking bad. These results suggest that the presence of characterological factors (i.e., a personality disorder), rather than malingering, contributes to exaggerated results in a forensic setting. Implications for future research are addressed.

Fox, Gerson and Lees-Haley (1995) studied a sample of MMPI-2 data from worker's compensation and personal injury cases (N = 289) to examine the relationship of various indicators of exaggeration. Intercorrelations of the F, F-K, the MMPI Dissimulation Scale-revised (Ds-r), total of obvious minus subtle scales (O-S), Fake Bad Scale (FBS), VRIN, and TRIN were computed and the relative sensitivity of each score calculated using various cut-offs. Factor analysis suggests that malingering may take the form of inconsistent responding as well as symptom exaggeration. Patients evaluated at the request of plaintiff attorneys showed a seemingly greater degree of symptom exaggeration and inconsistent responding than did those referred by defense counsel.

Millis (personal communication, May 31, 1999) reported that for N = 23 patients with severe TBI (GCS 3-8) FBS scores were as follows:

Mean 15.087
Median 14.00
Std. Dev 5.51
Skewness 2.11
Kurtosis 5.76
Minimum 9.00
Maximum 34.00
Percentiles:
25 11.00
50 14.00
75 17.00

David Fox, Ph.D. of Consultants in Psychological Assessment in Glendale, California (fox@psychotherapy.org) has collected normative data on several populations described below (personal communication, August 8, 2000). Fox calculated the mean score for females in the MMPI-2 standardization sample as 13.8, and males as 11.7, and estimated the SD for females as 6.84 and the SD for males as 7.48 based on the samples below. Fox reported the following data for workers' compensation benefit applicants, HMO psychiatric clinic outpatients, persons convicted of domestic violence charges of widely varying degrees of violence who are on

probation and involved in involuntary treatment, and persons applying for work in positions such as police officer, priest or deacon or missionary, and a few in other occupations.

Fox Data

SETTING		FBS	AGE	SEX (M=1)	EDUC
Workers Comp	Mean	22.48	39.3	1.61	12.6
	N	258	258	258	258
	Std. Deviation	7.12	10.27	.49	4.13
Psych Patients	Mean	18.70	37.14	1.53	12.77
	N	88	88	88	88
	Std. Deviation	5.62	14.56	.50	5.35
Probation	Mean	13.91	34.53	1.36	11.26
	N	45	45	45	45
	Std. Deviation	5.35	8.91	.48	5.21
Job Applicants	Mean	12.73	40.62	1.15	8.64
	N	132	132	132	132
	Std. Deviation	3.02	12.19	.36	8.0
Total	Mean	18.65	38.87	1.46	11.5
	N	523	523	523	523
	Std. Deviation	7.28	11.59	.50	5.9

Following are data from a handout accompanying a lecture entitled, "Overview of the MMPI-2 Fake Bad Scale" by Dr. Grant Iverson in September 2000, using content from Iverson and Binder (2000).

Table 1. Fake Bad Scale performances in known groups.

Group	Number of Subjects	Average Score (Mean)	Standard Deviation (SD)
--------------	---------------------------	-----------------------------	--------------------------------

Normal Control Subjects

MMPI-2 Normative Sample-Males	1,138	11.7	3.8
MMPI-2 Normative Sample-Females	1,462	13.8	4.1
Federal Prison Inmates-Minimum Security-Males	25	9.8	4.1

Medical Patients

Moderate-Severe Traumatic Brain Injury	20	16.1	4.9
Medical Outpatients-Transplant Candidates-Males	20	15.3	7.6

Psychiatric Patients

MMPI-2 Psychiatric Sample-Estimated scores for males ¹	---	16.9	---
MMPI-2 Psychiatric Sample-Estimated scores for females ¹	---	19.1	---
Patients with Chronic Psychiatric Problems	42	22.1	5.8
Male Veterans-Inpatient Substance Abuse Unit	25	15.6	6.4

Child Custody Litigants²

No Alleged Abuse-Males	---	13.8	3.0
No Alleged Abuse- Females	---	13.8	3.0
Alleged Physical Abuse-Males	---	13.6	3.1
Alleged Physical Abuse-Females	---	15.1	3.6
Alleged Sexual Abuse-Males	---	13.5	3.1
Alleged Sexual Abuse-Females	---	16.0	3.5

Personal Injury Litigants

Presumed Credible Litigants	20	15.7	4.1
Head Trauma, Chronic Pain, or Both	95	22.3	6.9
Unselected, Consecutive Litigants	492	20.8	6.7

Analog Malingerers

Medical Outpatients-Simulate emotional distress from MVA	16	26	6.7
Medical Outpatients-Simulate emotional distress from toxic exposure	15	21	9.6
Medical Outpatients-Simulate emotional distress from job stress	36	27	8.0
Medical Outpatients-Total analog malingerers	67	25	8.5
Patients with Psychiatric Problems-Malinger severe psychiatric problems	42	24.8	5.7
Federal Inmates-Malinger severe psychiatric problems	25	21.6	5.8

Suspected Malingerers

Mild Head Injury Litigants	20	29.0	5.1
----------------------------	----	------	-----

¹The total number of psychiatric patients reported in the manual was 540. ²In Posthuma & Harper (1998), there were 188 individuals involved in custody evaluations.

Comments on the 2003 Archives of Clinical Neuropsychology article by Butcher, Arbisi, Atlis, and McNulty (Butcher, J. N., Arbisi, P. A., Atlis, M. M., & McNulty, J. L. (2003). The Construct Validity of the Lees-Haley Fake-Bad Scale (FBS): Does this Scale Measure Somatic Malingering and Feigned Emotional Distress? Archives of Clinical Neuropsychology, 18, 473-485.). (Note: the following is mostly drawn from Lees-Haley and Fox, 2004, a shortened and somewhat edited version of the following comments. Also see another rebuttal of Butcher et al. by Greve & Bianchini, 2004).

Butcher, Arbisi, Atlis, and McNulty (2003) attempted to critique the FBS but ignored the literature, misstated the purpose of the FBS, misunderstood the purpose of the FBS, and applied it so inappropriately that their critique failed to provide any real rebuttal of the FBS. In fact, their findings provided evidence *supporting* the validity of the scale, although they do not appear to have recognized the implications of their own data.

Butcher et al. indicated that the purpose of their study was to evaluate the structure of the FBS, investigate the relationship of the scale to other MMPI-2 measures and evaluate the proportion of individuals classified as “somatic malingerers.” Their methodology and reporting

are very unusual for a research study. Not even basic descriptive statistics were reported, e.g., means, standard deviations, or ranges. Five of the six samples they studied were not in the setting for which the scale was designed, and only one – their smallest sample -- was personal injury litigants. As Dr. Roger Greene noted, applying the FBS in other contexts is analogous to applying the Marital Distress Scale to single persons. The authors selected cases from psychiatric inpatients, correctional facility participants, general medical patients, and a chronic pain sample in a National Computer Systems archival database containing 119,672 cases. Butcher et al. excluded 89,675 cases (75%) without explanation and without stating the exclusion criteria. They excluded another 10,881 cases because their profiles appeared invalid, leaving 19,116 cases to study. Excluding 10,881 cases with invalid profiles is precisely the incorrect methodology for evaluating a scale designed to detect invalid profiles. This is analogous to testing the MMPI-2's capacity for detecting depression after first removing depressed patients from the sample studied. The selection bias in this methodology is calculated to produce a misleadingly low impression of the scale being evaluated.

Although the authors reported an attempt to include only valid profiles for analysis, in fact, the cutoffs did not eliminate protocols that were likely invalid. The exclusion criteria were selected without any justification. Inspection of their cutoff criteria reveals that the cutoff scores were so unusually high that many people normally thought to be exaggerating their symptoms would not be excluded. For example, the authors used a cutoff of $T = 110$ for the F and F(b) scales. As noted by Greene (2000, p. 70), such scores represent “extreme” distortion and scores on these scales as low as 81 may represent invalid profiles due to symptom exaggeration. Similarly, Graham (1990) indicates that F scores as low as 80 indicate a person “may be malingering” (p. 26). In this article, Butcher et al. used dramatically different standards to criticize the FBS than Butcher and his colleagues published in the MMPI-2 manual and its revision (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989; Butcher, Graham, Ben-Porath, Tellegen, Dahlstrom & Kaemmer, 2001). They provided no explanation or rationale for using different standards when evaluating the FBS than Butcher et al. recommended in the MMPI-2 manuals. The first MMPI-2 manual indicates malingering as the first in the list of suspected sources of elevation of profiles with an F scale over 70, and says that such scales are of questionable validity. The revised MMPI-2 manual changes the interpretation without citing new evidence, and makes the odd claim that in non-clinical settings, F scale scores over 65 may be both exaggerated and valid! Nevertheless, the authors treat their data set as if only valid cases remain, an erroneous assumption.

Butcher et al. selected several research samples and assumed that they possessed certain characteristics without citation or reasoning. For example, they present no evidence to indicate the level of medical complaints, psychiatric complaints, or malingering in their samples from the Inpatient Psychiatric Group, Chronic Pain Program or General Medical Program. Instead of external criteria, they relied on speculative assumptions. Because no independent measure of symptom exaggeration (e.g., symptom validity test, clinical estimate, test score or external variable) was provided for any of these groups, it is impossible to evaluate the accuracy of the FBS with their methodology.

Butcher et al. then address whether the FBS represents a single homogeneous dimension. This attempt is pointless because Lees-Haley et al. (1991) from the beginning clearly defined the FBS as measuring an invalid response style that is a mixture of both faking good and faking bad behaviors. It makes no sense statistically to expect the FBS to correlate with a mixture of inversely correlated measures. Not surprisingly, the authors concluded that the FBS does not measure a single dimension, just as Lees-Haley et al. explained in 1991.

Despite the claims of Butcher et al., the data they report actually support the construct validity of the FBS. They report that the highest alpha coefficient obtained in their samples (.85) was in the Personal Injury group, which is precisely the group that is expected to have the highest rate of symptom exaggeration. This is the group for which the FBS was designed. Note that their FBS alpha is higher than the alpha for F and F(b) and higher than most of the clinical scales reported in the MMPI-2 manuals.

Butcher et al. question the relationship of the FBS to existing MMPI-2 scales. They note the relative lack of item overlap between the FBS and the F, F(b) and F(p) scales. They use this as evidence that the FBS does not reflect the same test taking strategy as do these existing MMPI validity scales. What they do not take into account is that the FBS shows substantial correlation with these scales but contributes unique variance (Fox, Gerson, and Lees-Haley, 1995). Butcher et al. then review individual items and claim, without specific citation, that those items indicate somatic problems and not necessarily somatic malingering. Although they characterize the FBS as a somatic symptom scale, according to their own analysis this is incorrect. In Table 1, on page 475, only 14 of 43 items are somatic and a larger number, 15, are denial of deviant attitudes or beliefs. Consequently, the data they present actually *supports* the construct validity of the FBS as a measure of symptom over-reporting in medical-legal cases. Based on this contradictory reasoning, Butcher et al. conclude that the developers of the FBS have only demonstrated that the scale measures physical symptoms and not necessarily “false claims.”

Criticism was made of the fact that the FBS has lower correlations with traditional validity scales than with many clinical scales. Examination of the data they present undermines this conclusion. For example, the correlation between FBS and F and F(b) for the Personal Injury group is .533 and .551, respectively. These are substantial correlations. There are even higher correlations in this group between the FBS and some of the basic clinical scales but this discrepancy is expected because the FBS takes a different approach to identifying symptom invalidity than the traditional validity scales. It is common knowledge that if someone exaggerates their problems on the MMPI-2 then they also show elevations on clinical scales. Later in the article, they reiterate that the FBS does not work in the same fashion as the traditional measures of symptom exaggeration, in effect arguing that if it doesn't do it the old-fashioned way it isn't worth doing at all. In addition, they ignore the previously published data revealing substantive but not extreme correlations between the FBS and these other measures in other samples (e.g., Fox, Gerson & Lees-Haley, 1995; Lees-Haley, 1997). One of the reasons the FBS has value is that it has been shown to be more sensitive to certain forms of symptom

exaggeration than the more traditional scales (Larrabee, 1998). Judging the FBS by its similarity to the F and F(b) scales undercuts the very purpose of developing new scales that remedy current scale deficiencies.

Finally, the authors conclude that the FBS over-predicts malingering in clinical and forensic samples, but offer no data to support this claim. Because no rate of malingering and no independent measure of symptom exaggeration were provided for any of their samples, it is impossible to evaluate the accuracy of the FBS with their methodology. The authors ignore the fact that the FBS is designed for civil forensic cases, so measuring the rates of “malingering” in other samples is interesting but not a refutation of the FBS. When complaining that the FBS over-predicts malingering, Butcher et al. ignored the extensive research on the incidence of malingering and relied on their speculation. Instead of utilizing independent criteria to identify malingering, they used unusually high cutoffs on traditional scales to define valid and invalid profiles. In contrast, other researchers have used a variety of external criteria to establish the validity of the FBS (e.g., see Greiffenstein, Baker, Gola, 1996; Greiffenstein, Baker, Gola, Donders, & Miller, 2002; Iverson, Henrichs, Barton, & Allen, 2002; Larrabee, 2003a, 2003b; Larrabee, 1998; Larrabee, 2000; Martens, Donders, and Millis, 2001; Meyers, Millis, and Volkert, 2002; Miller and Donders, 2001; Millis, Putnam, & Adams, 1995; Posthuma & Harper, 1998; Putnam, Millis & Adams, 1998; Ross, Millis, Krukowski, Putnam, & Adams, 2004; and Tsushima and Tsushima, 2001). Because Butcher et al. neglected to provide any external symptom exaggeration validity data and failed to review the literature, they confuse their speculation and their casual use of convenience data with meaningful validating criteria. Butcher et al. make the assumption that their samples, after excluding grossly exaggerated profiles, represent true psychological and physical disorders. On the contrary, there is ample evidence (e.g., Gervais et al., 2001) that both emotional and cognitive symptomatology can be grossly exaggerated in patients complaining of chronic pain and other medical conditions.

Butcher et al. then engage in wide-ranging speculation regarding the meaningfulness of their results. For example, they complain that the relatively high rate of malingering detected by the FBS in the VA sample is inaccurate in assessing PTSD “since the base rate of genuine PTSD in VA settings should be greater than other settings given the high rate of traumatic experiences associated with military service.” The authors inexplicably equate being a psychiatric inpatient in the VA with having PTSD, and ignore the extensive literature indicating that malingering of PTSD is a frequent problem in VA settings (e.g., see Campbell & Tueth, 1997; Frueh, Gold, and de Arellano, 1997; Frueh, Hamner, Cahill, Gold, & Hamlin, 2000; Frueh, Smith, & Barker, 1996). Prior to studying the Veteran’s Administration sample, 36% of the cases were excluded because their profiles were grossly invalid. Once again, this methodology is analogous to excluding depressed patients before evaluating the sensitivity of a depression scale, and it is calculated to produce biased findings consistent with the authors’ conclusions. The purpose of the FBS is to assist in detecting invalid response styles. The authors did not collect data regarding the actual rate of PTSD in order to address the accuracy of their speculation.

They further speculated that the FBS is more associated with the expression of

psychopathology in which physical symptoms are experienced. The data belie their speculation. If actual physical symptoms accounted for the greatest variance in the FBS, it would be anticipated that the Chronic Pain program and General Medical groups should have rates of “malingering” comparable to those in the Personal Injury group. Instead, just the opposite is the case. In precisely the group that is likely to have the highest rate of symptom exaggeration, there is the highest rate of classification of symptom over-reporting. Further, it cannot be assumed that all of the patients in the Chronic Pain and General Medical samples were not exaggerating. The actual rate of symptom exaggeration or malingering in these samples is unknown and the authors present no data to address this issue. The authors continue to speculate that the FBS represents physical complaints and that those with multiple sclerosis or neuralgia are likely to produce extreme elevations on the FBS but they present absolutely no data to support this speculation. They go on to state that mental health patients who have psychologically based disorders or those who have chronic medical conditions are likely to have elevated FBS scores without presenting any data whatsoever. Once again, the authors have not obtained any validating criteria as has been done in previous studies regarding the FBS.

Butcher et al. expand their speculation, stating that a “large number of general patients” from the various samples are unlikely to be malingering but present no justification whatsoever for their speculation and they ignore the substantial literature indicating that malingering is common in diverse samples. Their complaint that the FBS finds an “unacceptably” high rate of malingering -- between 2.4% and 30.6% -- betrays a dramatic lack of awareness of the current literature concerning base rates of malingering. With no empirical basis at all, Butcher et al. arbitrarily presume malingering to be less common than voluminous empirical research literature has indicated. Although estimates of the incidence of malingering have varied from less than 5% to over 60% of personal injury patients referred for neuropsychological evaluation (Hayes, Hilsabeck, & Gouvier, 1999), a general trend has been for estimates to fall at around 30% or more as a reasonable base rate for estimated malingering among an outpatient population referred for evaluation in the context of compensation seeking, which is the context for which the FBS was designed (Allen, Conder, Green, & Cox, 1997; Binder, 1997; Frederick, Sarfaty, Johnston, & Powel, 1994; Gervais et al., 1999; Gervais, Allen, Green, & Cunningham, 1998; Greiffenstein, Baker, & Gola, 1994; Gouvier, Lees-Haley, & Hammer, 2003; Green, Rohling, Lees-Haley, & Allen, 2001; Heaton, Smith, Lehman, & Vogt, 1978; Larrabee, 2000; Millis, 1992; Trueblood & Schmidt, 1993). Mittenberg, Patton, Canyock, & Condit (2002) studied 33,531 cases involved in personal injury, disability, criminal or medical matters and found that 29% of personal injury, 30% of disability, 19% of criminal, and 8% of medical cases involved probable malingering and symptom exaggeration. Mittenberg et al. found that 39% of mild head injury, 35% of fibromyalgia/chronic fatigue, 31% of chronic pain, 27% of neurotoxic, and 22% of electrical injury claims resulted in diagnostic impressions of probable malingering. In their study, diagnosis was supported by multiple sources of evidence, as distinct from no evidence at all in the Butcher et al. study, in which malingering was simply presumed absent or unlikely. Other authors have documented that the reality of the base rate of malingering falls in a range Butcher et al. find “unacceptable.” For example, based on a review of a variety of studies, Gouvier, Lees-Haley, and Hayes-Hammer estimated the base rate of malingering to be 30% or more (2003).

Butcher et al. complain that the FBS scale cannot be used until genuine psychiatric illness, emotional distress or somatic problems are ruled out. Once again their lack of understanding of malingering is dramatic. It is preposterous to argue that malingering cannot occur unless the plaintiff is mentally healthy (i.e., has no genuine mental disorder). Moreover, the data presented by Butcher et al. in no way support this conclusion. They inexplicably state that personal injury claimants have emotional distress and imply that this population is essentially identical to a regular psychiatric population. This is pure speculation that is contrary to the findings of the last decade and a half of research illustrating that claimants behave differently on validity scales.

Finally, in a bit of social commentary, Butcher et al. complain that the scale classifies women too often as malingerers, apparently implying that misdiagnosis of men is of lesser consequence. What they ignore, in addition to the previous published research and the earlier part of their own article, is that Lees-Haley and others have suggested different cutoffs for men and women, which is standard practice for MMPI-2 scales. Even if some differential effect of cutoffs could be demonstrated, it cannot be assumed that this represents unrealistic bias. The MacAndrews alcoholism scale, for example, classifies more men than women as at risk for substance abuse (Butcher et al., 2001). In a fit of righteous indignation, the authors claim that victims of PTSD will suffer further if they are falsely labeled as malingerers, expressing no concern at all for the victims of malingerers who are endorsed by psychologists as having valid injuries. As noted above, Butcher et al. in no way have established the true or false positive rate in any of their samples. While implying a looming cataclysm, the authors extol the reader to reject the FBS because it would destroy the therapeutic alliance and deny needed care. It should be evident to the reader that this portion of the article is politics, not science. Remarkably, Butcher et al. criticize an MMPI-2 scale as biased because men and women may respond differently, knowing full well that different norms have been used for men and women for over a half century. Note that Dr. Butcher uses different norms for men and women in the interpretive reports he sells through NCS. If Butcher et al. are serious about their claim that the use of different norms for men and women for the FBS means the measure is biased, then they should rethink the use of the MMPI-2 in any context, because if their logic is correct then the MMPI-2 is a discriminatory instrument. There obviously would be numerous legal implications for publishers and users of the MMPI-2 if their logic were found to be true in court. However, this superficial criticism by Butcher et al. ignores the fact that differential norms are common for various demographic characteristics, including gender, throughout psychology and neuropsychology. It also ignores the fact that many of the empirical studies cited herein make no recommendation of differential cutoffs for men and women.

Development of useful psychological scales is an ongoing process. Butcher et al. have sidetracked this process by misunderstanding the basis of the FBS, misapplying it to inappropriate populations and misinterpreting their own data. Does their study have any value? Yes -- it provides support for the FBS by showing that, as expected, patients with the highest risk of symptom over-reporting score higher on the FBS than do medical, pain or psychiatric patients.

Comments on the article by Alison S. Bury and R. Michael Bagby entitled “The Detection of Feigned Uncoached and Coached Posttraumatic Stress Disorder with the MMPI-2 in a Sample of Workplace Accident Victims.” (The following comments explain why the Bury and Bagby (2002) study is uninterpretable and neither provides support for the scales the authors advocate nor provides evidence arguing against the use of the scales the authors reject or minimize i.e., the results of the Bury and Bagby study are inconclusive regarding any of the validity scales.)

Bury and Bagby (2002) performed a study to determine whether a number of MMPI-2 validity scales could distinguish between 22 year old ($SD = 5.2$) college students faking Posttraumatic Stress Disorder (PTSD) and 40 year old ($SD = 9.34$) claimants who were classified as having “bona fide PTSD” based on a self report procedure administered by clinical psychology graduate students and a master’s level psychometrist whose “interrater agreement was not formally assessed” (p. 474). In lieu of measuring interrater reliability at diagnosing PTSD, the raters were supervised by a registered clinical psychologist. The self-report procedure used as a basis for diagnosing PTSD was the Structured Clinical Interview for DSM-IV Axis I Disorders I/P Form (SCID-I/P), i.e., the patient form of the SCID (First, Spitzer, Gibbon & Williams, 1995).

With all due respect, one psychologist’s clinical opinion, no matter how esteemed, is not a generally accepted criterion for establishing reliability or validity of diagnostic procedures in a research paper, and supervision by a psychologist is not a substitute for normal methods of measurement of interrater reliability. Specifically in connection with the diagnostic procedure used in this paper, note that according to the SCID senior author Michael B. First, in association with the other authors of the SCID, “The range in reliability [of the SCID] is enormous, depending on the nature of the sample and research methodology” (First, 2002, p. 1). No reliability data were provided for the methodology used in the study by Bury and Bagby.

As the American Psychological Association Board of Scientific Affairs Task Force on Statistical Inference has emphasized, “Naming a variable is almost as important as measuring it” and a phrase such as “retrospective self-report of childhood sexual abuse” is preferable to “childhood sexual abuse” because “Without such precision, ambiguity in defining variables can give a theory an unfortunate resistance to empirical falsification. Being precise does not make us operationalists. It simply means that we try to avoid excessive generalization” (Wilkinson & Task Force on Statistical Inference, 1999, p. 5). Following this recommendation, instead of “bona fide PTSD claimants” (p. 475), a more descriptive and precise characterization of the people studied by Bury and Bagby would be “compensation seeking claimants making retrospective self reports of PTSD symptoms.”

The development of symptom validity measures requires the use of criterion groups that are genuine (e.g., see Greiffenstein, Baker and Gola, 1994; Kazdin, 1998). Therein lies another

major weakness with this study: There is clear evidence raising questions about the authenticity of the PTSD group. The entire PTSD sample consisted of claimants seeking compensation through Worker's Compensation claims. The authors' confidence in "bona fide" PTSD symptoms rests on self-reported data in a context fraught with biased reporting and despite evidence of validity problems indicated by their MMPI-2 data. Compensation seeking is a notorious confounder in psychological claims (e.g., see meta-analyses by Binder & Rohling, 1996; Rohling, Binder, & Langhinrichsen-Rohling, 1995). Exaggeration and malingering in compensation contexts have been documented as problematic specifically with respect to PTSD cases (e.g., see Campbell & Tueth, 1997; Frueh, Gold, and de Arellano, 1997), and specifically in cases studying the MMPI-2 (e.g., see Berry, Wetter, Baer, Youngjohn, Gass, Lamb, Franzen, MacInnes, & Bucholz, 1995; Frueh, Hamner, Cahill, Gold, & Hamlin, 2000; Frueh, Smith, & Barker, 1996; Youngjohn, Davis, & Wolf, 1997). Malingering is so frequently an issue in the diagnosis of PTSD in forensic and compensation settings that it is explicitly named in the differential diagnostic process in DSM-IV (APA, 1994) and DSM-IV-TR (APA, 2000).

Thus it is a serious problem to presume the validity of the SCID data, as if it were difficult or unlikely for claimants to endorse PTSD symptoms in the absence of actually suffering these symptoms. The SCID has no built-in safeguards against response sets such as "yea saying" or malingering, and untrained persons asked to guess PTSD symptoms are usually able to do so. For example, Burges and McMillan (2001), replicating similar findings by Lees-Haley and Dunn (1994), found that the vast majority of untrained subjects were able to claim the symptoms of PTSD merely through uninformed guessing. Moreover, it may be naïve of us as experts to presume that Bury and Bagby's claimants were naïve, as they were enmeshed in the workers' compensation system, and therefore were likely exposed to symptom-related information during the course of their claim. Bury and Bagby were aware of this problem with their population and correctly pointed out that a claimant motivated to do so "can easily learn what symptoms must be reported to qualify for the diagnosis" (p. 472).

The "bona fide" group's MMPI-2 scores cast doubt on the authenticity of their symptoms. Table 1 of the study shows the majority of the "bona fide PTSD claimants" scored in ranges associated with exaggeration on MMPI-2 validity scales ($F \underline{M} = 75.6$, $\underline{SD} = 22.5$; $Fb \underline{M} = 82.4$, $\underline{SD} = 24.5$) (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989; Butcher, Graham, Ben-Porath, Tellegen, Dahlstrom & Kaemmer, 2001). Note too that almost half of the "bona fide claimants" scored in a range on the F-K that Graham says indicates the need to consider possible malingering (1999). The fact that Bury and Bagby did not use the SCID version with psychotic screening implies that psychotic features cannot be used as the explanation for their elevated F family scores, and they explicitly ruled out illiteracy. In short, the number of elevated F, Fb, and F-K scores indicates that the majority of the MMPI-2 profiles produced by the "bona fide PTSD claimants" are exaggerated rather than accurate measures of psychological symptomatology.

The SCID was administered only to the PTSD group, and not the analogue faking group. Because the study did not include a control group, the reader has no data with which to assess how effective the SCID might be at classifying genuine versus simulated PTSD, even if one were

to presume the PTSD claimants were bona fide. This is a fatal error. All of the authors' conclusions about MMPI-2 validity scales rest on the assumption that simulators and bona fide PTSD victims were accurately distinguished and sorted into two separate groups for comparison. Because the SCID was not administered to the coached group, there is no way to determine whether the SCID discriminated claimants from students, and no way to make any comparison of the base rate of PTSD symptoms among the groups studied. As in any study reaching conclusions about differences between two groups, measuring both the control group and the experimental group is essential.

Because this study used an older version of the SCID, some participants in the PTSD group may not meet modern criteria for PTSD even if one were to presume hypothetically that their self-report data were all valid (an assumption with which not even the authors agree). Presently, Criterion A, the "gatekeeper" criterion, includes two requirements: "(1) the person experienced, witnessed, or was confronted with an event or events that involved actual or threatened death or serious injury, or a threat to the physical integrity of self or others (2) the person's response involved intense fear, helplessness, or horror" (APA, 1994, 2000). Looking at the methods section, the reader learns that only 8% of those studied experienced something that clearly appeared to meet Criterion A: Witnessing another's death. The remaining stressors were too vague and general to draw any conclusions. Does "assault" mean being stabbed multiple times, or does it mean being pushed against a wall during a brief argument? These distinctions are critical for passing Criterion A before moving on to the symptom criteria.

There is further cause for concern about using the participants in this study as a basis for generalizations. Although one would expect college students to have a lower than average number of lifetime traumatic experiences than the total adult population due to their age, 62% of this particular group had a past traumatic event that the authors indicated would satisfy DSM-IV Criterion A for PTSD. This is a higher lifetime prevalence than is reported in epidemiological studies of adults, more than 75% of whom were older than the average college student in this study. For example, Kessler, Sonnega, Bromet, Hughes, Nelson, and Kessler (1995) found that 60.7% of men and 51.2% of women reported at least one such stressor, a finding similar to those of previous researchers (e.g., see Kilpatrick & Resnick, 1992).

An N of 61 bona fide PTSD patients is also striking in view of the best available research on causation of PTSD, e.g., findings that even large scale disasters only produce a modest increase in the incidence of psychopathology (e.g., see Gist & Lubin, 1999; Rubonis & Bickman, 1991), and the fact that the majority of PTSD victims fully recover within a few months of their trauma (APA, 1994). The conclusion that a high rate of Criterion A trauma in the college students did not cause them to have any PTSD present at the time of the study, despite their average age of only 22 years, juxtaposed with the conclusion that a large N of claimants' Criterion A trauma did cause PTSD, despite the evidence of exaggeration on the MMPI-2, reinforces the conclusion that the claimants' self-report data are being relied upon to an extent that may merit reconsideration. It would have been informative to know the total number of claimants from which "66 consecutive PTSD claimants who completed the MMPI-2 and the interview segments

of the assessment” (p. 474) were drawn, in order to better understand the base rate at which this practice assigns this diagnosis in relation to other diagnoses. Further concern about the possibility of excessive false positive diagnoses in the “bona fide” group arises from the high rate of diagnosis of comorbidity. The *lifetime* prevalence for even a *single* mental disorder is substantially lower than the comorbidity (two or more mental disorders) findings in this study, which most likely covered only a few years of the lives of the participants, since they were making workers’ compensation claims (e.g., see Kessler and Walters, 2002).

Bury and Bagby in effect admit that the sample they characterize as “bona fide PTSD claimants” included data produced by malingering. For example, they wrote, “it is almost certain that the PTSD sample included some people who at the very least exaggerated their symptoms. This is an inherent limitation in experimental studies that use compensation-seeking comparison samples, because many of the patients may be responding similarly to those research participants instructed explicitly to feign” (p. 482). They also stated “all PTSD claimants were seeking continuation or reinstatement of compensation. In the context of these incentives, symptom exaggeration is expected, and the comparatively low classification rates may be a result of the presence of individuals in the workplace PTSD comparison sample who were actually exaggerating or malingering their condition. Indeed, there is some support for this observation in the current study...” (p. 483). These are confusing admissions. Given these observations indicating invalidity in the PTSD claimants, why were the claimants repeatedly characterized as “bona fide”? And why was there no SCID comparison group? The authors went on to explain that the range of scores was comparable between the bona fide claimants and the fakers, and that the “bona fide claimants” displayed an upper limit of $T = 120$ on the Fp scale. They conclude, “This suggests that some of the [“bona fide”] PTSD claimants were likely exaggerating their symptoms” (p. 483). To include data with such an extreme upper limit of $T = 120$ in the “bona fide” (valid) group seems to suggest that the Fp scale is not a measure of validity, or the data from claimants with elevated Fp scales would have been excluded from the “bona fide” sample. However, the authors include the Fp in the family of F scales and single it out as especially valid when they say, “The family of F scales (i.e., F, Fb, Fp), particularly Fp, produced consistently high rates of positive and negative predictive power” (p. 472).

In other words, this study was not based on group of claimants who can plausibly be presumed to have bona fide PTSD. The unintended product of relying so heavily on self-reported symptoms was contradictory and circular reasoning. The solution to this problematic design is to conduct a group comparison between nonlitigating persons seen in the clinic following unambiguous Criterion A trauma, versus a sample of compensation seeking PTSD claimants. For example, similar methodology was used to good effect in Greiffenstein, Baker, Gola, Donders, and Miller (2002). They constructed a group of minor head injury litigants with atypical symptom histories: Delayed onset of memory complaints, progressive or fluctuating symptom course, and autobiographical, procedural and semantic memory loss. These authors’ findings were quite different than those in the present study. For example, their atypical group scored much higher on the Fake Bad Scale (FBS) than persons with severe closed head injuries who were undergoing rehabilitation and were not in litigation, and they found the MMPI-2 F

scale insensitive to group differences. However, Greiffenstein et al. studied a more specialized form of trauma – head injury – rather than the relatively broad range of traumas seen in Bury and Bagby’s PTSD research.

In conclusion, given the questionable validity and failure to assess the interrater reliability of the process for defining “bona fide” PTSD, the lack of a control group on the most critical measure in the study (the SCID), the evidence of invalidity of the “bona fide” group (much of which the authors admit), the vagueness of definition of Criterion A experiences, the evidence of false positives, and the excessive and uncritical reliance on self-reported symptoms as a basis for diagnosing PTSD in the claimants juxtaposed with finding an absence of PTSD caused by Criterion A stressors in the student group, the results of this study are uninterpretable. If anything, to the extent that MMPI-2 validity scales are capable of detecting PTSD faking, many of the members of the “bona fide” sample should have been treated as members of a malingering comparison sample. In summary, the findings of this study neither provide support for the scales the authors advocate nor provide evidence arguing against the use of the scales the authors reject or minimize. The results of the Bury and Bagby study are inconclusive for any of the validity scales.

References

- Allen, L., Conder, R.L., Green, P., & Cox, D.R. (1997). CARB’97 Manual for the Computerized Assessment of Response Bias. Durham, NC: CogniSyst.
- American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington DC: Author.
- American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders (Text Revision) (4th ed.-TR). Washington DC: Author.
- Berry, D., Wetter, M., Baer, R., Youngjohn, J., Gass, C., Lamb, D., Franzen, M., MacInnes, W., & Bucholz, D. (1995). Overreporting of closed head injury symptoms on the MMPI-2. Psychological Assessment, *7*, 517-523.
- Bianchini, K. J., Houston, R. J., Greve, K. W., Irvin, T. R., Black, F. W., Swift, D. A., and Tamimie, R. J. (2003). Malingered Neurocognitive Dysfunction in Neurotoxic Exposure: An Application of the Slick Criteria. Journal of Occupational and Environmental Medicine, *45*, 1087-1099.
- Binder, L. (1997). A review of mild head trauma. Part II. Clinical implications. Journal of Clinical and Experimental Neuropsychology, *19*, 432-457.
- Binder, L. M., & Rohling, M. L. (1996). Money matters: Meta-analytic review of the effects of financial incentives on recovery after closed-head injury. American Journal of Psychiatry, *153*, 7-10.
- Binder, L. (1997). A review of mild head trauma. Part II. Clinical implications. Journal of Clinical and Experimental Neuropsychology, *19*, 432-457.
- 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.

- Bury, A. S., & Bagby, R. M. (2002). The detection of feigned uncoached and coached posttraumatic stress disorder with the MMPI-2 in a sample of workplace accident victims. Psychological Assessment, *14*, 472-484.
- 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.
- 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., Graham, J.R., Ben-Porath, Y. S., Tellegen, A., Dahlstrom, W.G., & Kaemmer, B. (2001). MMPI-2: Manual for administration, scoring and interpretation. Revised Edition. Minneapolis: University of Minnesota Press.
- Campbell, W.H., & Tueth, M.J. (1997). Misplaced rewards: Veterans' Administration System and symptom magnification. Clinical Orthopaedics and Related Research, *336*, 42-46.
- 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*, 449-463.
- Elhai, J. D., Gold, S. N., Sellers, A. H., & Dorfman, W. I. (2001). The detection of malingered posttraumatic stress disorder with MMPI-2 Fake Bad Indices. Assessment, *8*, 221-236.
- Elhai, J.D. (2001). "Cross-validation of the MMPI-2 in detecting malingered post-traumatic stress disorder": Errata. (2001). Journal of Personality Assessment, *77*, 189.
- First, M. B. (Ed.) (2002). SCID-Frequently Asked Questions. Available at <http://cpmnet.columbia.edu/dept/scid/scidfaq.htm>.
- First, M. B., Spitzer, R. L., Gibbon, M. & Williams, J. B. W. (1995). Structured Clinical Interview for DSM-IV Axis I Disorders – Patient Edition, Version 2.0. New York: New York State Psychiatric Institute Biometric Research Department. [cited in Bury & Bagby, 2002]
- Fox D.D., Gerson A., & Lees-Haley P.R. (1995). Interrelationship of MMPI-2 validity scales in personal injury claims. Journal of Clinical Psychology, *51*, 42-47.
- Frederick, R. I., Sarfaty, S. D., Johnston, J. D., & Powel, J. (1994). Validation of a detector of response bias on a forced-choice test of nonverbal ability. Neuropsychology, *8*, 118-125.
- Frueh, B. C., Gold, P., & de Arellano, M. A. (1997). Symptom overreporting in combat veterans evaluated for PTSD: Differentiation on the basis of compensation status. Journal of Personality Assessment, *68*, 369-384.
- Frueh, B. C., Hamner, M. B., Cahill, S. P., Gold, P. B., & Hamlin, K. L. (2000). Apparent symptom overreporting in combat veterans evaluated for PTSD. Clinical Psychology Review, *20*, 853-885.
- Frueh, B.C., Smith, D.W., & Barker, S.E. (1996). Compensation seeking status and psychometric assessment of combat veterans seeking treatment for PTSD. Journal of Traumatic Stress, *9*, 427-439.
- Gervais, R. O., Allen, L.M., Green, P., & Cunningham, S. (1998, November). The effects of coaching DSM-IV pain disordered patients on the Computerized Assessment of Response Bias. Presented at the annual meeting of the National Academy of Neuropsychology, Washington, DC. Archives of Clinical Neuropsychology, *14*, 97-98.

- Gervais, R., Green, P., Allen, L. M. & Iverson, G. Effects of coaching on symptom validity testing in chronic pain patients presenting for disability assessments. *Journal of Forensic Neuropsychology*, 2001, 2 (2), 1-19.
- Gervais, R.O., Green, P., & Allen, L.M. (1999, November). Differential sensitivity to symptom exaggeration of verbal, visual, and numerical symptom validity tests. Presented at the annual meeting of the National Academy of Neuropsychology, San Antonio, TX. Archives of Clinical Neuropsychology, 14, 746-747. (Abstract).
- Gist, R. & Lubin, B. (Eds.) (1999). Response to disaster. Philadelphia, PA: Brunner/Mazel.
- Gouvier, W. D. (1999). Base rates and clinical decision making in neuropsychology. In J. J. Sweet (Ed.), Forensic neuropsychology. Fundamentals and practice (pp. 27-37). Lisse, The Netherlands: Swets & Zeitlinger.
- Gouvier, W. D., Lees-Haley, P. R., and Hammer, J. H. (2003). The neuropsychological examination in the detection of malingering in the forensic arena: Costs and benefits. In Prigatano, G. P. and Pliskin, N. H. Clinical neuropsychology and cost outcome research: A beginning. New York: Psychology Press, Inc. Pp. 405-424.
- Graham, J. R. (1990). MMPI-2: Assessing Personality and Psychopathology. New York: Oxford University Press.
- Graham, J.R. (1999). MMPI-2 Assessing personality and psychopathology (3rd ed.). New York: Oxford University Press.
- Green, P., Rohling, M. L., Lees-Haley, P. R. and Allen III, L. M. III (2001). Effort Has a Greater Effect on Test Scores than Severe Brain Injury in Compensation Claimants. Brain Injury, 15, 1045-1060.
- Greene, R.L. (2000). MMPI-2/MMPI: An interpretive manual (2nd ed.). Needham Heights, MA: Allyn & Bacon.
- Greiffenstein, M.F., Baker, W.J., Axelrod, B., Peck, T.A., & Gervais, R. (in press, 2004). The Fake Bad Scale and MMPI-2 F-Family in Detection of Implausible Psychological Trauma Claims. The Clinical Neuropsychologist.
- Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. Psychological Assessment, 6, 218-224.
- Greiffenstein, M. F., Baker, W. J., & Gola, T. (1996). What Kind of Faking Does the *Fake Bad Scale* Measure? American Psychology-Law Society Newsletter (APA Convention Issue).
- Greiffenstein, M. F., Baker, W. J., Gola, T., Donders, J. & Miller, L. J. (2002). The FBS in atypical and severe closed head injury litigants. Journal of Clinical Psychology, 58, 1591-1600.
- Greve, K. W. & Bianchini, K. J. (2004). Response to Butcher et al., The construct validity of the Lees-Haley Fake Bad Scale... Archives of Clinical Neuropsychology, 19, 337-339. Letter to the editor.
- Grillo, J., Brown R.S., Hilsabeck R., Price J.R., & Lees-Haley P.R. (1994). Raising doubts about claims of malingering: implications of relationships between MCMI-II and MMPI-2 performances. Journal of Clinical Psychology, 50, 651-655.
- Hayes, J.S., Hilsabeck, R., & Gouvier, W.D. (1999). Assessment of malingering among personal injury litigants. In N. Varney and R. Roberts (Eds.) Neurobehavioral Dysfunction Following Mild Head Injury. Mahwah, NJ: Lawrence Erlbaum Associates.

- Heaton, R. K., Smith, H. H., Jr., Lehman, R. A., & Vogt, A. T. (1978). Prospects for faking believable deficits on neuropsychological testing. Journal of Consulting and Clinical Psychology, 46, 892-900.
- Iverson, G. L., & Binder, L. M. (2000). Detecting exaggeration and malingering in neuropsychological assessment. Journal of Head Trauma Rehabilitation, 15, 829-858.
- Iverson, G. L., Henrichs, T. F., Barton, E. A., & Allen, S. (2002). Specificity of the MMPI-2 Fake Bad Scale as a marker for personal injury malingering. Psychological Reports, 90, 131-136.
- Kazdin, A. E. (1998). Methodological issues & strategies in clinical research (2nd ed.). Washington, DC: American Psychological Association.
- Kessler, R. C. & Walters, E. (2002). The national comorbidity study. In Tsuang, M. T. & Tohen, M. (Eds.) Textbook in psychiatric epidemiology. New York: Wiley-Liss, Inc.
- Kessler, R. C., Sonnega, A., Bromet, E., Hughes, M., Nelson, C. B., & Kessler, R. C. (1995). Posttraumatic stress disorder in the National Comorbidity Survey. Archives of General Psychiatry, 52, 1048-60.
- Kilpatrick, D. G., & Resnick, H. S. (1992). Posttraumatic Stress Disorder associated with exposure to criminal victimization in clinical and community populations. In Davidson, J. R. T. & Foa, E. B. (Eds.) Posttraumatic stress disorder: DSM-IV and beyond. Washington, DC: American Psychiatric Press, Inc.
- Lanyon, R. I. (2001). Dimensions of self-serving misrepresentation in forensic assessment. Journal of Personality Assessment, 76, 169-179.
- Larrabee, G. J. (1997). Neuropsychological outcome, post concussion symptoms, and forensic considerations in mild closed head trauma. Seminars in Clinical Neuropsychiatry, 2, 196-206.
- Larrabee, G. J. (1996). Somatic malingering on the MMPI/MMPI-2 in litigating subjects. Presented to the National Academy of Neuropsychology annual conference, New Orleans, Louisiana, November.
- Larrabee, G. J. (1998). Somatic malingering on the MMPI and MMPI-2 in personal injury litigants. The Clinical Neuropsychologist, 12, 179-188.
- Larrabee, G. J. (2000). Exaggerated MMPI-2 symptom report in probable malingerers. Paper presented at the 28th Annual Meeting of the International Neuropsychological Society, February 11, Denver, Colorado.
- Larrabee, G. J. (2000, October). Assessing Malingering. Presented at The American Academy of Clinical Neuropsychology, Chicago, IL.
- Larrabee, G. J. (2003a). Detection of symptom exaggeration with the MMPI-2 in litigants with malingered neurocognitive dysfunction. The Clinical Neuropsychologist, 17, 54-68.
- Larrabee, G. J. (2003b). Exaggerated MMPI-2 Symptom Report in Personal Injury Litigants with Malingered Neurocognitive Deficit. Archives of Clinical Neuropsychology, 18, 673-686.
- Larrabee, G. J. (2003c). Exaggerated pain report in litigants with malingered neurocognitive dysfunction. The Clinical Neuropsychologist, 17, 395-401.
- Larrabee, G. J. (2003d). Detection of malingering using atypical performance patterns on standard neuropsychological tests. The Clinical Neuropsychologist, 17, 410-425.

- 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*, 681-689.
- Lees-Haley P. R. (1997). MMPI-2 base rates for 492 personal injury plaintiffs: Implications and challenges for forensic assessment. Journal of Clinical Psychology, *53*, 745-755.
- Lees-Haley, P. R. & Fox, D. D. (2004). Commentary on Butcher, Arbisi, Atlis, and McNulty (2003) on the Fake Bad Scale. Archives of Clinical Neuropsychology, *19*, 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. **The answer key to the FBS appears in an appendix to this paper (also see below).**
- Lees-Haley, P.R. & Dunn, J.T. (1994). The ability of naive subjects to report symptoms of mild brain injury, PTSD, major depression and generalized anxiety disorder. Journal of Clinical Psychology, *50*(2), 252-256.
- Martens, M., Donders, J., & Millis, S.R. (2001). Evaluation of invalid response sets after traumatic head injury. Journal of Forensic Neuropsychology, *2*, 1-18.
- Meyers, J. E., Millis, S. R., & Volkert, K. (2002). A validity index for the MMPI-2. Archives of Clinical Neuropsychology, *17*, 157-169.
- Miller, L. J. & Donders, J. (2001). Subjective symptomatology after traumatic head injury. Brain Injury, *15*, 297-304.
- Millis, S. R., Putnam, S. H., & Adams, K.M. (1995). Neuropsychological malingering and the MMPI-2: Old and new indicators. Paper presented at the 30th Annual Symposium on Recent Developments in the Use of the MMPI, MMPI-2 and MMPI-A, St Petersburg Beach, FL, March 19.
- Mittenberg, W., Patton, C., Canyock, E. M., & Condit, D. C. (2002). Base rates of malingering and symptom exaggeration. Journal of Clinical and Experimental Neuropsychology, *24*, 1094-1102.
- Posthuma, A. B. & Harper, J. F. (1998). Comparison of MMPI-2 responses of child custody and personal injury litigants. Professional Psychology: Research and Practice, *29*, 437-443.
- Putnam, S. H., Millis, S. R., & Adams, K. M. (1998). Consideration of impression management in the neuropsychological examination with the MMPI-2. Paper presented at the Annual Meeting of the American Psychological Association, San Francisco, CA, August.
- Rogers, R. (1997). Introduction. In R. Rogers (ed.), Clinical assessment of malingering and deception. New York: Guilford, 1-19.
- Rogers R., Sewell K. W., & Salekin R. T. (1994). A meta-analysis of malingering on the MMPI-2. Assessment, *1*, 227-237.
- Rogers R., Sewell K. W., & Salekin R. T. (1995). "A meta-analysis of malingering on the MMPI-2": Erratum. Assessment, *2*, 201.
- Rogers R., Sewell K.W., & Ustad L.L. (1995). Feigning among chronic outpatients on the MMPI-2: A systematic examination of fake-bad indicators. Assessment, *2*, 81-89.
- Rohling, M. L., Binder, L. M. & Langhinrichsen-Rohling, J. (1995). Money matters: A meta-analytic review of the association between financial compensation and the experience and treatment of chronic pain. Health Psychology, *14*, 537-547.

- Ross, S. R., Millis, S. R., Krukowski, R. A., Putnam, S. H., and Adams, K. M. (2004). Detecting probable malingering on the MMPI-2: An examination of the Fake bad Scale in mild head injury. Journal of Clinical and Experimental Neuropsychology, *26*, 115-124.
- Rubonis, A.V., & Bickman, L. (1991). Psychological impairment in the wake of disaster: The disaster-psychopathology relationship. Psychological Bulletin, *109*, 384-399.
- Slick, D. J., Hopp, G., Strauss, E. & Spellacy, F. J. (1996). Victoria Symptom Validity Test: Efficiency for detecting feigned memory impairment and relationship to neuropsychological tests and MMPI-2 validity scales. Journal of Clinical and Experimental Neuropsychology, *18*, 911-922.
- Trueblood, W., & Schmidt, M. (1993). Malingering and other validity considerations in the neuropsychological evaluation of mild head injury. Journal of Clinical and Experimental Neuropsychology, *15*, 578-590.
- Tsushima, W.T. & Tsushima, V.G. (2001). Comparison of the Fake Bad Scale and other MMPI-2 validity scales with personal injury litigants. Assessment, *8*(2), 205-212.
- Van Gaasbeek, J. K., Denney, R. L., & Harmon, J. (2001). Another look at an MMPI-2 neurocorrective factor in forensic cases: utility of the Fake Bad Scale. Archives of Clinical Neuropsychology, *16*, 813.
- Wilkinson, L. and Task Force on Statistical Inference, APA Board of Scientific Affairs. (1999). Statistical Methods in Psychology Journals: Guidelines and Explanations American Psychologist, *54*, 594-604.
- Youngjohn, J. R., Davis, D., & Wolf, I. (1997). Head injury and the MMPI-2: Paradoxical severity effects and the influence of litigation. Psychological Assessment, *9*, 177-184.

Answer Key:

Following are the MMPI-2 items and scored direction of answering for the Fake Bad Scale (FBS):

Add one point if marked True: 11, 18, 28, 30, 31, 39, 40, 44, 59, 111, 252, 274, 325, 339, 464, 469, 505, 506

Add one point if marked False: 12, 41, 57, 58, 81, 110, 117, 152, 164, 176, 224, 227, 248, 249, 250, 255, 264, 284, 362, 373, 374, 419, 433, 496, 561