Commentary on Butcher, Arbisi, Atlis, and McNulty (2003) on the Fake Bad Scale

Paul R. Lees-Haley a,*, David D. Fox b

a 3021 Panorama Drive, Huntsville, AL 35801, USA
b Consultants in Psychological Assessment, Glendale, CA, USA

Accepted 17 December 2003

Butcher, Arbisi, Atlis, and McNulty (2003) misunderstand the Fake Bad Scale (FBS) and make naïve, inaccurate assumptions about malingering. Their reasoning and methodology do not support their conclusions. Their methodology and reporting are very unusual for a research study. Not even basic descriptive statistics were reported, for example, means, standard deviations, or ranges. Five of their six samples were not in the setting for which the FBS was designed, and only one—their smallest sample—was personal injury litigants. Applying the FBS in other contexts is analogous to applying the Marital Distress Scale to single persons. The authors used data in a National Computer Systems archival database containing 119,672 cases but 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 when evaluating a scale designed to detect invalid profiles is analogous to testing the MMPI-2’s capacity to detect depression after first removing depressed patients from the sample. The selection bias in this methodology is calculated to produce a misleadingly low impression of the scale being evaluated.

Their cutoff criteria were not justified and they are so unusually high that many people normally thought to be exaggerating were presumed to be valid cases instead of being 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. Butcher et al. used dramatically different standards to criticize the FBS than Butcher and his colleagues use in the MMPI-2 manuals (Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989; Butcher et al., 2001). 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

*Corresponding author. Tel.: +1-256-551-1024; fax: +1-256-551-1036.
E-mail address: paul@lees-haley.com (P.R. Lees-Haley).
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.

Butcher et al. criticize the FBS for not measuring the same behavior as their preferred scales, \( F \), \( F(b) \), and \( F(p) \). The FBS was never intended to replace the \( F \) family and was not created in the same manner. The FBS is based on empirical observations of plaintiffs rather than the Butcher et al. assumption that someone who gives unusual (infrequent) answers is malingering. One of the reasons the FBS has value is that it has been shown to be more sensitive to somatic exaggeration than the \( F \) family (Larrabee, 1998). Judging the FBS by its similarity to \( F \) and \( F(b) \) undermines the development of new scales that remedy current MMPI-2 deficiencies.

The authors conclude that the FBS over-predicts malingering in clinical and forensic samples, but offer no data to support this claim. 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’s actual usefulness. 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. In contrast, others have used 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, 1998, 2003a, 2003b; Martens, Donders, & Millis, 2001; Meyers, Millis, & Volkert, 2002; Miller & Donders, 2001; Millis, Putnam, & Adams, 1995; Posthuma & Harper, 1998; Putnam, Millis, & Adams, 1998; Ross et al., 2004; Millis, Krukowski, Putnam, & Adams, 1995; Tsushima & Tsushima, 2001).

The authors’ 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 literature on malingering. With no empirical basis at all, Butcher et al. arbitrarily presume malingering to be less common than voluminous empirical research literature has indicated. Estimates of the incidence of malingering have varied from less than 5 to over 60% in personal injury neuropsychology (Hayes, Hilsabeck, & Gouvier, 1999), but a general trend is estimates at about 30% or more as a reasonable base rate for estimated malingering in the FBS setting, that is, compensation seeking contexts (Binder, 1997; Greiffenstein, Baker, & Gola, 1994; Gouvier, Lees-Haley, & Hammer, 2003; Green, Rohling, Lees-Haley, & Allen, 2001; Heaton, Smith, Lehman, & Vogt, 1978; Lees-Haley, 1997). Mittenberg, Patton, Canyock, and 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.’s study, in which malingering was simply presumed absent.

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 transparent. It is absurd to argue that malingering cannot occur unless the plaintiff is mentally healthy.
Butcher et al. address whether the FBS represents a single homogeneous dimension. The highest alpha coefficient in their samples (.85) was in the Personal Injury group. This is the group expected to have the highest rate of symptom exaggeration and 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. 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 concern somatic symptoms and a larger number, 15, are denial of deviant attitudes or beliefs.

In a fit of righteous indignation, the authors claim that persons with PTSD will suffer further if they are falsely labeled as malingerers, but express no concern for the victims of malingerers who are endorsed by psychologists as having valid injuries. They also complain that women score higher than men. It should be evident to the reader that this portion of the article is ad hominem 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.

Does the Butcher et al. study have any value? Yes—it provides support for the FBS by showing that, as expected, those with the highest risk of symptom over-reporting score higher on the FBS than other groups.

References


