



Commentary on “The Detection of Feigned Uncoached and Coached Posttraumatic Stress Disorder with the MMPI-2 in a Sample of Workplace Accident Victims”

Paul R. Lees-Haley⁺

Health Education Services, Alabama, United States of America

Keywords: Feigning; Malingering; PTSD; MMPI-2; Compensation

INTRODUCTION

In the course of performing forensic psychological evaluations, consideration of possible malingering has become routine practice. Thus studies of tools for detecting or ruling out malingering are of particular interest to forensic psychologists. The above-named study by Bury and Bagby (2002) examined data that potentially could determine whether a number of MMPI-2 validity scales could distinguish between 22-year-old college students (age $M = 22.3$, $SD = 5.2$) faking Posttraumatic Stress Disorder (PTSD) and 40-year-old workers' compensation claimants (age $M = 40.0$, $SD = 9.34$) 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).

The comments below on this article are organized as follows:

1. Unfounded presumption of validity and reliability
2. Does this sample represent a larger population well enough to permit generalizations?
3. Confounding of the data by malingering/exaggeration
4. Conclusion

Unfounded Presumption of Validity and Reliability

In lieu of measuring interrater reliability at diagnosing PTSD, the diagnostic 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 patient form (SCID-I/P; First, Spitzer, Gibbon & Williams, 1995). One psychologist's clinical opinion 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 generally accepted methods for measurement of interrater reliability. Specifically in connection with the diagnostic procedure used in this paper, note that according to the SCID senior author Michael 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

⁺Address for correspondence: Paul, R., Lees-Haley, PhD: 3021 Panorama Drive Huntsville, Alabama 35801, USA; Ph: 256-551-1024; Facsimile: 256-551-1036; Email: paul@lees-haley.com

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 subject of this study would be “self-reported claims of PTSD from compensation-seeking claimants.”

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 that the PTSD symptoms were bona fide 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 settings that it is explicitly named in the differential diagnostic process in DSM-IV (APA, 1994) and DSM-IV-TR (APA, 2000).

It is problematic to presume the validity of SCID data in forensic settings, as it is neither difficult nor unlikely for claimants to endorse PTSD symptoms in the absence of actually suffering these symptoms (Lees-Haley, Price, Williams, & Betz, 2001). 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 is unreasonable of us as experts to presume that workers’ compensation

claimants are naïve, as they are likely exposed to symptom-related information during the course of pursuing a claim through the workers’ compensation system. Bury and Bagby were well 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 claimants’ MMPI-2 scores cast doubt on the authenticity of their symptoms. Table 1 of the study shows the majority of the 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 claimants scored in a range on the F-K of greater than zero, which Graham suggests as an indication for considering malingering (1999). The fact that Bury and Bagby did not use the SCID version with psychotic screening or mention any psychosis implies that psychosis is not the explanation for their elevated F family scores. Moreover, 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 those classified as 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 conclusions in this study 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 Bury and Bagby 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 the majority of the claimants studied experienced stressors that were too vaguely defined to meet Criterion A, thus undermining conclusions based on symptoms and test data predicated on the assumption that PTSD was present.

Does this Sample Represent a Larger Population well enough to Permit Generalizations?

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 (and thus, other things being equal, would be expected to have had more traumatic experiences). For example, Kessler, Sonnega, Bromet, Hughes, and Nelson (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). Despite the high rate of exposure of students to Criterion A events, PTSD was conspicuously absent in the students yet presumed present in the claimants.

Confounding of the data by Malingering/Exaggeration

Bury and Bagby accurately acknowledge that the sample they characterized 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). Given these observations indicating invalidity in the PTSD claimants, we need to re-think the conceptualization of the claimants as “bona fide”. 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 a group of claimants who can plausibly be presumed to have bona fide PTSD. The 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 non-litigating 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. Greiffenstein et al. studied a more specialized form of trauma – head injury – rather than the broad range of traumas seen in PTSD research.

CONCLUSION

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 acknowledge), the vagueness of definition of Criterion A experiences, the evidence of false positives, and the 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 and inconclusive. 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 cannot be relied upon as a basis for evaluating any of the validity scales.

ACKNOWLEDGEMENTS

Thank you to Dr. Manfred Greiffenstein for his thoughtful suggestions.

REFERENCES

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.

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.

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.

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.

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.

First, M. B. (Ed.) (2002). SCID-Frequently Asked Questions. Available at <http://cpmcnet.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.

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.
- Graham, J.R. (1999). *MMPI-2 Assessing personality and psychopathology (3rd ed.)*. New York: Oxford University Press.
- Greene, R.L. (1999). *MMPI-2/MMPI: An interpretive manual (2nd ed.)*. Needham Heights, MA: Allyn & Bacon.
- 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.
- Greiffenstein, M. F., Baker, W. J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment, 6*, 218-224.
- Kazdin, A. E. (1998). *Methodological issues & strategies in clinical research (2nd ed.)*. Washington, DC: American Psychological Association.
- 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*, 1048-60.
- 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.
- 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*, 252-256.
- Lees-Haley, P., Price, J. R., Williams, C. W., and Betz, B. P. (2001). Use of the impact of events scale in the assessment of emotional distress and PTSD may produce misleading results. *Journal of Forensic Neuropsychology, 2*, 45-52.
- 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.
- 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.