Utility of the Trauma Symptom Inventory's Atypical Response Scale in Detecting Malingered Post-Traumatic Stress Disorder

Jon D. Elhai
University of South Dakota

Matthew J. Gray
University of Wyoming

James A. Naifeh
Jimmie J. Butcher
University of South Dakota

Joanne L. Davis
University of Tulsa

Sherry A. Falsetti
University of Illinois at Chicago

Connie L. Best
Medical University of South Carolina

The authors examined the Trauma Symptom Inventory's (TSI) ability to discriminate 88 student post-traumatic stress disorder (PTSD) simulators screened for genuine PTSD from 48 clinical PTSD-diagnosed outpatients. Results demonstrated between-group differences on several TSI clinical scales and the Atypical Response (ATR) validity scale. Discriminant function analysis using ATR revealed 75% correct patient classification but only 48% correct simulator classification, with an overall correct classification rate of 59% (positive predictive power [PPP] = .71; negative predictive power [NPP] = .51). Individual ATR cutoff scores did not yield impressive classification results, with the optimal cutoff (T score = 61) correctly classifying only 61% of simulators and patients (PPP = .66, NPP = .54). Although ATR was not developed as a malingered PTSD screen, instead serving as a general validity screen, caution is recommended in its current clinical use for detecting malingered PTSD.

Keywords: malingering; symptom overreporting; posttraumatic stress disorder; trauma symptom inventory

U.S. national surveys indicate that post-traumatic stress disorder (PTSD) affects at least 7% of the general population (Kessler, Sonnega, Bromet, Hughes, & Nelson, 1995; H. S. Resnick, Kilpatrick, Dansky, Saunders, & Best, 1993). However, PTSD may be frequently fabricated in some clinical and forensic settings, usually to reduce

Assessment, Volume 12, No. 2, June 2005 210-219
DOI: 10.1177/1073191105275456

criminal charges or for financial gain (e.g., disability or personal injury payments) (P. J. Resnick, 1997). Furthermore, distinguishing genuine from fabricated PTSD is quite complicated (Frueh, Hamner, Cahill, Gold, & Hamlin, 2000) because PTSD is relatively easy to feign and its symptoms are largely based on subjective report.

In recent years, several studies have investigated malingered PTSD by comparing individuals instructed to fabricate PTSD with genuine PTSD patients on psychological test scores. The majority of these studies used the Minnesota Multiphasic Personality Inventory–2 (MMPI-2) (Butcher et al., 2001), including genuine PTSD-diagnosed comparison groups composed of survivors of combat exposure (Elhai et al., 2002; Elhai, Gold, Frueh, & Gold, 2000), sexual abuse (Elhai et al., 2004; Elhai, Gold, Sellers, & Dorfman, 2001), workplace accidents (Bury & Bagby, 2002), and general traumatic events from adulthood (Wetter, Baer, Berry, Robison, & Sumpter, 1993). Although these studies have demonstrated that the MMPI-2's "fake bad" scales yield respectable classification rates in distinguishing genuine from simulated PTSD, effect sizes are significantly smaller than those seen in malingering studies of other types of psychopathology (e.g., depression and schizophrenia) (Rogers, Sewell, Martin, & Vitacco, 2003). One possible explanation for these smaller effect sizes is that the MMPI-2 fake bad scales have demonstrated confounding with trauma-related psychopathology and distress (Flitter, Elhai, & Gold, 2003) and therefore may not be as effective as validity scales from trauma-specific instruments in detecting fabricated PTSD.

Of the many screening instruments that assess the long-term effects of trauma victimization, including PTSD (Briere, 2004), the 100-item multiscale Trauma Symptom Inventory (TSI; Briere, 1995) is unique in its inclusion of validity scales. It contains the Inconsistent Response (measuring random responding), Response Level (measuring symptom underreporting), and Atypical Response (measuring symptom overreporting) scales. The TSI has 10 clinical scales, possessing strong psychometric properties. Reliability is adequate, with coefficient alpha values ranging from .84 to .87 across studies (Briere, Elliott, Harris, & Cotman, 1995). In addition, criterion validity is adequate, with victims of interpersonal trauma scoring higher than control subjects on all TSI scales. Convergent validity has been found with the Clinician-Administered PTSD Scale, PTSD Checklist, Civilian Mississippi PTSD Scale, and Personality Assessment Inventory's Traumatic Stress Scale (McDevitt-Murphy, Weathers, & Adkins, in press). A statistically derived discriminant function (based on TSI scales) demonstrated 92% sensitivity to diagnosing PTSD and 91% specificity (Briere, 1995), and regression analyses demonstrated 86% overall PTSD diagnostic classification (McDevitt-Murphy et al., in press).

At present, only one published study (Edens, Otto, & Dwyer, 1998) has examined the ability of the TSI to detect fabricated PTSD. The authors compared students simulating PTSD with students completing the test honestly. They found that a score of 61T on the Atypical Response scale (ATR) was optimal in correctly identifying PTSD simulators (82% sensitivity) and honest participants (92% specificity) (with cross-validation producing a slightly lower sensitivity rate of 74% but equal specificity). The cutoff score of 61T produced various specificity rates when the authors applied it to descriptive data from several clinical samples reported in the literature. However, the study's lack of a direct clinical PTSD patient comparison sample limits the generalizability and significance of its findings.

The aim of this study is to examine the ability of the TSI validity scale (ATR) to detect fabricated from genuine PTSD, comparing a clinical PTSD patient sample with individuals instructed to feign PTSD. This study is important, based on PTSD's general prevalence, the large incentives available to PTSD malingerers, and significant problems distinguishing genuine from feigned PTSD. We attempted to use rigorous methodology (Rogers, 1997), including PTSD training and testing materials, monetary incentives to enhance motivation, cautionary statements about believability, situational malingering contexts (i.e., providing a scenario in which to malinger, such as avoiding prosecution for a crime), and posttest queries of training comprehension. We also attempted to enhance internal validity by ensuring that genuine PTSD was not present in the PTSD simulators (prior to simulation) but was present in the PTSD patients, using standardized assessments of both trauma exposure and PTSD symptom Criteria B through D, recommended in trauma research (Briere, 2004). Our research question was as follows: How effective is the TSI in distinguishing PTSD simulators from patients? We hypothesized that the TSI would be fairly effective, although less effective than Edens et al. (1998)

Training materials used for the post-traumatic stress disorder simulation group are available by request from the first author. Portions of this article were presented at the 20th annual meeting of the International Society for Traumatic Stress Studies, New Orleans, Louisiana (2004, November). This study was funded by a Research Enhancement Award to the first author, from the Office of Research and Graduate Education, the University of South Dakota; and a National Institute of Mental Health–funded postdoctoral fellowship to the fifth author. Requests for reprints may be addressed to John D. Elhai (preferably via e-mail) at Disaster Mental Health Institute, University of South Dakota, 414 East Clark Street–SDU 114, Vermillion, SD 57069-2390; e-mail: jonelhai@hotmail.com or jelhai@usd.edu.

discovered because the present investigation used tighter methodological constraints, including a PTSD patient comparison group.

METHOD

Participants

PTSD simulation group. A sample of 88 participants (41 men, 47 women) enrolled in introductory psychology courses served as subjects for the PTSD simulation condition. These students were at least age 18 and attending college at one of two medium-sized state universities in the midwestern and western United States. Participants were recruited in groups from their psychology classes (for extra credit) or from their departmental research pool (for research credit).

For demographic characteristics, age ranged from 18 to 39 years ($M = 20.05$, $SD = 2.84$). Educational level ranged from 12 to 16 years, with an average of 12.75 ($SD = 1.07$). Parental annual household income ranged from $12,000 to $180,000 ($M$ = $66,405.17, SD = $35,481.33); personal income ranged from $0.00 to $40,000 ($M$ = $5,743.95, SD = $7,073.13). The majority was working part-time (50.00%) or not at all (47.73%). Most participants were Caucasian (93.18%). Relationship status for the majority was single (88.64%), whereas 7.95% were cohabitating with a significant other.

PTSD patient group. The clinical comparison group was a PTSD-diagnosed sample, drawn from archival data of 48 patients (3 men, 45 women).[1] These patients were age 18 and older who presented to an outpatient treatment program for victims of violent crime at a medical university in the southeastern United States. The group's trauma history (taken from interviews, described below) primarily involved sexual or physical victimization, including 56.30% reporting child sexual abuse victimization, 43.80% reporting child physical abuse, 45.80% reporting domestic violence, and 52.10% reporting adult sexual assault (total percentage exceeds 100%, due to multiple traumatic events reported).

Patients were assessed by supervised clinical psychology interns and postdoctoral fellows, who administered a battery of psychological assessments. Of relevance to the present study, trauma history was assessed with the Trauma Assessment for Adults—Interview Version (TAA) (H. S. Resnick, 1996; H. S. Resnick, Falsetti, Kilpatrick, & Freedy, 1996). The TAA is a behaviorally specific, structured interview of exposure to 13 potentially traumatic events, with empathic and contextually orienting preface statements, yielding similar trauma prevalence rates to other major epidemiological trauma studies (H. S. Resnick et al., 1996). Psychiatric diagnoses were made according to the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-IV*; American Psychiatric Association, 1994). All diagnoses were taken during initial intake examination, obtained from the Structured Clinical Interview for *DSM-IV* Axis I Disorders (SCID-I), a structured, diagnostic interview with established psychometric properties (First, Spitzer, Gibbon, & Williams, 1996). However, only the SCID-I's PTSD, Panic Disorder, and Major Depressive Disorder modules were administered.

Only patients diagnosed with PTSD were included in this study. In addition, panic disorder was diagnosed in 39.60% and major depressive disorder in 68.80% of the sample, reflecting the established high psychiatric comorbidity found in PTSD patients (Kessler et al., 1995). For demographic characteristics, age ranged from 18 to 53 years, with a mean of 36.19 ($SD = 9.33$). Educational level ranged from 8 to 22 years, with an average of 14.13 years ($SD = 2.50$). The majority was Caucasian (79.17%) or African American (14.58%). Relationship status was primarily single (40.43%), previously married (27.66%), currently married (19.15%), or cohabitating (12.77%).

Procedure

The current study was Institutional Review Board (IRB) approved, with participants treated according to the ethical standards for human subjects. For the PTSD patient group, the TSI was administered with standard instructions during intake evaluations. For the PTSD simulation group, procedures are described below.

For PTSD simulators, an experimenter first explained the purpose of the study (i.e., to detect fabricated PTSD). These participants, in groups typically ranging from 30 to 60 students, were given the following self-report forms printed with precoded subject numbers (unless otherwise noted): (a) informed consent form (not coded by subject number); (b) contact information form; (c) demographic questionnaire; (d) Life Events Checklist, a PTSD Criterion A1 measure of direct and indirect exposure to 16 potentially traumatic events (Blake et al., 1990), with excellent test-retest reliability and convergent validity (Gray, Litz, Wang, & Lombardo, 2004) (and modified to include empathic, contextually-orienting preface statements); (e) brief survey querying about several trauma-related characteristics (e.g., age at onset) for endorsed traumatic events, adapted from the TAA (H. S. Resnick, 1996) and including a PTSD Criterion A2 assessment of fear, helplessness, and horror; and (f) PTSD Checklist (Weathers, Litz, Herman, Huska, & Keane, 1993), a 17-item measure of PTSD symptoms, with excellent internal consistency,

convergent validity, and concurrent validity in diagnosing PTSD (Ruggiero, Del Ben, Scotti, & Rabalais, 2003).

After completing these forms, students were given training materials, composed of information on PTSD and a worksheet to assist in mastering information for a subsequent quiz. Participants were informed about monetary incentives for their simulation performances (described in detail below).

The training materials included clear descriptions of the 17 symptoms of PTSD with several case studies, available to the lay person and adapted from the *DSM-IV* and various popular abnormal psychology textbooks. To increase the specificity of participants' feigning performances, they were instructed to only choose one case study to use in simulation. Twenty-five minutes were allotted for participants to study the information and to complete the PTSD worksheet. An experimenter subsequently removed the materials and administered a 10-question true-false quiz, based on content from the materials. The experimenter then collected and scored all quizzes, telling all participants that they passed (to avoid potential embarrassment for those who failed) and could complete the TSI while simulating PTSD. However, it was determined that data from participants who scored less than 70% on the quiz would not be used in data analysis.

At this point, participants were given TSI test booklets and answer sheets with their precoded research numbers, given written instructions to feign PTSD that were also recited aloud by an experimenter, primarily taken from Elhai et al. (2001) (see the appendix).

To further increase feigning specificity, participants were asked to select which malingering scenario (obtaining money or avoiding prosecution for PTSD) they intended to use when feigning PTSD. After completing the TSI, test booklets and answer sheets were collected and a 6-point Likert-type scale (1 = *not at all*, 6 = *very*) was administered, measuring participants' perceptions of training material clarity, understanding of PTSD, confidence in portrayal of PTSD's symptom criteria (B through D), motivation based on the monetary incentives, and previous familiarity with PTSD. In addition, they were again asked to identify which malingering scenario and case study they used when simulating PTSD and to provide details on any malingering strategies they implemented.

Prizewinners were determined from participants whose total summed score on the Intrusive Experiences, Defensive Avoidance, and Anxious Arousal scales (corresponding to the three PTSD symptom criteria) most closely resembled that of the PTSD patient group's mean score, while not elevating the ATR validity scale above a *T* score of 90 (Briere, 1995). Prizes were awarded by mail to the three participants per group of 30 with the closest scores to the patient group.

Data Analysis Exclusion Criteria

TSI data from participants in either the PTSD simulation or PTSD patient groups were excluded from analyses if there were 10 or more items left blank ($n = 0$ from each group) or if random responding was evident from Inconsistency scale (INC) *T* scores greater than or equal to 75 (simulators $n = 7$, patients $n = 1$). PTSD simulators were additionally excluded if they scored less than 70% on the PTSD quiz ($n = 1$).

Last, simulators were excluded for having "possible PTSD" if they met all of the following criteria: (a) endorsement of at least one experienced or witnessed traumatic event from the Life Events Checklist (PTSD Criterion A1); (b) the report of intense fear, helplessness, or horror during at least one traumatic event from the trauma characteristics survey (Criterion A2); and (c) a PTSD Checklist score of 44 or higher (Ruggiero et al., 2003) (Criteria B through D) ($n = 17$). (We strategically used a liberal definition of "possible PTSD" to ensure that no true PTSD cases were included in the simulation group. Based on epidemiological PTSD studies mentioned above, it is unlikely that as many as 17 of 88—19%—college students would be diagnosed with PTSD.) Overall, these criteria resulted in the exclusion of 25 participants in the PTSD simulation group and 1 participant in the PTSD patient group. The remaining 63 PTSD simulators and 47 PTSD patients served as the overall sample for the study.

Among student participants, we assessed whether the resulting sample represented the larger sample by comparing excluded and nonexcluded participants on all demographic variables. Using *t* tests (for continuous variables) and Pearson chi-square tests (categorical variables), groups were only significantly different on employment status, $\chi^2(3) = 8.29$, $p < .05$, with excluded participants more likely to be unemployed.

RESULTS

First, compliance with instructions can be documented among simulators based on several posttest questions asked, with possible answers ranging from 1 (*not at all*) to 6 (*very*). In general, simulators reported that the training materials were clear ($M = 4.86$, $SD = .91$), acquiring a solid understanding of PTSD ($M = 4.62$, $SD = 1.03$). Overall confidence in PTSD simulation was slightly lower ($M = 3.94$, $SD = 1.27$). Specifically, confidence was highest in portraying PTSD's Criteria B ($M = 4.19$, $SD = 1.18$) and C symptoms ($M = 4.14$, $SD = 1.26$) and lowest for Criterion D ($M = 3.62$, $SD = 1.22$). Paired-samples *t* tests indicated significant differences between confidence for Criteria B and D, $t(62) = 4.01$, $p < .001$, and between Criteria C and

D, $t(62) = 4.47$, $p < .001$, but not between Criteria B and C, $t(62) = .44$, $p > .05$. Simulators reported only a mild amount of PTSD knowledge prior to the study ($M = 2.73$, $SD = 1.18$) and that the monetary incentive was a relatively strong motivator of simulation performances ($M = 4.05$, $SD = 1.45$).

Next, we tested for group differences in TSI PTSD simulation (clinical scales and ATR) based on university setting and gender. No significant differences were found for university or gender (all *p*s > .05). In terms of case studies chosen by simulators, 61% reported choosing the motor vehicle accident, 27% chose the gang physical attack, and 11% chose the sexual assault case study (11%). Surprisingly, the case study chosen did not significantly affect any TSI scale scores (all *p*s > .05), and none reached the threshold for a "medium" effect size (where for Cohen's *d*, small = .2, medium = .5, large = .8) (Cohen, 1990). In addition, we tested for differences in TSI simulation based on possible PTSD status (prior to excluding possible PTSD participants), using *t* tests. Only the Anxious Arousal, Anger-Irritability, and Sexual Concerns scales were significant ($p < .05$), with those judged to have possible PTSD scoring higher (only the latter scale reached the medium effect size threshold).

A one-way multivariate analysis of variance (MANOVA) (two-tailed) was conducted to assess whether PTSD simulators and patients scored differently on the TSI's ATR scale and all clinical scales. MANOVA results demonstrated a significant multivariate effect of group membership, $F(11, 98) = 4.47$, $p < .001$ (Wilks's lambda = .67). Table 1 displays univariate ANOVA test results demonstrating that simulators scored significantly higher on ATR, Dysfunctional Sexual Behavior (DSB), and Tension Reduction Behavior (TRB) scales, whereas patients scored higher on the Anxious Arousal (AA) and Sexual Concerns (SC) scales ($p < .05$). Effect sizes were small to medium for the significant scales (see Table 1).

Next, ATR was entered as a single predictor variable in a discriminant function analysis (DFA), with group status as the criterion variable (1 = PTSD patients, 2 = PTSD simulators). Overall, the model yielded a canonical correlation of .23 ($p = .02$), indicating that it accounted for 5% of the variance in discriminating PTSD simulators from patients. In terms of group classification, the DFA correctly classified only 47.6% of simulators (sensitivity) and 74.5% of patients (specificity) for an overall correct classification rate of 59.1%. Using a Jackknifed Mahalanobis D^2 procedure (which conducts as many DFAs as there are subjects, excluding a different subject each time to simulate cross-validation with new samples), correct classification of simulators and patients remained exactly the same.

TABLE 1
Trauma Symptom Inventory Mean *T* Scores, Standard Deviations, and ANOVA Results for Post-Traumatic Stress Disorder (PTSD) Simulators and PTSD Patients

	PTSD Simulators (n = 63)		*PTSD Patients (n = 47)*		*ANOVA*	
Scale	M	SD	M	SD	F*(1, 108)*	*Cohen's* d
Atypical Response	69.23	18.47	61.60	13.01	5.85*	.48
Anxious Arousal	63.75	8.83	67.70	10.85	4.42*	.40
Depression	61.66	8.48	62.74	9.79	.39	.12
Anger/ Irritability	58.60	10.24	57.43	11.34	.32	.11
Intrusive Experiences	69.66	9.61	69.98	10.47	.03	.03
Defensive Avoidance	63.92	6.90	66.51	6.80	3.84	.38
Dissociation	65.00	10.50	64.79	10.87	.01	.02
Sexual Concerns	55.26	9.60	59.77	13.32	4.26*	.39
Dysfunctional Sexual Behavior	59.66	14.81	53.74	12.61	4.86*	.43
Impaired Self-Reference	60.21	9.54	61.10	10.04	.29	.09
Tension Reduction Behavior	65.16	12.67	57.77	12.16	9.49**	.60

NOTE: For Cohen's *d*, .2 = small, .5 = medium, and .8 = large. *T* scores were computed, using separate norms for men and women, and for separate age groups (age 18 to 54, versus age 55 and older), as per the Trauma Symptom Inventory's manual.
*$p < .05$. **$p < .01$.

We calculated diagnostic efficiency statistics for our DFA (see Table 2). We also included adjustments for three different malingering base rates estimated in the literature: clinical (7%) and forensic settings (16% and 21%) (Rogers, Salekin, Sewell, Goldstein, & Leonard, 1998; Rogers, Sewell, & Goldstein, 1994). Results demonstrated that as the malingering base rate was lowered (e.g., for clinical settings), ATR's positive predictive power drastically reduced to unacceptable levels (i.e., less than .70 and thus approaching chance predictions).

We also computed an additional DFA, dummy-coding ATR based on its recommended *T* score cutoff of 90 (Briere, 1995) (1 = ATR ≤ 90, 2 = ATR > 90), in predicting group status (1 = PTSD patients, 2 = PTSD simulators). The model obtained a canonical correlation of .25 ($p = .01$), accounting for 6% of the group status variance. Sensitivity was unacceptable, with only 22.2% of simulators correctly classified, but specificity revealed a respectable

TABLE 2
Classification Accuracy for the Atypical Response Scale in Detecting Post-Traumatic Stress Disorder (PTSD) Simulators and PTSD Patients

Original DFA: 50% malingering base rate	
Sensitivity	.48
Specificity	.75
PPP	.71
NPP	.51
OCC	.59
Original DFA: 21% malingering base rate	
Sensitivity	.48
Specificity	.75
PPP	.33
NPP	.84
OCC	.69
Original DFA: 16% malingering base rate	
Sensitivity	.48
Specificity	.75
PPP	.28
NPP	.88
OCC	.71
Original DFA: 7% malingering base rate	
Sensitivity	.48
Specificity	.75
PPP	.13
NPP	.95
OCC	.73
Second DFA (predictor variable: 1 = ATR ≤ 90, 2 = ATR > 90): 50% malingering base rate	
Sensitivity	.22
Specificity	.96
PPP	.88
NPP	.48
OCC	.54
Cutoff *T* score of 58 (observed base rate)	
Sensitivity	.70
Specificity	.43
PPP	.60
NPP	.53
OCC	.55
Cutoff *T* score of 61 (observed base rate)	
Sensitivity	.65
Specificity	.55
PPP	.66
NPP	.54
OCC	.61
Cutoff *T* score of 63 (observed base rate)	
Sensitivity	.52
Specificity	.70
PPP	.70
NPP	.52
OCC	.60

NOTE: DFA = discriminant function analysis. PPP = positive predicative power; NPP = negative predictive power; OCC = overall correct classification. Sensitivity = true positives / (true positives + false negatives); specificity = true negatives / (true negatives + false positives); PPP = true positives / (true positives + false positives); NPP = true negatives / (true negatives + false negatives); OCC = (true positives + true negatives) / (true positives + true negatives + false positives + false negatives).

level of 95.7% in correctly classifying patients (overall correct classification rate was 53.6%; see Table 2).

Clinicians may benefit from knowing optimal individual cutoff scores for the ATR scale, derived from the present study to discriminate between PTSD simulators and patients. A *T* score of 58 yielded sensitivity of 69.5% in detecting PTSD simulators and specificity of 42.6%. A *T* score of 61 yielded sensitivity of 65.1% and specificity of 55.3%. A *T* score of 63 yielded sensitivity of 52.4% and specificity of 70.2%. However, no cutoff score minimized both false positive and false negative rates to below 30% (see Table 2 for diagnostic efficiency estimates). For clinicians concerned with misclassifying a high-scoring patient as a PTSD malingerer in a clinical setting, fewer than 5% of PTSD patients scored higher than 87, and fewer than 2% scored higher than 92 on ATR. Thus, a *T* score above 92 is likely to yield few misclassifications of PTSD patients as malingerers.

Because of the relatively small sample size, we conducted a post hoc power analysis (using SPSS Sample Power 2.0). Assuming an alpha level of .05, two tailed, we did not achieve adequate (80%) power for several of our ANOVAs. However, the main concern for power is failing to find a significant relationship when one actually exists (Type II error). For nonsignificant ANOVAs, we would have needed an average of nearly 20,000 participants (range was 2,500 to 78,000) to achieve adequate power, suggesting that our nonsignificant scales reflected the lack of true significant group differences.

DISCUSSION

This unique study directly examined the ability of the TSI to discriminate PTSD simulators from clinical PTSD patients. Overall, PTSD simulators scored significantly higher than PTSD patients on the ATR, DSB, and TRB scales. On the other hand, PTSD patients scored higher on the AA and SC scales. ATR was inadequate in correctly classifying simulators and patients.

Thus, the TSI evidenced few scale differences in the expected direction (i.e., simulators patients). It could be argued that the lack of more robust differences was due to the ineffectiveness of our simulation instructions. However, we found more robust effects and better classification accuracy in our previous studies using very similar methodology, assessing the MMPI-2 validity scales (e.g., Fp and Fptsd) (Elhai et al., 2000, 2001, 2002). Thus, it appears that the present study's manipulation may not have accounted for the less robust effects; instead, problems with the construct validity of the TSI in detecting simulated PTSD may be at fault. It should be noted, however, that ATR was not developed specifically as a malingered

PTSD screen (Briere, 1995), instead serving as a screen of general validity. Furthermore, the TSI itself is a general measure of posttraumatic symptomatology rather than being strictly a PTSD symptom measure (Briere, 1995). Thus, although clinically used and empirically investigated as a malingered PTSD screen, it was not apparently intended for this specific purpose.

Regarding the TSI clinical scales, DSB and TRB were the only ones on which PTSD simulators scored higher than patients. Interestingly, these two scales seem to be related, both rationally and empirically. Rationally, these scales both involve maladaptive, risk-taking behaviors. DSB taps inappropriate, dysfunctional, or indiscriminate sexual behavior (Briere, 1995). TRB involves dysfunctional behaviors used to modulate negative affect, including inappropriate sexual behavior as well as self-mutilation, aggression, and suicidality (Briere, 1995). Empirically, TSI exploratory and confirmatory factor analyses have demonstrated that DSB and TRB both belong to the same factor (Self-Dysfunction for exploratory, and Self for confirmatory factor analyses) (Briere, 1995). Thus, on the TSI it appears that PTSD simulators overestimated the extent to which PTSD patients suffer from these maladaptive, risky behaviors. However, it is difficult to explain why the simulator group overendorsed problematic sexual behaviors because the sample's majority reported using case studies in their simulation that were not related to sexual trauma.

There were two clinical scales in which PTSD patients scored higher than simulators: AA and SC. AA measures autonomic hyperarousal symptoms (e.g., trembling, exaggerated startle response), resembling PTSD's Criterion D (Briere, 1995). Hyperarousal symptoms are also often seen in patients diagnosed with panic disorder. In fact, panic disorder was diagnosed in nearly 40% of the PTSD patient sample, which is quite common in PTSD patients in general (Kessler et al., 1995). However, PTSD simulators underestimated the extent to which these hyperarousal symptoms would be present in PTSD patients and reported the least confidence in this cluster of PTSD symptoms on the feedback survey.

SC was the other clinical scale that was higher in patients than simulators. SC involves distress and dysfunction concerning sex, including sexual confusion, unwanted preoccupation with sex, and sexual problems in relationships (Briere, 1995). Thus, simulators underestimated the extent of patients' sexual concerns. However, it is noteworthy that simulators underendorsed concerns with sexual functioning (SC) because they overendorsed problematic sexual behaviors (DSB).

In evaluation contexts where compensation is clearly at stake (e.g., disability assessments), the Atypical Response scale may not be the best measure for detecting symptom overreporting. In fact, ATR's cutoff scores never achieved less than 30% error rates for both sensitivity and specificity. The optimal cutoff score of 61 was the same as that found in Edens et al.'s (1998) examination of simulated PTSD on the TSI. However, their sensitivity and specificity rates were much higher than in the present study, probably due to the fact that their comparison group was composed of students instructed to complete the TSI honestly (thus yielding lower ATR scores than a clinical group and making it easier to discriminate from a simulation group). Another explanation for differences in sensitivity and specificity rates across these studies is that this study used standardized assessments of trauma exposure and PTSD to screen out potential simulators who might have already had PTSD (resulting in lower simulator group mean scores that are more difficult to detect from a clinical patient group; see the possible PTSD status analyses, above).

The ATR cutoff scores of the study also performed much worse than those derived from MMPI-2 fake bad scales in previous PTSD simulation studies. For example, the MMPI-2's Infrequency-Psychopathology (Fp) scale optimally classified 74% of PTSD simulators and 89% of patients in one study (Wetter et al., 1993) and between 75% and 83% of simulators and 76% and 78% of patients in another study (Elhai et al., 2001). The Infrequency (F) scale correctly classified 75% of PTSD simulators and 77% of patients in an additional study (Elhai et al., 2000). These classification rates are much higher than that found for ATR in the present study, which yielded an overall correct classification rate of 59%, sensitivity of .48, and specificity of .75 in DFA. Thus, these MMPI-2 fake bad scales have an established record of detecting simulated PTSD, with greater accuracy than ATR.

Some consideration should be given regarding why ATR did not perform well in the study. ATR was developed from a pool of items tapping bizarre, unusual, and psychotic phenomena, of which its 10 items were those least frequently endorsed in the TSI's standardization sample, and subsequently combined with the clinical items (Briere, 1995; J. Briere, personal communication, January 23, 2005). Thus, two problems exist with its development: (a) No specific infrequency criterion (e.g., 10%) was used to select ATR items, possibly resulting in less robust infrequency items, and (b) item infrequency was determined exclusively from healthy individuals' responses, whereas the malingering literature has revealed that infrequency scales developed with psychiatric patients are significantly better in detecting malingered psychopathology (Arbisi & Ben-Porath, 1995, 1998; Elhai et al., 2001, 2002, 2004; Storm & Graham, 2000).

Because of the unimpressive hit rates for ATR, clinicians should be strongly cautioned when using ATR to de-

termine malingered PTSD. As revealed in our cutoff score and DFA, although using the TSI manual's recommended ATR cutoff score of 90 would result in a very low false positive rate for identifying genuine patients as malingerers, it would likely allow for numerous individuals who are actually feigning PTSD to be identified as honest responders. However, in clinical settings false positives are probably more detrimental than false negatives, because a false positive may potentially result in denying a genuine patient treatment that is sorely needed. Moreover, our DFAs demonstrate that within the context of clinical settings (i.e., having low malingering base rates), ATR results in relatively low false positive rates. Nonetheless, we again emphasize that in forensic settings, ATR may seriously underestimate malingering, and missing a feigned test protocol results in clinical interpretation of an invalid profile.

It is strongly advised that the TSI's ATR scale not be used alone in evaluating the possibility of feigned PTSD. Instead, other measures with established general indicators of malingering, such as the MMPI-2 or Structured Interview of Reported Symptoms (Rogers, Bagby, & Dickens, 1992), should be used when assessing psychiatric symptom accuracy in PTSD patients. This multimethod assessment should help minimize false positive and false negative classifications.

Several limitations were present in the study. First, our sample (which was relatively small) was composed of college students; we do not know how equivalent their simulation performances are to those of community participants' simulation (or to those of actual malingerers, for that matter). Specifically, because our student sample was quite young, and relatively well-educated compared to the general population, our conclusions regarding the TSI's malingering detection utility may not be generalizable to nonclinical populations. Nonetheless, college student samples are viewed as appropriate when validating a new assessment measure or procedure (Butcher, Graham, & Ben-Porath, 1995). Given the dearth of relevant TSI literature, we judged that our college sample inclusion was appropriate in avoiding unnecessary costs and resources in recruiting community participants at this early stage of validating the TSI in malingered PTSD detection.

Additional limitations were present as well. For example, our method of screening simulators for genuine PTSD prior to simulation was not comprehensive. The Life Events Checklist is not intended for group administration, because follow-up queries are recommended (Weathers & Keane, 1999). Furthermore, the PTSD Checklist is only an approximation of PTSD diagnosis and does not allow for clinical judgment, as in a structured PTSD interview. Next, despite our screening for PTSD among simulators prior to simulation, we were unable to screen for external incentives to malinger among the PTSD patient sample. Although no external incentive appeared available for fabricating or exaggerating problems at the mental health center to which they presented, it is possible that at least some genuine patients had overreported their psychiatric symptoms. Also, although the PTSD patient sample had significant psychiatric comorbidity, simulators were not trained or instructed to simulate any specific psychiatric condition other than PTSD, thus potentially limiting the degree to which simulators were analogous to patients.

Last, although specific case studies chosen for simulation did not affect TSI scores, we must acknowledge the limitation that the majority of PTSD patients reported sexual victimization, whereas the majority of simulators chose the motor vehicle accident case study. This problem is especially relevant because two of the TSI scales that differed between patients and simulators involved maladaptive sexual functioning. Moreover, the majority of the patient sample was composed of women, who generally endorse higher rates of sexual victimization than do men. In fact, higher scores on several TSI scales (e.g., AA) are noted for women, and higher scores on scales such as SC and AA are noted for sexual abuse victims, which could have influenced our simulation findings. Replication of these findings is necessary, providing simulators with a case study more relevant to that of the clinical comparison group.

APPENDIX
Post-Traumatic Stress Disorder (PTSD) Simulation Instructions

In this experiment, we are interested in exploring how well people can pretend to have post-traumatic stress disorder (PTSD). Please read the instructions printed at the beginning of the test booklet (the standard instructions). However, when completing the test, instead of responding to the items honestly, we would like you to answer the questions as if you have the experiences of PTSD, described in the information packet. Your goal is to deceive a psychological examiner by pretending that you have PTSD. Imagine that you are attempting to convince this examiner, based on your test results, that you are sufficiently disturbed with PTSD to either (a) obtain money for the PTSD problems you claim to have (e.g., such as in a lawsuit against a person or company you claim hurt you psychologically and caused you to have PTSD, or to collect social security disability payments from the government for the PTSD you claim to have) or (b) to avoid prosecution for a crime that you say you committed because of the PTSD you claim to have (e.g., the insanity defense). Try to do your best, because the 3 best PTSD fakers in each group of about 30 participants will receive checks for $50, $45, or $40 (respectively) later this month. The odds of someone winning one of these prizes is approximately 1 in 10. However, please also realize that if you present too dramatically, it will look fake, and you will not be believed. Therefore, use caution to maintain a convincing, realistic profile.

NOTE

1. These data were used in a previous, unrelated article by Davis et al. (in press).

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Arbisi, P. A., & Ben-Porath, Y. S. (1995). An MMPI-2 infrequent response scale for use with psychopathological populations: The Infrequency Psychopathology scale, F(p). *Psychological Assessment*, *7*, 424-431.

Arbisi, P. A., & Ben-Porath, Y. S. (1998). The ability of Minnesota Multiphasic Personality Inventory–2 validity scales to detect fake-bad responses in psychiatric inpatients. *Psychological Assessment*, *10*, 221-228.

Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Klauminser, G., Charney, D. S., et al. (1990). A clinician rating scale for assessing current and lifetime PTSD: The CAPS-1. *Behavior Therapist*, *18*, 187-188.

Briere, J. (1995). *Trauma Symptom Inventory professional manual*. Odessa, FL: Psychological Assessment Resources.

Briere, J. (2004). *Psychological assessment of adult posttraumatic states: Phenomenology, diagnosis, and measurement*. Washington, DC: American Psychological Association.

Briere, J., Elliott, D. M., Harris, K., & Cotman, A. (1995). Trauma Symptom Inventory: Psychometrics and association with childhood and adult victimization in clinical samples. *Journal of Interpersonal Violence*, *10*, 387-401.

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., Graham, J. R., & Ben-Porath, Y. S. (1995). Methodological problems and issues in MMPI, MMPI-2, and MMPI-A research. *Psychological Assessment*, *7*, 320-329.

Butcher, J. N., Graham, J. R., Ben-Porath, Y. S., Tellegen, A., Dahlstrom, W. G., & Kaemmer, B. (2001). *MMPI-2 (Minnesota Multiphasic Personality Inventory-2): Manual for administration, scoring, and interpretation* (Rev. ed.). Minneapolis: University of Minnesota Press.

Cohen, J. (1990). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillside, NJ: Lawrence Erlbaum Associates.

Davis, J. L., Borntrager, C., Combs-Lane, A., Wright, D., Elhai, J. D., Falsetti, S. A., et al. (in press). Comparison of racial groups on trauma and post-trauma functioning. *Journal of Trauma Practice*.

Edens, J. F., Otto, R. K., & Dwyer, T. J. (1998). Susceptibility of the Trauma Symptom Inventory to malingering. *Journal of Personality Assessment*, *71*, 379-392.

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., Naifeh, J. A., Zucker, I. S., Gold, S. N., Deitsch, S. E., & Frueh, B. C. (2004). Discriminating malingered from genuine civilian posttraumatic stress disorder: A validation of three MMPI-2 infrequency scales (F, Fp, and Fptsd). *Assessment*, *11*, 139-144.

Elhai, J. D., Ruggiero, K. J., Frueh, B. C., Beckham, J. C., Gold, P. B., & Feldman, M. E. (2002). The Infrequency-Posttraumatic Stress Disorder scale (Fptsd) for the MMPI-2: Development and initial validation with veterans presenting with combat-related PTSD. *Journal of Personality Assessment*, *79*, 531-549.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. (1996). *Structured Clinical Interview for DSM-IV Axis I Disorders, Clinician Version (SCID-CV)*. Washington, DC: American Psychiatric Press.

Frueh, B. C., Hamner, M. B., Cahill, S. P., Gold, P. B., & Hamlin, K. (2000). Apparent symptom overreporting among combat veterans evaluated for PTSD. *Clinical Psychology Review*, *20*, 853-885.

Gray, M. J., Litz, B. T., Wang, J., & Lombardo, T. W. (2004). Psychometric properties of the Life Events Checklist. *Assessment*, *11*, 330-341.

Kessler, R. C., Sonnega, A., Bromet, E., Hughes, M., & Nelson, C. B. (1995). Posttraumatic stress disorder in the National Comorbidity Survey. *Archives of General Psychiatry*, *52*, 1048-1060.

Flitter, J. M. K., Elhai, J. D., & Gold, S. N. (2003). MMPI-2 F scale elevations in adult victims of child sexual abuse. *Journal of Traumatic Stress*, *16*, 269-274.

McDevitt-Murphy, M. E., Weathers, F. W., & Adkins, J. W. (in press). The use of the Trauma Symptom Inventory in the assessment of posttraumatic stress disorder. *Journal of Traumatic Stress*.

Resnick, H. S. (1996). Psychometric review of Trauma Assessment for Adults (TAA). In B. H. Stamm (Ed.), *Measurement of stress, trauma, and adaptation* (pp. 362-364). Lutherville, MD: Sidran Press.

Resnick, H. S., Falsetti, S. A., Kilpatrick, D. G., & Freedy, J. R. (1996). Assessment of rape and other civilian trauma-related PTSD: Emphasis on assessment of potentially traumatic events. In T. W. Miller (Ed.), *Theory and assessment of stressful life events* (2nd ed., pp. 235-271). Madison, CT: International Universities Press.

Resnick, H. S., Kilpatrick, D. G., Dansky, B. S., Saunders, B. E., & Best, C. L. (1993). Prevalence of civilian trauma and posttraumatic stress disorder in a representative national sample of women. *Journal of Consulting and Clinical Psychology*, *61*, 984-991.

Resnick, P. J. (1997). Malingering of posttraumatic disorders. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed., pp. 130-152). New York: Guilford.

Rogers, R. (1997). Researching dissimulation. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (2nd ed., pp. 398-426). New York: Guilford.

Rogers, R., Bagby, R. M., & Dickens, S. E. (1992). *Structured Interview of Reported Symptoms: Professional manual*. Tampa, FL: Psychological Assessment Resources.

Rogers, R., Salekin, R. T., Sewell, K. W., Goldstein, A., & Leonard, K. (1998). A comparison of forensic and nonforensic malingerers: A prototypical analysis of explanatory models. *Law and Human Behavior*, *22*, 353-367.

Rogers, R., Sewell, K. W., & Goldstein, A. (1994). Explanatory models of malingering: A prototypical analysis. *Law and Human Behavior*, *18*, 543-552.

Rogers, R., Sewell, K. W., Martin, M. A., & Vitacco, M. J. (2003). Detection of feigned mental disorders: A meta-analysis of the MMPI-2 and malingering. *Assessment*, *10*, 160-177.

Ruggiero, K. J., Del Ben, K., Scotti, J. R., & Rabalais, A. E. (2003). Psychometric properties of the PTSD Checklist–Civilian Version. *Journal of Traumatic Stress*, *16*, 495-502.

Storm, J., & Graham, J. R. (2000). Detection of coached general malingering on the MMPI-2. *Psychological Assessment*, *12*, 158-165.

Weathers, F. W., & Keane, T. M. (1999). Psychological assessment of traumatized adults. In P. A. Saigh & J. G. Bremner (Eds.), *Posttraumatic stress disorder: A comprehensive text* (pp. 219-247). Boston: Allyn & Bacon.

Weathers, F. W., Litz, B. T., Herman, D. S., Huska, J. A., & Keane, T. M. (1993, October). *The PTSD checklist: Reliability, validity, & diagnostic utility*. Paper presented at the annual meeting of the International Society for Traumatic Stress Studies, San Antonio, TX.

Wetter, M. W., Baer, R. A., Berry, D. T. R., Robison, L. H., & Sumpter, J. (1993). MMPI-2 profiles of motivated fakers given specific symptom information: A comparison to matched patients. *Psychological Assessment*, *5*, 317-323.

Jon D. Elhai, Ph.D., is an assistant professor in the Disaster Mental Health Institute at the University of South Dakota. He specializes in assessment and health services research among traumatic event victims.

Matthew J. Gray, Ph.D., is an assistant professor of psychology at the University of Wyoming. He trained at the National Crime Victims Center and completed a postdoctoral fellowship at the National Center for PTSD, Behavioral Sciences Division, in Boston.

James A. Naifeh, B.A., is a doctoral student in the Disaster Mental Health Institute and Clinical Psychology Training Program at the University of South Dakota. His research interests are in the assessment of traumatic event exposure and PTSD.

Jimmie J. Butcher, M.A., is a doctoral candidate in the Clinical Psychology Training Program and a Morgan Fellow in the Disaster Mental Health Institute at the University of South Dakota.

Joanne L. Davis, Ph.D., is an assistant professor at the University of Tulsa. She specializes in the assessment, prevention, and treatment of trauma and its sequelae.

Sherry A. Falsetti, Ph.D., is an associate professor and director of Behavioral Sciences in the Department of Family and Community Medicine at the University of Illinois College of Medicine at Rockford. She specializes in the treatment of PTSD and panic attacks.

Connie L. Best, Ph.D., is a professor in the National Crime Victims Research and Treatment Center at the Medical University of South Carolina. Her research interests are in posttraumatic stress disorder and the psychological impact of criminal victimization and other trauma.