Journal of Psychopathology and Behavioral Assessment, Vol. 28, No. 3, September 2006 (© 2006)
DOI: 10.1007/s10862-005-9006-5

Detection of Feigned Posttraumatic Stress Disorder: A Comparison of the MMPI-2 and PAI

David E. Eakin,[1] Frank W. Weathers,[1,2] Trisha B. Benson,[1] Christopher F. Anderson,[1] and Brandice Funderburk[1]

Published online: 28 December 2005

The Minnesota Multiphasic Personality Inventory-2 (MMPI-2) and Personality Assessment Inventory (PAI) were compared for detecting feigned posttraumatic stress disorder (PTSD) in a simulation research design. Participants were 85 undergraduates in one of three groups: PTSDs ($n = 23$), Fakers ($n = 31$), and Controls ($n = 31$). As expected, both the MMPI-2 and PAI discriminated PTSDs and Controls, with PTSDs scoring significantly higher on fake-bad validity scales and PTSD-relevant clinical scales. However, only the MMPI-2 discriminated Fakers and PTSDs, with Fakers scoring significantly higher on all MMPI-2 scales considered, but on only one PAI scale. Further, in logistic regression analyses the MMPI-2 demonstrated higher overall correct classification of PTSDs and Fakers than did the PAI. Although the MMPI-2 outperformed the PAI in detecting feigned PTSD, a substantial proportion of Fakers avoided detection by MMPI-2 fake-bad validity scales, suggesting that both tests are vulnerable to feigning of PTSD by motivated respondents with relatively limited coaching.

KEY WORDS: malingering; posttraumatic stress disorder; MMPI-2; PAI.

Since its introduction in the DSM-III (American Psychiatric Association; APA, 1980), posttraumatic stress disorder (PTSD) has gained wide recognition as a valid clinical syndrome. PTSD has also gained status as a compensable disorder, unfortunately increasing the potential for malingering within Veterans Affairs and in forensic settings and civil litigation. Consequently, a large body of research has been conducted to develop psychometrically sound methods for detecting malingered PTSD (Guriel & Fremouw, 2003). Much of this research has centered on multiscale personality inventories, especially the Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1943, 1951) and MMPI-2 (Butcher et al., 2001), as well as the newer Personality Assessment Inventory (PAI; Morey, 1991).

The MMPI/MMPI-2 is the most popular psychological assessment instrument (Groth-Marnat, 2003) and the most frequently employed in forensic evaluations (Pope, Butcher, & Seelen, 2000). Most MMPI/MMPI-2 PTSD malingering research has focused on the validity scales, especially F, F_B, F_P, O-S, and Ds (Rogers, Sewell, Martin, & Vitacco, 2003). These scales have demonstrated utility in some contexts, particularly in studies involving uncoached dissimulators of severe psychopathology (e.g., Bagby et al., 1997; Rogers, Sewell, & Salekin, 1994), including faking of PTSD by Vietnam combat veterans (e.g., Baldrachi, Hilsenroth, Arsenault, Sloan, & Walter, 1998; McCaffrey & Bellamy-Campbell, 1989; Fairbank, McCaffrey, & Keane, 1985). Typically, participants in such studies are provided with information about symptoms, but not about the presence or function of validity scales. When participants are informed about validity scales, however, the utility of these scales for detecting dissimulators may be diminished (Walters & Clopton, 2000; Bagby et al., 1997; Lamb, Berry, Wetter, & Baer, 1994; Rogers, Bagby, & Chakraborty, 1993a).

The more recently developed PAI is another multiscale inventory that may be useful in the detection of malingered PTSD. Like the MMPI-2, the PAI contains

[1]Department of Psychology, Auburn University, 226 Thach Hall, AL 36849.
[2]To whom correspondence should be addressed at e-mail: weathfw@auburn.edu.

several response validity scales, including the Inconsistency (ICN), Infrequency (INF), Negative Impression (NIM), and Positive Impression (PIM) scales. Morey (1996) identified the NIM scale as the beginning point for the detection of malingered psychopathology, although NIM was not specifically developed as a malingering scale. Rogers, Ornduff, and Sewell (1993b) investigated the ability of NIM to detect feigned schizophrenia, depression, and generalized anxiety disorder (GAD) in a simulation study. NIM was successful at detecting feigned schizophrenia, but was much less successful at detecting feigned depression and GAD.

In addition to NIM, two PAI indexes have been derived specifically for the detection of malingering. The Malingering Index (MAL; Morey, 1996) consists of eight indicators seen with substantially greater frequency in individuals attempting to simulate severe psychopathology than in actual clinical samples. In contrast to NIM, MAL has shown greater specificity for detecting malingering of severe psychopathology. The Rogers Discriminant Function index (RDF; Rogers, Sewell, Morey, & Ustad, 1996) was developed from a two-stage discriminant analysis comparing sophisticated simulators, naïve simulators, and a clinical sample. Although the RDF showed strong discrimination of simulators and patients in the original study, it performed more poorly in a subsequent study conducted in an actual forensic setting, demonstrating only chance levels of detection of known feigners (Rogers, Sewell, Cruise, Wang, & Ustad, 1998).

Two studies have investigated the PAI and malingering of PTSD. Liljequist, Kinder, and Schinka (1998) compared the PAI profiles of undergraduate feigners and a clinical sample of Vietnam veterans with PTSD. Among the validity scales, NIM demonstrated the greatest sensitivity to feigned PTSD, while MAL correctly classified all actual PTSD patients and approximately half of feigners. Although feigners and patients both produced significant elevations on most PAI scales relative to controls, seven PAI scales effectively discriminated feigners from patients. In a similar study, Calhoun, Earnst, Tucker, Kirby, and Beckham (2000) found that NIM and MAL demonstrated moderate effectiveness in discriminating feigners from patients.

To date, few studies have directly compared the MMPI-2 and PAI with respect to the detection of feigned psychopathology. In one such study, Bagby, Nicholson, Bacchiochi, Ryder, and Bury (2002) reported that RDF was the best PAI predictor in discriminating attempts to feign a mental disorder by individuals who had been coached about the presence of validity scales. In this study, the RDF index performed as well or better than the family of F scales on the MMPI-2. No other PAI validity scale or index was found to offer predictive validity in this study.

Most current studies of malingering are simulation studies, in which normal participants are instructed to pretend to have a psychological disorder and are given some incentive for successful faking (Rogers, 1997; Schretlen, 1988). However, the applicability of such studies to real-world clinical and forensic settings is a concern. To improve the relevance and clinical utility of simulation studies, Rogers (1997) identified several guidelines that should be incorporated into the research design. First, instructions must be clear, concise, and easily grasped by participants. Vague instructions, such as "pretend to have a mental illness," have limited external validity. Second, participants must be adequately prepared and coached. Most malingerers are knowledgeable about the syndrome they are feigning and are often informed about validity scales on tests when referred for a psychological evaluation (Wetter & Corrigan, 1995). Third, adequate incentives must be included, in part to elicit an analog for the motivation that real-world incentives supply, and in part because external incentive is an integral aspect of the diagnostic criteria for malingering. Fourth, debriefing should be conducted to ensure that participants understood their role and fully complied with instructions.

A potentially crucial issue regarding the ecological validity of research on malingering and PTSD is trauma exposure in research participants. Surprisingly, few simulation studies of malingered PTSD have used trauma-exposed but asymptomatic participants in the faking groups (e.g., Bury & Bagby, 2002; Scheibe, Bagby, Miller, & Dorian, 2001; Hickling, Taylor, Blanchard, & Devineni, 1999). In most clinical and forensic settings, unless an individual claiming to have PTSD can document trauma exposure, there is no need to determine whether any reported symptoms are genuine or malingered. Further, trauma-exposed dissimulators pose a more realistic challenge for psychometric detection than do non-trauma-exposed dissimulators (Elhai, Gold, Sellers, & Dorfman, 2001).

The present investigation involved a simulation study comparing the ability of the MMPI-2 and PAI to discriminate three groups of trauma-exposed individuals: participants with PTSD (PTSDs) and well-adjusted participants instructed either to fake PTSD (Fakers) or to answer honestly (Controls). The study was designed to address as many of the design issues raised by Rogers (1997) as possible. Instructions were written to be concise, comprehensible, and standardized across groups. Symptom information was provided in clear and simple language and was supplemented with videotapes in which actual trauma survivors described the nature and impact of their

trauma exposure and PTSD symptoms. To evaluate their comprehension of the PTSD syndrome participants completed a brief quiz in each phase of the study. Information was also provided about the presence and purpose of validity scales on the MMPI-2 and PAI. Finally, a $50 cash incentive was offered as a potential reward for carefully attending to the instructions.

The specific hypotheses tested in the present study were based on previous research and focused primarily on two comparisons: PTSDs v. Controls and PTSDs v. Fakers. First, for both the MMPI-2 and PAI it was predicted that PTSDs would score higher than Controls on fake-bad validity scales and PTSD-relevant clinical scales, reflecting the sensitivity of these scales to PTSD. Second, it was predicted that Fakers would score higher than PTSDs on the same scales, reflecting the tendency for malingerers to overendorse psychopathology relative to bona fide clinical samples on multiscale inventories. Third, it was predicted that logistic regression analyses involving these scales would replicate the moderate to strong correct classification rates (.60–.80) for PTSDs and Fakers previously reported in the literature. Finally, the primary exploratory question for the study involved the relative effectiveness of the MMPI-2 and PAI for the detection of dissimulated PTSD.

METHOD

Participants

Participants were 85 (26 male, 59 female) undergraduates at a large southeastern university. They were predominantly European American (87%), unmarried (96%), and in their first or second year of college (80%). Mean age was 19.6 years ($SD = 1.2$). The Life Events Checklist (LEC), the trauma history portion of the Clinician-Administered PTSD Scale (CAPS; Blake et al., 1995), was administered to assess exposure to a traumatic life event. All participants reported experiencing at least one event that satisfied the two-part definition of a trauma in Criterion A for PTSD in the DSM-IV (APA, 1994).

Participants were screened for possible PTSD with the PTSD Checklist (PCL; Weathers, Litz, Herman, Huska, & Keane, 1993), a 17-item, DSM-correspondent self-report measure of PTSD symptom severity. On the PCL, participants indicate how much they have been bothered in the past month by each PTSD symptom, using a 5-point scale ranging from 1 (not at all) to 5 (extremely). The status of probable PTSD was determined by treating PCL items rated as 3 (moderately) or higher as symptom endorsements and following the DSM-IV PTSD diagnostic rule (one reexperiencing symptom, three avoidance and numbing symptoms, two hyperarousal symptoms).

Participants who met criteria for probable PTSD were further evaluated with the CAPS, which was administered by advanced doctoral students in clinical psychology or by the second author, a licensed clinical psychologist and co-author of the CAPS. All interviewers were trained and supervised in the use of the CAPS by the second author. Fifty-four participants screened positive for possible PTSD on the PCL, of whom 39 agreed to further participation and were administered the CAPS. Of these 39, 23 met criteria for PTSD based on the F1/I2 CAPS scoring rule (Weathers, Ruscio, & Keane, 1999) and were included in the PTSD group.

To be included in the Faker or Control groups, participants had to (a) report having experienced a Criterion A event within the past 3 years, (b) screen negative for possible PTSD on the PCL, and (c) agree to further participation. The requirement that the index event have occurred within the past 3 years was intended to assure that the event would be sufficiently recent in memory to serve as the basis for faking PTSD symptoms. Sixty-two participants met all of these criteria and were randomly assigned to the fake group (Fakers; $n = 31$) or the control group (Controls; $n = 31$). Fakers and Controls were not administered the CAPS.

The three groups did not differ on any of the demographic variables, but they did differ with respect to the types of trauma experienced. The LEC lists 17 categories of traumatic events, including a category of "other" to capture traumatic experiences that do not fit any of the other 16 categories. Because of the low frequencies in many of the trauma categories, five summary categories were created for this analysis, including (a) accidents (combining transportation accident and other serious accident); (b) physical assaults (combining physical assault with and without a weapon); (c) sexual assaults (combining sexual assault and other unwanted sexual experience); (d) sudden death of someone close (combining sudden violent death and sudden death of someone close to you); and (e) all other (combining natural disaster, fire or explosion, exposure to toxic substances, combat, captivity, severe human suffering, life threatening illness, death you caused to someone else, and other stressful experiences).

Chi-square analysis revealed significant differences ($\chi^2 = 19.34$; $p < .01$; $df = 8, 85$) in the types of trauma experienced among the three groups. The types of experiences reported by Controls and Fakers were quite similar. Within these two groups the most frequent trauma types were transportation accident (48% of Controls and 42% of Fakers) and sudden death of someone close (29% of both groups). In contrast, among

PTSDs, the most frequent trauma types were sexual assault (39%), physical assault (17%) and sudden death of someone close (17%). This pattern is consistent with the results of the National Comorbidity Survey (Kessler, Sonnega, Bromet, Hughes, & Nelson, 1995), which found that sexual assault and physical assault are the types of traumatic experiences most likely to result in PTSD. Thus, although these group differences in trauma exposure are significant, it is likely they are characteristic of mixed civilian trauma samples.

Procedure

All participants completed the MMPI-2 and PAI in counterbalanced order over two sessions. At the beginning of each session, all participants were read standardized instructions that varied by group. PTSDs and Controls were informed about the presence and nature of response validity scales on the tests and were told that their eligibility for a chance at earning the cash incentive depended upon their test scores indicating adequate validity. The cash incentive consisted of being entered into a drawing at the end of the study for one of five $50 cash prizes. They were then read the standard test instructions and were asked to follow them in completing the test.

In their first session, Fakers were given a narrative lecture, along with a written copy of the lecture content, that outlined common reasons for malingering of PTSD (e.g., fraudulently obtaining disability or insurance payments, avoidance of criminal prosecution), and were given comprehensive descriptions of PTSD symptoms and common comorbid problems such as substance abuse and depression. The lecture was supplemented by a 20 min video, consisting of selected excerpts from a video entitled *Understanding Psychological Trauma, Part 1: Understanding Survivors* (Doepel, 1989), in which actual trauma survivors described their symptoms. At the completion of training in the first session, the written materials were collected and a 10-item True-False quiz was administered to evaluate participants' comprehension of the material. In the second session, the symptom information was briefly reviewed and an alternate form of the quiz was administered before administration of the second personality inventory. A score of 90% on each quiz was required for eligibility to continue in the study. No participant was excluded on this basis.

Fakers were then instructed to respond to MMPI-2 and PAI items as if they had PTSD, using their own traumatic event as a reference, and were encouraged to respond to test items in a way that would make them appear to have PTSD. Fakers were also informed about the validity scales and were instructed to feign PTSD without being detected by these scales. They were told that their eligibility for a chance at earning the cash incentive depended upon their test scores indicating that they carefully followed instructions, i.e., attempted to appear to have PTSD without being detected by the validity scales.

After completing participation, participants were debriefed regarding the purpose of the study and were provided with a list of mental health resources to refer to in the event their participation resulted in psychological distress. All participants received course extra credit for their participation. At the end of data collection, all participants who completed all phases of the study were included in the drawing for the cash incentive regardless of their test scores, per institutional review board requirements.

RESULTS

Following the procedure used in previous research (Bagby et al., 2002; Storm & Graham, 2000), validity profiles of each test were examined for inconsistent or invalid overall responding. Using a cutoff of TRIN < 5 or ≥ 13 and a VRIN ≥ 13 (raw scores) on the MMPI-2 and ICN ≥ 73 and INF ≥ 75 on the PAI, one Control and two Fakers were excluded on this basis. Therefore, group sizes used throughout the remainder of the analyses were: PTSDs, $n = 23$; Controls, $n = 30$; and Fakers, $n = 29$.

Group Means and Comparison Effect Sizes

To test hypothesized group differences, descriptive statistics and Cohen's *d* effect sizes (Rosenthal, Rosnow, & Rubin, 2000) were calculated. MMPI-2 results are presented in Table I and PAI results are presented in Table II. For both the MMPI-2 and PAI it was predicted that PTSDs would score significantly higher than Controls on the fake-bad and PTSD-relevant scales, and that Fakers would score significantly higher than PTSDs on the same scales. For the MMPI-2 both hypotheses were supported. For PTSDs v. Controls, large effect sizes were found for all the scales tested, including the fake-bad validity scales (F, F_B, F_P, Ds) and PTSD-relevant clinical scales (2, 7, 8, Pk). For PTSDs v. Fakers, large effect sizes were found for all scales except Pk, which had a moderate effect size.

For the PAI, the first hypothesis, regarding PTSDs v. Controls, was generally supported, with large effect sizes obtained on the NIM validity scale and all PTSD-relevant clinical scales (ANX, DEP, BOR, and ARD-T). Regarding the other two validity indexes, the effect size was moderate for MAL and nonsignificant for RDF. The

Table I. MMPI-2 Group Means and Comparison Effect Sizes (d)

	Group								
	Control ($n = 30$)		PTSD ($n = 23$)		Faker ($n = 29$)		Effect size (Cohen's *d*)		
Scale	M	*SD*	M	*SD*	M	*SD*	Faker v. Control	PTSD v. Control	Faker v. PTSD
F	50.6	12.5	72.5	19.7	91.1	22.4	2.28***	1.38***	.89**
F_B	50.4	10.2	70.2	24.3	95.5	25.0	2.42***	1.12***	1.04**
F_P	51.5	10.2	60.1	14.4	77.8	21.3	1.61***	.71*	.97**
Ds_2	53.4	11.3	72.2	16.0	86.6	18.2	2.23***	1.40***	.84**
2	52.8	15.6	70.2	13.3	82.2	14.2	2.00***	1.20***	.88**
7	53.4	15.0	70.8	12.2	81.8	10.7	2.21***	1.26***	.98**
8	53.6	13.3	71.6	17.3	87.0	15.4	2.36***	1.20***	.96**
Pk	52.6	15.0	75.1	15.4	85.5	13.6	2.20***	1.50***	.59*

Note. F: infrequency scale; F_B: F-back scale; F_P: infrequency-psychopathology scale; Ds_2: Gough dissimulation scale; 2: depression; 7: psychasthenia; 8: schizophrenia; Pk: Keane PTSD scale.
$^{*}p = <.05$. $^{**}p = <.01$. $^{***}p = <.001$.

second hypothesis, however, regarding Fakers v. Controls, was generally not supported in that Fakers scored significantly higher than PTSDs only on ARD-T.

Regression Analyses

In accordance with previous studies (e.g., Bagby et al., 2002; Bury & Bagby, 2002; Scheibe et al., 2001), a series of regression analyses was performed, using group as the criterion variable and the fake-bad validity scales and PTSD-relevant clinical scales of the MMPI-2 and PAI as the predictor variables. As shown in Table III, results of the PTSD v. Control comparison indicate that for both the MMPI-2 and PAI, all PTSD-relevant clinical scales were significant predictors of group status, both individually and when entered as a block. Validity scales were not tested for the PTSD v. Control comparison, as they are not designed to discriminate bona fide patients from non-patients.

The primary comparison of interest, however, was Fakers v. PTSDs. Validity scale analyses for this comparison are presented in Table IV and clinical scale analyses are presented in Table V. As shown in Table IV, for Fakers v. PTSDs all MMPI-2 fake-bad scales were significant predictors of group status, both individually and when entered as a block. Similarly, as shown in Table V, all PTSD-relevant MMPI-2 clinical scales were

Table II. PAI Group Means and Comparison Effect Sizes (d)

	Group								
	Control		PTSD		Faker		Effect size (Cohen's *d*)		
Scale	M	*SD*	M	*SD*	M	*SD*	Faker v. Control	PTSD v. Control	Faker v. PTSD
Validity indexes and PTSD-relevant clinical scales									
NIM	48.8	6.8	61.5	12.8	63.9	12.8	1.50***	1.30***	.19
MAL	50.6	9.1	59.0	15.3	59.6	12.0	0.86**	0.70*	.04
RDF	48.9	8.2	52.7	11.1	56.7	11.5	0.79**	0.39	.36
ANX	54.5	13.3	71.3	12.6	75.5	14.7	1.52***	1.30***	.32
DEP	50.2	13.9	71.0	14.7	75.2	12.0	1.95***	1.46***	.32
BOR	54.6	12.4	71.1	12.3	67.4	12.1	1.05***	1.34***	.30
ARD-T	54.0	12.1	75.5	13.3	82.8	12.3	2.39***	1.71***	.57*

Note. NIM: negative impression; MAL: malingerig index; RDF: Rogers discriminant function; ANX: anxiety; DEP: depression; BOR: borderline features; ARD-T: anxiety-related disorders-traumatic stress.
$^{*}p = <.05$. $^{**}p = <.01$. $^{***}p = <.001$.

Table III. Logistic Regression Analyses of PTSD-relevant MMPI-2 and PAI Clinical Scales for PTSD v. Control Comparison

	Parameter							
						Correct classification %		
	Model fit χ^2	Wald's Λ	*SE*	Exp(β)	R^2	PTSD	Control	Total
MMPI-2								
2	15.37***	11.41	.02	.93	.25	78	83	81
7	16.86***	11.96	.02	.92	.27	70	77	74
8	16.27***	10.79	.03	.92	.26	74	80	77
Pk	21.61***	14.29	.02	.92	.35	78	83	81
Block:	21.70***				.34	78	83	81
2	–	.07	.05	1.01	–	–	–	–
7	–	.05	.07	.98	–	–	–	–
8	–	.00	.05	1.00	–	–	–	–
Pk	–	3.28	.05	.92	–	–	–	–
PAI								
ANX	17.64***	12.07	.03	.91	.28	70	80	76
DEP	21.98***	12.76	.03	.90	.34	74	87	81
BOR	18.34***	12.47	.03	.91	.29	57	77	68
PAI								
ARD-T	26.69***	14.42	.03	.89	.40	74	83	79
Block	30.38***				.44	70	90	81
ANX	–	.92	.05	1.05	–	–	–	–
DEP	–	1.68	.05	.94	–	–	–	–
BOR	–	.42	.05	.97	–	–	–	–
ARD-T	–	6.97	.04	.90	–	–	–	–

Note. MMPI-2—2: depression; 7: Psychasthenia; 8: Schizophrenia; Pk: Keane PTSD scale. PAI—ANX: Anxiety; DEP: depression; BOR: borderline features; ARD-T: anxiety-related disorders, traumatic stress. For individual scale analyses $df = 1$, $n = 53$. For block analyses $df = 3$, $n = 53$. Parameters for scales analyzed in a block are from the final model
***$p = <.001$.

significant predictors of group status. For the PAI, however, as shown in Table IV, no validity indexes either alone or in combination emerged as a significant predictor of group status for the comparison of Fakers v. PTSDs. Similarly, as shown in Table V, of the PAI PTSD-relevant scales only ARD-T emerged as a significant predictor of group status.

Cut Scores

As shown in Table VI, classification rates for specific cut scores were examined for the PTSD v. Faker comparison. Only the validity scales of the MMPI-2 were included in this analysis. PAI scales were not included because with the exception of ARD-T they did not significantly differentiate the groups. Validity scales were examined because they are the primary scales for detecting fake-bad response sets and have previously established cut scores for detecting malingering. Following Bury and Bagby (2002), classification rates are presented for cut scores based on Rogers et al. (1994). The optimal cut scores based on the present sample are also provided, but given the relatively small sample size and the artificially created base rate, the generalizability of these scores is a concern.

As in Bury and Bagby (2002), the established cut scores demonstrated generally low sensitivity and high specificity, indicating that almost all PTSDs and most Fakers scored below the cutoffs. Thus, in this sample the established cut scores are associated with a low rate of false positives (i.e., are unlikely to incorrectly classify someone with genuine PTSD as a Faker) but a high rate of false negatives (i.e., are likely to incorrectly classify a Faker as having genuine PTSD). This suggests that if these established cut scores were used, a substantial proportion of motivated malingerers with relatively little coaching could evade detection. The optimally efficient cut scores (Kraemer, 1992) derived from the present sample were lower to substantially lower than the established scores and demonstrated a better balance between sensitivity and specificity. Again, however, these are sample-dependent and should not be used without thorough cross-validation.

Table IV. Logistic Regression Analyses of MMPI-2 and PAI Fake-Bad Validity Scales for PTSD v. Faker Comparison

	Parameter							
						Correct classification %		
	Model fit χ^2	Wald's Λ	*SE*	Exp(β)	R^2	PTSD	Faker	Total
MMPI-2								
F	9.08**	7.52	.02	1.04	.16	70	72	71
F_B	11.82**	9.56	.01	1.04	.20	65	76	71
F_P	11.03**	8.32	.02	1.06	.19	65	69	67
Ds_2	8.55**	6.95	.02	1.05	.15	65	72	69
Block	14.11**				.24	61	76	69
F	–	.44	.04	.98	–	–	–	–
F_B	–	2.12	.03	1.04	–	–	–	–
F_P	–	2.12	.03	1.04	–	–	–	–
Ds_2	–	.00	.04	1.00	–	–	–	–
PAI								
NIM	.48	.69	.28	1.26	.01	17	90	57
MAL	.02	.02	.28	1.04	.00	0	100	56
RDF	1.66	1.95	.03	1.03	.03	39	76	60
Block	1.75				.03	35	72	56
NIM	–	.09	.03	1.01	–	–	–	–
MAL	–	.02	.36	.95	–	–	–	–
RDF	–	1.15	.03	1.03	–	–	–	–

Note. MMPI-2—F: infrequency scale; F_B: F-back scale; F_P: infrequency-psychopathology scale; Ds_2: Gough dissimulation scale. PAI—NIM: negative impression management; MAL: malingering index; RDF: Rogers discriminant function. For single scale analyses $df = 1$, $n = 52$. For MMPI-2 block analysis, $df = 4$, $n = 52$; for PAI block analysis $df = 3$, n =52. Parameters for scales analyzed in a block are from the final model.
**$p = <.01$.

DISCUSSION

This study directly compared the ability of the MMPI-2 and PAI fake-bad validity scales and PTSD-relevant clinical scales to detect coached attempts to feign PTSD in a mixed trauma, mixed gender civilian population. The study included a PTSD group from the same population as the dissimulating and control groups. An effort was made to implement each of Rogers's (1997) guidelines for conducting an ecologically valid simulation study, including using concise, comprehensive instructions, coaching participants about the presence of validity scales, creating relatively homogenous comparison groups (including the clinical sample), and providing adequate incentives.

Effect sizes and logistic regression analyses revealed that the fake-bad validity scales and selected clinical scales of the MMPI-2 and PAI were generally comparable in their ability to distinguish PTSDs from Controls. However, for the key comparison of PTSDs v. Fakers, the MMPI-2 outperformed the PAI. All MMPI-2 fake-bad validity scales and PTSD-relevant clinical scales significantly discriminated PTSDs and Fakers, whereas for the PAI only the ARD-T scale did so. Nonetheless, despite the superior performance of the MMPI-2 relative to the PAI, a substantial proportion of Fakers avoided detection by MMPI-2 scales validity scales. This suggests that even with relatively limited coaching and a modest incentive, dissimulators of PTSD can be trained to avoid detection by the validity scales of these multiscale inventories. These results are somewhat divergent from previous studies, which found higher correct classification of malingerers for the MMPI-2 (e.g., Storm & Graham, 2000; Elhai et al., 2001; Lamb et al., 1994). Similarly, the correct classification rates reported for the PAI in previous studies (Bagby et al., 2002; Calhoun et al., 2000; Liljequist et al., 1998) are greater than those found in the present study.

It is important to recognize that substantial differences in research designs exist among previous studies of malingering. For example, in the studies by Bagby et al. (2002) and Storm and Graham (2000), participants were asked to feign a mental disorder. Rogers et al. (1993a, 1993b) and Lamb et al. (1994) asked participants to feign schizophrenia or a closed head injury, respectively. In cases where individuals are asked to feign a mental disorder or to feign psychological distress in general, the

Table V. Logistic Regression Analyses of PTSD-relevant MMPI-2 and PAI Clinical Scales for PTSD v. Faker Comparison

	Parameter							
						Correct classification %		
	Model fit χ^2	Wald's Λ	*SE*	Exp(β)	R^2	PTSD	Faker	Total
MMPI-2 scale								
2	9.05**	7.18	.02	1.07	.16	65	79	73
7	10.96**	8.32	.03	1.09	.19	65	76	71
8	10.72**	8.12	.02	1.06	.19	65	76	71
Pk	4.34*	3.89	.02	1.04	.08	39	79	62
Block	15.94**				.34	78	83	81
2	–	1.50	.04	1.06	–	–	–	–
7	–	.50	.06	1.04	–	–	–	–
8	–	.08	.04	1.08	–	–	–	–
Pk	–	.08	.05	.91	–	–	–	–
PAI scale								
ANX	1.24	1.21	.02	1.02	.02	26	79	56
DEP	1.34	1.28	.02	1.03	.03	39	83	64
BOR	1.21	1.19	.02	.98	.02	30	83	60
ARD-T	4.04*	3.75	.02	1.05	.08	52	76	65
Block	9.05				.16	57	69	64
ANX	–	.51	.04	1.03	–	–	–	–
DEP	–	.62	.04	1.03	–	–	–	–
BOR	–	4.41	.04	.92	–	–	–	–
ARD-T	–	.89	.03	1.03	–	–	–	–

Note. MMPI-2—2: depression; 7: psychasthenia; 8: schizophrenia; Pk: Keane PTSD scale. PAI—ANX: anxiety; DEP: depression; BOR: borderline features; ARD-T: anxiety-related disorders, traumatic stress. For individual scale analyses $df = 1$, $n = 52$. For block analyses $df = 3$, $n = 52$. Parameters for scales analyzed in a block are from the final model.
* $p = <.05$. ** $p = <.01$.

ambiguity of these instructions may tend to amplify respondents' tendency to overendorse symptoms that will elevate validity scale profiles. It seems reasonable to conclude that instructions about the specific disorder to be feigned may enhance participants' ability to replicate response patterns produced by bona fide patients. However, there is some evidence that suggests this may be only partially true. Bagby et al. (1997) found that clinical psychology interns and psychiatric residents were able to moderate their overendorsement of MMPI-2 items when asked to feign schizophrenia, although they were unable to accurately replicate the profiles of clinical patients. In

Table VI. Classification Rates for Cut Scores on MMPI-2 Fake-Bad Validity Scales for PTSD v. Faker Comparison

		Classification rate				
Scale	Cut score	SENS	SPEC	PPP	NPP	OCC
Cut scores reported in Bury & Bagby (2002)						
F	>23	.21	.96	.86	.49	.54
F_B	>16	.41	.91	.86	.55	.63
F_P	>8	.10	1.00	1.00	.47	.50
Ds_2	>35	.34	.96	.91	.54	.62
Cut scores based on present sample						
F	>11	.71	.70	.74	.67	.71
F_B	>7	.79	.65	.74	.71	.73
F_P	>4	.52	.87	.83	.59	.67
Ds_2	>25	.66	.74	.76	.63	.69

Note. Base rate of malingering = 56%; SENS: sensitivity, SPEC: specificity, PPP: positive predictive power; NPP: negative predictive power; OCC: overall correct classification; F: infrequency scale; F_B: F-back scale; F_P: infrequency-psychopathology scale; Ds_2: Gough dissimulation scale.

another study, in which PTSD symptom information was provided, Bury and Bagby (2002) reported that providing specific symptom information did not improve participants' ability to successfully feign PTSD on the MMPI-2 over that of participants provided with validity scale information or both.

This study partially replicated results reported by Bagby et al. (2002), in which the NIM and MAL scales were not significant predictors of feigning, but the RDF index performed as well as the F family of scales on the MMPI-2. In the present study, the RDF scale of the PAI did not appear as a significant predictor in any regression model tested. However, the largest (though moderate) effect size on the PAI for the PTSD v. Faker comparison was demonstrated by the RDF index. Therefore, it is possible that the results reported by Bagby et al. (2002) might have been replicated with a larger sample. Nevertheless, in the current investigation, none of the PAI fake-bad validity scales significantly predicted feigned PTSD. The results of this study, taken together with the results of previous studies in which NIM and MAL were ineffective predictors, suggest that the PAI validity scales may not be operating as they are intended for detecting dissimulation by coached feigners, particularly for neurotic disorders such as PTSD.

One possible explanation for the better performance of the MMPI-2 relative to the PAI in the detection of feigned PTSD has to do with construction of the tests. Because the MMPI-2 was developed using an empirical criterion-keyed method, many items are not face-valid, making it a daunting task to identify items as belonging to certain scales and strategically endorse items to create a particular profile, even for trained individuals (Mehlman & Rand, 1960; Bagby et al., 1997). In contrast, because the PAI was developed using a construct validation approach, many of its items possess substantial face validity, which could enhance the ability of coached feigners to recognize and endorse items that are associated with the disorder being feigned. Evidence for this can be inferred from the success of feigners in this study at replicating the PAI clinical scale profiles of the PTSD group. The only clinical scale that Fakers scored significantly higher on than PTSDs was ARD-T, lending further support for the interaction effect of symptom training and face-validity of test items.

Another feature of the MMPI-2 that may make it more difficult for feigners to avoid detection is its True-False response format, as opposed to the four-option item rating scale of the PAI. Because item responses are dichotomous on the MMPI-2, the only way to moderate responding is to endorse only a proportion of the relevant items. However, this requires that respondents (a) recognize what scale items load on, (b) know how many such items are presented across the entire test, and (c) know what proportion of items should be endorsed in order to be seen as having PTSD while avoiding detection by validity scales. In contrast, on the PAI respondents could moderate responding by adopting the relatively straightforward heuristic of simply avoiding extreme ratings on individual PAI items, i.e., avoiding item ratings of 0 and 3 and relying primarily on ratings of 1 and 2, while selectively endorsing items that are clearly consistent with the symptom information provided.

Two other findings are of note. First, for both the MMPI-2 and PAI, combining scales into blocks resulted in little incremental predictive capacity over that found for the most discriminating individual scales. This reflects the impact of multicollinearity among the validity scales and among the PTSD-relevant clinical scales on both tests. The only substantial incremental improvement was found for the MMPI-2 clinical scales in the Faker v. PTSD comparison, where the combined scales showed an increase both in R^2 and overall correct classification. The combined PAI clinical scales also showed an increase in R^2 for the Faker v. PTSD comparison, but not in correct classification. Second, for the both the MMPI-2 and PAI, the combined PTSD-relevant clinical scales yielded better discrimination of Fakers and PTSDs than did the combined fake-bad validity scales. Although this finding needs to be replicated, it suggests that clinical scales may prove to be a valuable supplement to validity scales in the detection of malingering.

Limitations of this study include the relatively small sample size. As a cautionary note, regression models that involved block entry of four predictors were conducted as exploratory analyses, with the recognition that overfitting of the data was a likely problem due to the small sample. As with all simulation malingering research designs, an additional limitation is whether the performance of trained student malingerers is truly comparable to that of highly motivated real-world malingerers. No formal assessment of Fakers' motivation or effort to simulate PTSD was conducted. However, during debriefing Fakers often spoke enthusiastically about their engagement in the task, describing various strategies for "beating the test" and winning the cash prize. Further, only two Fakers were excluded due to inconsistent or irrelevant responding on the tests, which suggests that they attempted to adhere closely to the instructions.

Finally, it should be noted that although the PTSD sample in this study met DSM-IV criteria for PTSD, it was not a treatment-seeking sample. Therefore, the severity of symptoms and corresponding scale elevations were probably lower than might be observed in a clinical

setting. However, this suggests that the coached feigners in this study might be even more difficult to detect were they compared to a clinical sample. For example, with respect to the MMPI-2 validity scales, although PTSDs showed mean elevations ($T > 70$) on all fake-bad validity scales except Fp, Fakers scored even higher, and thus at least moderate levels of discrimination were found. However, clinical PTSD samples often demonstrate substantially higher elevations on fake-bad validity scales (e.g., Elhai et al., 2001; Frueh, Hamner, Cahill, Gold, & Hamlin, 2000), making it substantially more difficult to discriminate genuine and malingered PTSD on the basis of these scales.

Future research should be directed toward replication and extension of this and other similar studies into the relative vulnerability of these tests to feigning of specific disorders by coached, motivated individuals. There is some conflicting evidence regarding the effectiveness of the PAI scales for detecting motivated feigning by individuals who are armed with symptom information and some understanding of validity scale construction. More information is needed to determine if these scales are particularly vulnerable to feigning in such instances.

REFERENCES

American Psychiatric Association. (1980). Diagnostic and statistical manual of mental disorders (3rd edn.). Washington, DC: Author.

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

Bagby, R. M., Nicholson, R. A., Bacchiochi, J. R., Ryder, A. G., & Bury, A. S. (2002). The predictive capacity of the MMPI-2 and PAI validity scales and indexes to detect coached and uncoached feigning. *Journal of Personality Assessment*, *78*, 69–86.

Bagby, R. M., Rogers, R., Nicholson, R., Buis, T., Seeman, M. V., & Rector, N. (1997). Does clinical training facilitate feigning schizophrenia on the MMPI-2? *Psychological Assessment*, *9*, 106–112.

Baldrachi, R., Hilsenroth, M., Arsenault, L., Sloan, P., & Walter, C. (1998). MMPI-2 assessment of varying levels of posttraumatic stress in Vietnam combat veterans. *Journal of Psychopathology and Behavioral Assessment*, *21*, 109–116.

Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Gusman, F. D., Charney, D. S., & Keane, T. M. (1995). The development of a clinician-administered PTSD scale. *Journal of Traumatic Stress*, *8*, 75–90.

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., Tellegen, A. M., Dahlstrom, W. G., & Kaemmer, B. (2001). MMPI-2 Minnesota Multiphasic Personality Inventory-2: Manual for administration, scoring, and interpretation (Rev. edn.). Minneapolis: University of Minnesota Press.

Calhoun, P. S., Earnst, K. S., Tucker, D. D., Kirby, A. C., & Beckham, J. C. (2000). Feigning combat-related posttraumatic stress disorder on the Personality Assessment Inventory. *Journal of Personality Assessment*, *75*, 338–350.

Doepel, D. (Producer/director) (1989). Understanding psychological trauma (Video). Available from Baxley Media Group at www.baxleymedia.com..

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.

Fairbank, J. A., McCaffrey, R. J., & Keane, T. M. (1985). Psychometric detection of fabricated symptoms of posttraumatic stress disorder. *American Journal of Psychiatry*, *142*, 501–503.

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.

Groth-Marnat, G. (2003). *Handbook of psychological assessment* (4th edn.). New York, NY: John Wiley & Sons.

Guriel, J., & Fremouw, W. (2003). Assessing malingered posttraumatic stress disorder: A critical review. *Clinical Psychology Review*, *23*, 881–904.

Hathaway, S. R., & McKinley, J. C. (1943). *The Minnesota Multiphasic Personality Inventory manual*. Minneapolis: University of Minnesota Press.

Hathaway, S. R., & McKinley, J. C. (1951). *An atlas for the clinical use of the MMPI*. Minneapolis: University of Minnesota Press.

Hickling, E., Taylor, A., Blanchard, E., & Devineni, T. (1999). Simulation of motor-vehicle accident-related PTSD: Effects of coaching with DSM-IV criteria. In E. J. Hickling & E. B. Blanchard (Eds.), The international handbook of road traffic accidents and psychological trauma: Current understanding, treatment, and law (pp. 304–320). Amsterdam: Elsevier Press.

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

Kraemer, H. C. (1992). Evaluating medical tests: Objective and quantitative guidelines. Newbury Park, CA: Sage.

Lamb, D., Berry, D. T. R., Wetter, M., & Baer, R. (1994). Effects of two types of information on malingering of closed-head injury on the MMPI-2: An analog investigation. *Psychological Assessment*, *6*, 8–13.

Liljequist, L., Kinder, B. N., & Schinka, J. A. (1998). An investigation of malingering posttraumatic stress disorder on the Personality Assessment Inventory. *Journal of Personality Assessment*, *71*, 322–336.

McCaffrey, R. J., & Bellamy-Campbell, R. (1989). Psychometric detection of fabricated symptoms of combat-related post-traumatic stress disorder: A systematic replication. *Journal of Clinical Psychology*, *45*, 76–79.

Mehlman, B., & Rand, M. E. (1960). Face validity of the MMPI. *Journal of General Psychology*, *63*, 171–178.

Morey, L. (1991). Personality Assessment Inventory: Professional manual. Odessa, FL: Psychological Assessment Resources.

Morey, L. (1996). An interpretive guide to the Personality Assessment Inventory. Odessa, FL: Psychological Assessment Resources.

Pope, H. S., Butcher, J. N., & Seleen, J. (2000). *The MMPI, MMPI-2 & MMPI-A in court*. Washington, DC: American Psychological Association.

Rogers, R. (Ed.). (1997). Clinical assessment of malingering and deception (2nd edn.). New York: Guilford.

Rogers, R., Bagby, R. M., & Chakraborty, D. (1993a). Feigning schizophrenic disorders on the MMPI-2: Detection of coached simulators. *Journal of Personality Assessment*, *60*, 215–226.

Rogers, R., Ornduff, S. R., & Sewell, K. W. (1993b). Feigning specific disorders: A study of the Personality Assessment Inventory (PAI). *Journal of Personality Assessment*, *60*, 554–560.

Rogers, R., Sewell, K. W., Cruise, K. R., Wang, E. W., & Ustad, K. L. (1998). The PAI and feigning: A cautionary note on its use in forensic-correctional settings. *Assessment*, *5*, 399–405.

Rogers, R., Sewell, K. W., Martin, M. A., & Vitacco, M. J. (2003). Detection of feigned mental disorders. *Assessment*, *10*, 160–177.

Rogers, R., Sewell, K. W., Morey, L., & Ustad, K. L. (1996). Detection of feigned mental disorders on the Personality Assessment Inventory:

A discriminant analysis. *Journal of Personality Assessment*, *67*, 629–640.

Rogers, R., Sewell, K. W., & Salekin, R. T. (1994). A meta-analysis of malingering on the MMPI-2. *Assessment*, *1*, 227–237.

Rosenthal, R., Rosnow, R. L., & Rubin, D. B. (2000). Contrasts and effect sizes in behavioral research: A correlational approach. New York: Cambridge University Press.

Scheibe, S., Bagby, R. M., Miller, L. S., & Dorian, B. J. (2001). Assessing posttraumatic stress disorder with the MMPI-2 in a sample of workplace accident victims. *Psychological Assessment*, *13*, 369–374.

Schretlen, D. J. (1988). The use of psychological tests to identify malingered symptoms of mental disorder. *Clinical Psychology Review*, *8*, 451–476.

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

Walters, G. L., & Clopton, J. R. (2000). Effect of symptom information and validity scale information on the malingering of depression on the MMPI-2. *Journal of Personality Assessment*, *75*, 183–199.

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.

Weathers, F. W., Ruscio, A. M., & Keane, T. M. (1999). Psychometric properties of nine scoring rules for the Clinician-Administered Posttraumatic Stress Disorder Scale. *Psychological Assessment*, *11*, 124–133.

Wetter, M. W., & Corrigan, S. K. (1995). Providing information to clients about psychological tests: A survey of attorneys' and law students' attitudes. *Professional Psychology: Research and Practice*, *26*, 474–477.