JOURNAL OF PERSONALITY ASSESSMENT, *80*(2), 197–205

A Comparison of the PAI and MMPI–2 As Predictors of Faking Bad in College Students

Dorothy D. Blanchard and Robert E. McGrath
School of Psychology
Fairleigh Dickinson University

David L. Pogge
School of Psychology
Fairleigh Dickinson University and
Four Winds Hospital
Katonah, New York

Ali Khadivi
Bronx Lebanon Hospital, New York and
Department of Psychiatry
Albert Einstein Medical College

Both the Minnesota Multiphasic Personality Inventory–2 (MMPI–2; Butcher et al., 2001) and Personality Assessment Inventory (PAI; Morey, 1991) offer a large set of scales devoted to the identification of response styles. This study directly compared the effectiveness of the 2 inventories as indicators of overreporting. The 2 measures were administered to 52 college students instructed to fake bad under conditions describing either a forensic ($n = 24$) or psychiatric ($n = 28$) setting as well as to 432 psychiatric patients. Results indicated that the MMPI–2 F – K index and Fp Scale were the best single indicators of a faking bad response style and that the MMPI–2 scales were the better indicators as a set. However, the PAI scales demonstrated a significant level of incremental validity over the MMPI–2 indicators in every analysis conducted. The findings suggest that either inventory offers an effective approach to the detection of overreporting, and administering both inventories can enhance the accuracy of prediction further.

The ability to detect potentially invalid protocols is critical in clinical assessment, particularly when self-report measures are used. The detection of overreporting or the exaggeration of pathology is particularly important in settings where malingering can benefit the respondent, as is often the case in forensic evaluations (Rogers, Sewell, & Goldstein, 1994). Overreporting can occur for other reasons as well, as in instances of factitious disorder or distorted self-perceptions or as a means of drawing attention from psychiatric staff.

The authors of the original Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1942; Meehl & Hathaway, 1946) were well aware of this issue, and introduced a series of response style indicators early in the process of developing the inventory. Subsequent research has both introduced additional response style indicators and demonstrated their validity. As a consequence, the revised version of the MMPI (MMPI–2; Butcher et al., 2001) offers the largest, best validated set of response style indicators of any measure currently available. For example, two meta-analyses have reported mean *d* values above 1.0 for the MMPI indicators of overreporting (Berry, Baer, & Harris, 1991; Rogers, Sewell, & Salekin, 1994). Although most of the studies reviewed were analog studies, naturalistic studies with suspected malingerers who completed the MMPI under standard instructions on average also demonstrated large effect sizes.

The Personality Assessment Inventory (PAI; Morey, 1991) represents an alternative self-report measure of psychopathology with several advantages over the MMPI. The item pool is smaller, and items on average require a lower reading level. There is no item overlap between scales and between subscales, and scales were developed to be consistent with current diagnostic nosology. The PAI has its own set of response style indicators including four that have been specifically suggested for the identification of overreporting.

Although the literature demonstrating the validity of the PAI as an indicator of overreporting is much smaller than that for the MMPI, results generally support the use of the inventory for this purpose (Calhoun, Earnst, Tucker, Kirby, & Beckham, 2000; Liljequist, Kinder, & Schinka, 1998; Morey & Lanier, 1998; Rogers, Ornduff, & Sewell, 1993; Rogers, Sewell, Cruise, Wang, & Ustad, 1998; Rogers Sewell, Morey, & Ustad, 1996; Wang et al., 1997).

To date, only one study has directly compared the validity of the PAI and MMPI as predictors of overreporting. Bagby, Nicholson, Bacchiochi, Ryder, and Bury (2002) compared college students responding honestly to the two inventories, college students instructed to feign a mental disorder in the context of a disability claim who were coached on methods of escaping detection by the response style scales, college students instructed to feign without coaching, and psychiatric patients. They found the Rogers Discriminant Function (RDF; Rogers Sewell, Morey, & Ustad, 1996) was the best single indicator of overreporting, and slightly better than the MMPI scales evaluated in the study.

This study is similar in purpose to Bagby et al.'s (2002) study with several important differences. First, whereas Bagby et al. restricted their analysis to a subset of the MMPI overreporting scales, we considered all of the MMPI scales that have been recommended for the identification of overreporting. Second, whereas Bagby et al. compared fakers on whether they were coached on the response style scales, in this study they were compared on whether they were attempting to feign a mental disorder in a psychiatric or forensic setting. Finally, we included a much larger psychiatric comparison group than Bagby et al. Doing so not only should enhance the reliability of the findings; by combining this sample with much smaller samples of individuals instructed to feign, it also resulted in what is probably a more realistic representation of the base rate for overreporting in clinical settings.

The study was conducted with several questions in mind. Do the indicators of overreporting from the two inventories measure similar variables? Which indicator is the single best measure of overreporting? Is one inventory better than the other as an indicator of overreporting? Does each inventory offer incremental validity over the other as a measure of overreporting?

METHOD

Participants

The study involved two samples of participants. The student sample consisted of 52 undergraduates at Fairleigh Dickinson University, a midsize private university located in a suburb of New York City. Students were informed they would be paid $25 for their participation in the study should they meet eligibility requirements and complete the assessment. In addition, students were informed that 8 participants would earn additional incentives, with 2 earning as much as $100. Students were randomly assigned to a forensic or psychiatric feigning condition to be described later.

Students were eliminated from the analyses for one of three reasons. Before administering the questionnaires students were asked to read three fairly complex MMPI–2 items (Items 110, 262, and 442) and explain their meaning in writing. Participation was terminated if Dorothy Blanchard was concerned about fluency in English based on their responses. This criterion eliminated 1 participant. Second, to reduce the likelihood of actual pathology in the sample students were asked as part of a standard demographic form whether they were currently receiving psychotherapy or psychotropic medication. None were excluded on the basis of this criterion. Finally, to guard against random or careless responding, 7 students with scores above 80*T* on the MMPI–2 Variable Response Inconsistency scale (VRIN) and/or 73*T* on the PAI Inconsistency scale (ICN) were excluded from the sample. Cut scores were based on standard references concerning the optimal cut scores for these scales (Berry, 1995; Butcher et al., 2001; Morey, 1991). The remaining 52 students included 24 in the forensic condition and 28 in the psychiatric condition.

The second sample consisted of archival records for 432 inpatients from Four Winds Hospital, a private psychiatric facility located in a suburb of New York City, who had completed both the MMPI–2 and PAI during their hospitalization. An additional 25 patients with VRIN > 80*T* and/or ICN > 73*T*, or who had omitted more than 29 items, were excluded from the sample. It is unclear whether the higher rate of exclusion for random or careless responding in the student group is an artifact of their attempt to fake bad or stricter prescreening of patients for the capacity to complete the inventory.

Table 1 provides demographic data for the three groups. The gender composition differed significantly across groups, $\chi^2(2, N = 484) = 9.9$, $p < .01$. Subsequent chi-square tests using pairs of groups indicated the proportion of men was significantly lower in the students instructed to fake a psychiatric condition than among the patients, $\phi = .15$. There was no significant difference between the forensic fakers and patients, $\phi = .03$, or between the two student groups, $\phi = .26$.

An analysis of variance (ANOVA) also indicated a significant difference across groups for age, $F(2, 462) = 17.8$, $p < .01$. Subsequent t tests demonstrated the mean age was significantly lower in the forensic feigners than in patients, $d = 1.0$, and significantly lower in the psychiatric feigners than the patients, $d = 0.8$, but the difference between feigner groups was not significant, $d = 0.2$. Despite the use of college students as feigners, mean years of education did not differ significantly across the groups.

Measures

MMPI–2. The study focused on eight MMPI–2 scales and scale combinations that have been recommended for the

TABLE 1
Demographic Data for the Sample

Variable	Student Feigners		Patients
	Forensic	Psychiatric	
N	24	28	432
Gender			
% male	$41.7_{a,b}$	17.9_{a}	48.1_{b}
% female	58.3	82.1	51.9
Age			
M	22.3_{a}	24.3_{a}	34.4_{b}
SD	7.8	12.0	12.2
Education			
M	14.2	13.8	14.2
SD	1.8	1.1	2.9
Ethnicity			
% White	58.3	57.1	
% Black	25.0	14.3	
% Hispanic	8.3	10.7	
% other/unspecified	8.3	17.9	

Note. Although ethnicity data were not available for the patient sample, hospital census statistics indicate the general population is approximately 62% White, 18% Black, 14% Hispanic and 6% other/unspecified. Different subscripts in a row indicate a significant difference ($p < .01$).

identification of overreporting in previous studies. Three of these scales are currently included in the standard battery of MMPI response style indicators (Butcher et al., 2001). The Infrequency (F) scale consists of items rarely endorsed in the keyed direction by members of the MMPI's original normative sample. The Back F (Fb) scale consists of items not included in the F scale that were rarely endorsed in the keyed direction in the MMPI–2 normative sample. The Fb items are included in the second half of the MMPI–2 item set so that the Fb scale serves as an F scale for the later items in the inventory. The Infrequency-Psychopathology scale (Fp; Arbisi & Ben-Porath, 1995) is the most recent addition to the standard set. It is similar in format to F and Fb but is comprised of items rarely endorsed in the keyed direction by patients as well as by normals. These scales are sensitive to malingering because many of the items represent particularly rare symptoms of psychopathology.

There are five other approaches suggested for the detection of overreporting that are not included in the standard battery of response style indicators. The Gough Dissimulation Scale (Ds; Gough, 1954) was designed to identify individuals attempting to fake neuroticism. A similar scale, called the Ds–Revised (Ds–R; Gough, 1957), was created in conjunction with the development of the California Psychological Inventory (Gough, 1957). The MMPI–2 versions of these scales consist of those items retained from the original MMPI scales. The Fake Bad Scale (FBS; Lees-Haley, English, & Glenn, 1991) was developed to detect malingering in the context of personal injury evaluations. These scales were developed by selecting items that were significant predictors of putative malingering status.

Two indexes were also computed. The F minus K index (F – K) is computed by subtracting the raw score on the Correction (K) scale, a measure of underreporting, from the raw score on F (Gough, 1947). Finally, Weiner and Harmon developed the Obvious–Subtle index (O–S) based on the hypothesis that endorsing more obvious items predictive of psychopathology than subtle items would be indicative of overreporting (Weiner, 1948). The index is computed by summing the difference between the *T* scores for the obvious and subtle subscales from each of five clinical scales.

PAI. Four scales or scale combinations have been recommended to date for the assessment of overreporting with the PAI. The Negative Impression Management scale (NIM) was included with the original PAI set of scales (Morey, 1991) and consists of items expected to represent unusually severe elements of psychopathology. Morey also identified eight configurations of scales that tended to be observed more frequently in the profiles of respondents feigning psychopathology than in bona fide psychiatric patients. The Malingering Index (MAL; Morey 1991) is the number of those configural indicators present in a profile. The RDF (Rogers et al., 1996) represents the linear combination of PAI scales that best discriminated between overreporters and respondents instructed to answer honestly. Finally, the Cashel Discriminant Function (CDF; Cashel, Rogers, Sewell, & Martin-Cannici, 1995) was developed to identify faking good, but Morey and Lanier (1998) found it sensitive to both faking good and faking bad. It is interesting to note that NIM is not included in either discriminant function.

Procedure

Student sample. Students completed both the MMPI–2 and PAI with instructions to overreport. The instructions were developed in light of recent calls for the use of more realistic faking conditions in response style research (e.g., Rogers, 1998; Schretlen, 1988). First, students were offered a context for their faking. Overreporting has been identified as an issue in both forensic and psychiatric settings; therefore, two scenarios were developed to which students were randomly assigned. Those in the forensic condition (n = 24) were told they were trying to convince a jury to find them "Not Guilty by Reason of Insanity" for commission of a felony. Those in the psychiatric condition (n = 28) were told they were trying to convince a doctor they should be admitted to a psychiatric hospital.

Second, students were informed of the presence of scales intended to detect inaccurate responding, although they were not given any specific information about the manner in which the validity scales accomplished this goal. The instructions stated, "Both the MMPI and PAI have ways of telling if the person is faking the mental illness. Your goal in responding to these questionnaires is to convince a doctor that you have a mental illness *without being detected as faking this* condition."

Third, students were also provided detailed information about the symptoms most likely to achieve their goals based on descriptions found in the *Diagnostic and Statistical Manual for Mental Disorders* (4th ed.; American Psychiatric Association, 1994). We focused the coaching on symptoms rather than specific diagnoses because we believed malingerers are more likely to mimic symptoms that will lead to the desired response from professionals than to attempt to duplicate the diagnostic criteria for a particular disorder. In the forensic condition, students were provided with information about five symptoms indicative of severe psychosis (delusions, hallucinations, loss of interest, emotional numbing, and bizarre behavior). In the psychiatric condition, students were in addition given information about seven signs of major depression (depressed mood, loss of interest, sleep problems, fatigue, worthlessness/guilt, lack of concentration, and suicidal ideation).

Third, students were offered financial incentives for participation and for successful overreporting. Specifically, they were informed that the top four simulators in each condition would receive a reward of $100, $75, $50, or $25, for obtaining elevations on clinical scales while eluding detection by the response style indicators. Finally, the instructions asked the students to think about their roles before beginning the questionnaires.[1]

Students completed the PAI and MMPI–2 in counterbalanced order. Prizewinners were determined by subtracting each participant's *T* score on Scale F from the mean of their *T* scores on Scales 6 and 8 and ranked from highest to lowest (Wetter, Baer, Berry, Robinson, & Sumpter, 1993).

Inpatient sample. All adult inpatients at Four Winds Hospital capable of doing so are administered the MMPI–2 on admission. Patients referred for psychological testing also complete the PAI so that in all cases the PAI was completed after the MMPI–2. Both were completed under standard instructions without financial incentives.

RESULTS

Table 2 provides the mean scores on basic scales for each group of feigners and patients. Because these are provided primarily for descriptive purposes, significance tests and effect sizes were not computed. In almost every case the mean score for the psychiatric feigners exceeded the mean score among patients, with the mean for forensic feigners falling in between. The only exception involved measures of underreporting in which patient scores were higher than those for feigners and measures of manic tendencies in which forensic feigners produced the highest scores. In every case except for several validity scales not relevant to this study, feigners' mean scores were higher than those for patients.

Do the Two Inventories Measure Similar Variables?

Table 3 presents results relevant to the degree of overlap among the 12 measures of overreporting. Only the patient sample was used for these analyses. In general correlations were quite high even across inventories. Except for those involving RDF, CDF, or FBS, correlations were consistently .40 or greater. All correlations were significant ($p < .01$) except those between FBS and the two PAI discriminant functions.

A principal components analysis was then conducted using the 12 scales. Without rotation a substantial first component emerged. The eigenvalue was seven times that of the second component, and the first component accounted for

TABLE 2
Means and Standard Deviations for Basic MMPI and PAI Scales

	Forensic Feigners		*Psychiatric Feigners*		*Patients*	
Scale	*M*	*SD*	*M*	*SD*	*M*	*SD*
MMPI						
L	51.8	10.1	51.8	9.4	51.2	9.7
K	36.8	5.2	37.1	5.8	46.2	10.1
1	77.5	20.2	87.0	12.7	63.6	13.5
2	74.0	17.2	92.8	10.6	73.9	15.4
3	68.5	19.1	79.7	13.2	66.6	15.3
4	81.1	16.8	92.3	7.6	71.6	14.0
5	59.5	12.2	62.0	10.6	51.7	10.0
6	105.2	19.7	106.9	11.9	70.9	16.7
7	82.6	15.7	92.0	7.0	72.5	15.6
8	109.5	17.9	113.9	9.1	72.8	17.4
9	79.3	12.2	73.1	14.1	57.6	13.3
0	69.5	13.5	78.7	7.4	57.5	12.6
PAI						
SOM	80.2	21.4	88.9	15.0	60.6	13.1
ANX	77.4	16.5	88.3	9.3	65.8	14.7
ARD	77.4	14.9	84.8	14.4	62.8	15.0
DEP	81.8	19.1	98.5	12.4	71.8	16.9
MAN	69.3	14.3	55.4	13.9	54.2	12.2
PAR	93.8	15.7	94.1	13.8	58.5	13.4
SCZ	97.6	21.3	102.8	15.0	60.6	15.1
BOR	76.5	10.7	85.4	9.4	68.9	14.2
ANT	80.0	15.9	78.0	19.2	57.8	13.2
AGG	73.0	14.4	66.8	18.4	56.0	14.3
ICN	53.9	9.8	49.8	10.2	53.8	8.3
INF	68.9	14.1	74.5	18.7	53.4	9.4
PIM	39.3	13.3	32.9	7.6	40.2	11.7

Note. MMPI = Minnesota Multiphasic Personality Inventory; PAI = Personality Assessment Inventory; L = Lie; K = Correction; 1 = Hypochondriasis; 2 = Depression; 3 = Hysteria; 4 = Psychopathic Deviate; 5 = Masculinity–Feminity; 6 = Paranoia; 7 = Psychasthenia; 8 = Schizophrenia; 9 = Hypomania; 0 = Social Introversion; SOM = Somatization; ANX = Anxiety; ARD = Anxiety Related Disorders; DEP = Depression; MAN = Mania; PAR = Paranoia; SCZ = Schizophrenia; BOR = Borderline Features; ANT = Antisocial Attitides; AGG = Aggression; ICN = Inconsistency; INF = Infrequency; PIM = Positive Information Management.

[1]The complete set of instructions for each condition is available from Dorothy D. Blanchard.

TABLE 3
Pairwise Correlations for Overreporting Indicators

Scale	*F*	*Fb*	*Ds*	*Ds–R*	*F(p)*	*FBS*	*O–S*	*F – K*	*NIM*	*RDF*	*MAL*	*CDF*
MMPI												
Fb	0.80											
Ds	0.84	0.84										
Ds–R	0.81	0.82	0.96									
Fp	0.73	0.60	0.62	0.57								
FBS	0.48	0.51	0.52	0.55	0.27							
O–S	0.78	0.80	0.88	0.86	0.53	0.51						
F – K	0.91	0.81	0.88	0.85	0.66	0.42	0.89					
PAI												
NIM	0.72	0.73	0.69	0.66	0.53	0.42	0.72	0.74				
RDF	0.29	0.27	0.31	0.32	0.17	0.06	0.28	0.30	0.18			
MAL	0.52	0.50	0.47	0.44	0.45	0.25	0.48	0.53	0.63	0.14		
CDF	0.27	0.21	0.29	0.28	0.25	0.11	0.26	0.32	0.23	0.27	0.21	
Loadings	0.92	0.90	0.94	0.92	0.71	0.56	0.91	0.94	0.82	0.34	0.61	0.23

Note. These statistics are based on 432 psychiatric inpatients. All correlations are significant ($p < .01$) except those between FBS and the two PAI discriminant functions. Loadings are from the first unrotated factor generated via principal components analysis. The first component was substantial: eigenvalue = 7.1; percentage of total variance = 59.5. The corresponding values for the second component were 1.2 and 10.0, respectively. F = Infrequency scale; Fb = F Back scale; Ds = Dissimulation scale; Ds–R = Dissimulation Scale Revised; F(p) = Infrequency Psychopathology scale; FBS = Fake Bad scale; O–S = Obvious–Subtle Index; F – K = F Minus K Index; NIM = Negative Impression Management; RDF = Rogers Discriminant Function; MAL = Malingering Index; CDF = Cashel's Discriminant Function; MMPI = Minnesota Multiphasic Personality Inventory.

60% of total variance. The last row in Table 3 provides loadings for the first component on each scale. All loadings exceeded .50 with the exception of those for RDF and CDF.[2]

These findings suggest that despite the use of alternate approaches to the detection of malingering, 10 of the 12 scales are strongly influenced by the same latent variable. Only RDF and CDF seem to be relatively independent of this common source trait. Furthermore, RDF and CDF are not strongly correlated with each other, suggesting they each reflect different constructs than the other scales.

Which Is the Single Best Measure of Overreporting?

Table 4 provides information relevant to this question. For each of the 12 scales the omnibus ANOVA test of group differences was significant ($p < .01$). Subsequent t tests indicated that in every case except the CDF the patients were significantly different in the expected direction from both faking bad groups. For the CDF the students in the forensic condition generated significantly higher scores than the other two groups. The only other finding was that students in the psychiatric condition generated higher mean scores on FBS than students in the forensic condition. Except for CDF, the two scenarios did not result in substantially different responses to the overreporting measures.

Effect sizes were computed combining the two feigning conditions and for each feigning condition separately when compared with the patient group. In all three cases, the best single predictor was the F – K index, followed closely by Fp. With one exception involving the psychiatric feigners, the three PAI indicators excluding CDF represented the next best predictors of group membership, with MAL matching the effect size for Fp among the forensic feigners. Values for d were consistently large as is typical in analog studies of response style indicators (Berry et al., 1991; Rogers, Sewell, & Salekin, 1994). Although the effect size associated with the use of the RDF to predict forensic feigners was large, a previous naturalistic study (Rogers et al., 1998) suggested that the RDF may not be appropriate for use in forensic settings.[3]

Diagnostic efficiency statistics provide an alternative approach to identifying the best predictors. These are provided in Table 5. For each of the 12 predictors, a cut score was identified that maximized the hit rate between the patients and the combined group of feigners. It is worth noting the influence of the patient group on these findings. Given the large number of patients, the specificity and negative predictive power associated with a scale had a greater influence on the hit rate than its sensitivity and positive predictive power. This influence is to be expected given that the simulators represented

[2]Given the differences in outcomes possible from using alternate factor analytic strategies, this analysis was repeated using iterated principal axis factoring. The proportion of variance accounted for by the first factor, the disparity in the size of the eigenvalues, and the loadings all essentially replicated.

[3]Because feigners were instructed specifically in symptoms associated with depression and psychosis, these analyses were repeated with restricted samples of patients who met criteria for a depressive diagnosis (e.g., major depression, dysthymia) or a psychotic diagnosis (e.g., schizophrenia, major depression with psychotic features) at the time of discharge. Restricting the comparison in this way had little impact on the effect sizes.

TABLE 4
Descriptive Statistics for Overreporting Indicators

	Forensic Feigners		*Psychiatric Feigners*		*Patients*		*d*		
Scale	*M*	*SD*	*M*	*SD*	*M*	*SD*	*All Feigners*	*Forensic Feigners*	*Psychiatric Feigners*
MMPI–2									
F	115.0_a	16.5	119.1_a	3.4	72.1_b	20.7	2.24	2.07	2.35
Fb	112.7_a	20.2	119.6_a	1.7	77.9_b	24.8	1.62	1.41	1.71
Ds	41.0_a	10.0	43.7_a	7.8	19.9_b	9.5	2.42	2.19	2.54
Ds–R	22.8_a	4.6	25.5_a	4.5	12.1_b	6.2	2.01	1.78	2.22
Fp	114.7_a	16.9b	115.9_a	13.9	59.7_b	17.8	3.16	3.10	3.21
FBS	22.7_a	8.2	27.4_b	3.9	19.1_c	5.9	1.01	0.59	1.41
O–S	196.4_a	73.6	254.9_a	42.3	85.0_b	84.2	1.72	1.35	2.04
F – K	31.9_a	14.3	34.7_a	12.0	$–2.4_b$	10.0	3.44	3.35	3.67
PAI									
NIM	104.5_a	28.7	109.7_a	24.4	64.3_b	16.2	2.48	2.38	2.68
RDF	2.0_a	1.6	1.7_a	1.1	$–1.1_b$	1.2	2.48	2.65	2.41
MAL	4.3_a	1.9	3.8_a	2.1	1.0_b	1.0	2.61	3.10	2.54
CDF	1.9_a	1.9	0.5_b	1.4	$–0.1_b$	1.4	0.90	1.41	0.46

Note. Means with different subscripts in a row were significantly different ($p < .01$) based on *t* tests conducted after a significant analysis of variance. MMPI–2 = Minnesoata Multiphasic Personality Inventory–2; F = Infrequency scale; Fb = F Back scale; Ds = Dissimulation scale; Ds–R = Dissimulation Scale Revised; F(p) = Infrequency Psychopathology scale; FBS = Fake Bad scale; O–S = Obvious–Subtle Index; F – K = F Minus K Index; NIM = Negative Impression Management; RDF = Rogers Discriminant Function; MAL = Malingering Index; CDF = Cashel's Discriminant Function.

TABLE 5
Cut Score Analysis for Overreporting Indicators

Scale	*Cut Score*	*Hit Rate*	*Sensitivity*	*Specificity*	*PPP*	*NPP*
MMPI						
F[a]	120	.96	.90	0.97	.70	.99
Fb[a]	120	.87	.90	0.87	.45	.99
Ds	42	.96	.65	0.99	.92	.96
Ds–R	25	.93	.56	0.98	.76	.95
Fp[a]	120	.99	.89	1.00	.98	.99
FBS	34	.90	.04	1.00	.67	.90
O–S	244	.92	.48	0.98	.71	.94
F – K	23	.97	.83	0.99	.92	.98
PAI						
NIM[a]	110	.94	.54	0.99	.85	.95
RDF	1.80	.95	.60	0.99	.89	.95
MAL	5	.94	.48	1.00	.96	.94
CDF	2.73	.90	.23	0.98	.60	.91

Note. Feigning base rate = .11. Values ≥ the cut score were considered positive. PPP = positive predictive power; NPP = negative predictive power; F = Infrequency scale; Fb = F Back scale; Ds = Dissimulation scale; Ds–R = Dissimulation Scale Revised; F(p) = Infrequency Psychopathology scale; FBS = Fake Bad cale; O–S = Obvious–Subtle Index; F – K = F Minus K Index; NIM = Negative Impression Management; RDF = Rogers Discriminant Function; MAL = Malingering Index; CDF = Cashel's Discriminant Function.
[a]Cut score represents a *T* score.

only about 11% of the sample, a base rate that is consistent with estimates of faking in true clinical settings (Rogers, Sewell, & Goldstein, 1994).

Specificity and negative predictive power were quite good for all 12 scales. However, the sensitivity of FBS and CDF were very poor and those for O–S and MAL were less than desirable. In terms of positive predictive power and overall hit rate, Fp proved the best single choice, with F – K and Ds close behind. The results for MAL were also quite impressive. Taken together, results from these analyses suggest Fp and F – K as the best single predictors available from the MMPI, whereas the performance of MAL from the PAI is only slightly poorer.

Is One Inventory Better Than the Other? Does Each Offer Incremental Validity?

A series of hierarchical logistic regression analyses was conducted to address the overall effectiveness of each set of predictors as well as their incremental validity over the other set. Results are found in Table 6. ρ^2 values refer to McFadden's rho squared, an ordinal measure of effect size on a scale from 0 to 1 that tends to be more conservative than the traditional multiple R^2.

The first set of analyses, labeled "All feigners, all predictors," combined the two groups of feigners and compared all eight MMPI scales to all four PAI scales. ρ^2 was .79 for the eight MMPI scales alone, $\chi^2(8, N = 484) = 259.8$, $p < .01$. The addition of the PAI scales increased ρ^2 to .85 and represented a significant increment in fit, $\chi^2(4, N = 484) = 20.6$, $p < .01$. In contrast, ρ^2 for the four PAI scales alone was .66, $\chi^2(4, N = 484) = 216.9$, $p < .01$. The addition of the MMPI scales again increased ρ^2 to .85, representing a significant increment in fit, $\chi^2(8, N = 484) = 63.5$, $p < .01$.

The analyses were replicated for the forensic and psychiatric feigners separately. However, in those analyses that involved all 12 predictors the degree of collinearity among the variables precluded achieving the desired level of tolerance in the minimization of the log likelihood function (10^{-6}). In

TABLE 6
Results of Hierarchical Logistic Regression Analyses

				Change Analysis	
Comparison	ρ^2	χ^2	*df*	$\Delta\chi^2$	Δdf
All feigners, all predictors					
MMPI	.79	259.8*	8		
MMPI + PAI	.85	280.4*	12	20.6*	4
PAI	.66	216.9*	4		
PAI + MMPI	.85	280.4*	12	63.5*	8
Forensic feigners, all predictors[a]					
MMPI	.73	138.1*	5		
MMPI + PAI	.82	153.4*	9	15.3*	4
PAI	.64	121.2*	4		
PAI + MMPI	.82	153.4*	9	32.3*	5
Psychiatric feigners, all predictors[a]					
MMPI	.77	162.1*	5		
MMPI + PAI	.87	183.9*	9	21.7*	4
PAI	.70	147.4*	4		
PAI + MMPI	.87	183.9*	9	36.5*	5
All feigners, best stepwise predictors					
MMPI	.77	254.7*	3		
MMPI + PAI	.85	279.2*	6	24.4*	3
PAI	.66	216.9*	3		
PAI + MMPI	.85	279.2*	6	62.3*	3

Note. Best MMPI predictors include the F Back and Infrequency Psychopathology scales and the F – K Index. Best PAI predictors include Negative Impression Management, the Malingering Index, and Rogers Discriminant Function. In each case change analysis reflects the increment in fit associated with the addition of the second set of predictors. ρ^2 = McFadden's rho squared; MMPI = Minnesota Multiphasic Personality Inventory; PAI = Personality Assessment Inventory.

[a]Reduced for collinearity. The F Back scale and the Obvious–Subtle Index and F Minus K Indexes were excluded from these analyses to reduce multicollinearity.

*$p < .01$.

reviewing the correlations among the variables within the patient sample (the shared participants in these analyses), it was determined that four MMPI predictors (F, Fb, O–S, and F – K) demonstrated a high degree of overlap. Accordingly, Fb, O–S, and F – K were excluded from these analyses.

The findings within each subgroup were consistent with those for the feigners overall. In every case each inventory offered incremental validity over the other, and the MMPI was consistently associated with a larger ρ^2 than the PAI.[4]

In clinical settings it is unusual to consider all eight MMPI scales or all four PAI scales when attempting to identify response styles. To create a more realistic comparison, stepwise logistic regressions were conducted separately for the MMPI and PAI predictors. Because the results for the previous logistic analyses were essentially equivalent for the two feigning groups, they were combined for the stepwise analyses. The results for the MMPI suggest the most efficient combination of predictors consisted of the F – K index, Fp, and Fb. The results for the PAI suggested using MAL, NIM, and RDF. A final hierarchical regression was conducted comparing the best predictors from the MMPI and PAI. The results are provided at the bottom of Table 6. Again, each inventory offered incremental validity over the other, although the MMPI alone was associated with a larger effect size than the PAI alone.

[4]Analyses were also conducted evaluating the incremental validity of the two inventories over age because there was a significant difference in mean age between students and patients. Results indicated findings could not be accounted for by group differences in age.

DISCUSSION

Overall, the results suggest that if the choice of a single inventory is based on its effectiveness as an indicator of overreporting, the MMPI–2 is the better choice. The MMPI–2 set of indicators was consistently associated with larger effect sizes in regression analyses, and the best individual predictors of group status were the F – K index and Fp, although most of the other predictors were also associated with impressive effect sizes. Whereas effect size statistics were slightly higher for the F – K index than for Fp, the latter had the highest positive predictive power of any single indicator. The effectiveness of the F – K index is inconsistent with Butcher, Graham, and Ben-Porath's (1995) conclusion that the index has no advantage over Scale F alone but does mirror the results of prior meta-analyses.

Alhough the findings favor the MMPI, the PAI added incremental validity to the prediction of faking in every analysis conducted. The positive predictive power of MAL was particularly high. Finally, the MMPI FBS and the PAI CDF were particularly weak across all analyses. It is interesting to note that these scales were specifically developed for purposes other than the detection of overreporting in forensic or psychiatric settings (FBS was developed to detect overreporting in disability evaluations, whereas CDF was developed to detect underreporting).

Assuming these findings are reliable, the clinical implications are clear for any setting where the clinician is seriously concerned about the potential for faking self-report measures. Using either inventory by itself offers a valid approach to the detection of overreporting, although these findings suggest the MMPI as the better single choice. However, if the issue of faking is particularly important to the assessor, the administration of both the MMPI and PAI will provide a better indication than either inventory alone. In particular, the findings suggest the combined use of F – K, Fp, and Fb on the MMPI and NIM, RDF, and MAL on the PAI offers the most efficient and valid approach to the detection of feigning in the two situations examined.

There are several elements to the study's design that warrant mention for the limitations they place on the generalizability of the findings. Although students were notified of the existence of the response style scales, their coaching focused on describing the symptoms that were most

likely to help them achieve their goals. Previous research suggests coaching on avoiding detection by response style indicators can increase the likelihood of successful faking (Rogers, Bagby, & Chakraborty, 1993), although the effect is not always significant (e.g., Bagby et al., 2002). It is possible then that the results would have been different had we offered more training on the nature of the response style scales. Most likely this difference would have manifested itself in lower overall hit rates and perhaps in smaller differences in the effectiveness of the two inventories. Our decision in this study was based on an impression that in true clinical settings potential malingerers are more frequently familiar with the symptoms of severe pathology than they are with the response style scales used to detect faking. Even so, additional studies with fakers coached on the response style scales would clearly be worthwhile.

The second issue has to do with the analog nature of the study. Despite efforts to make the students' malingering more realistic, including offering incentives for successful faking, it is unclear how well these results mirror what would occur if a sample of true dissimulators was compared to a sample of patients known to be honest responders. As Rogers (1998) pointed out, however, in the absence of a gold standard for detecting malingerers analog research is usually the best alternative we have for evaluating the effectiveness of response style scales. The various elements of the methodology used to try to make the scenario more naturalistic, including the use of what may be a realistic base rate for faking, hopefully mitigate these concerns to some extent.

ACKNOWLEDGMENTS

Dorothy D. Blanchard is now at The City College of San Francisco. This research was supported in part by a grant from the Society for Personality Assessment. We thank the administration of Four Winds Hospital, Katonah, NY, for its support in the collection of data for this article. We are solely responsible for its content.

REFERENCES

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

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.

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.

Berry, D. T. R. (1995). Detecting distortion in forensic evaluations with the MMPI–2. In Y. S. Ben-Porath, J. R. Graham, G. C. N. Hall, R. D. Hirschman, & M. S. Zaragoza (Eds.), *Forensic applications of the MMPI–2* (pp. 82–102). Thousand Oaks CA: Sage.

Berry, D. T., Baer, R. A., & Harris, M. J. (1991). Detection of malingering on the MMPI: A meta-analysis. *Clinical Psychology Review, 11,* 585–598.

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: Manual for administration, scoring, and interpretation* (Rev. ed.). 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.

Cashel, M. L., Rogers, R., Sewell, K., & Martin-Cannici, C. (1995). The Personality Assessment Inventory (PAI) and the detection of defensiveness. *Assessment, 2,* 333–342.

Gough, H. (1947). Simulated patterns on the Minnesota Multiphasic Personality Inventory. *Journal of Abnormal and Social Psychology, 42,* 215–225.

Gough, H. (1954). Some common misconceptions about neuroticism. *Journal of Consulting Psychology, 18,* 287–292.

Gough, H. (1957). *California Psychological Inventory manual.* Palo Alto, CA: Consulting Psychologists Press.

Hathaway, S. R., & McKinley, J. C. (1942). *The Minnesota Multiphasic Personality Inventory.* Minneapolis: University of Minnesota Press.

Lees-Haley, P., English, L.T., & Glenn, W. (1991). A fake bad scale on the MMPI–2 for personal injury claimants. *Psychological Reports, 68,* 203–210.

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.

Meehl, P. E., & Hathaway, S. R. (1946). The K factor as a suppressor variable in the Minnesota Multiphasic Personality Inventory. *Journal of Applied Psychology, 30,* 525–564.

Morey, L. C. (1991). *The Personality Assessment Inventory professional manual.* Odessa, FL: Psychological Assessment Resources.

Morey, L. C., & Lanier, V. M. (1998). Operating characteristics of six response distortion indicators for the Personality Assessment Inventory. *Assessment, 5,* 203–214.

Rogers, R. (1998). Researching dissimulation. In R. Rogers (Ed.), *Clinical assessment of malingering and deception* (pp. 309–327). New York: Guilford.

Rogers, R., Bagby, R., & Chakraborty, D. (1993). 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. (1993). 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., & Goldstein, A. (1994). Explanatory models of malingering: A prototypical analysis. *Law and Human Behavior, 18,* 543–552.

Rogers, R., Sewell, K. T., Morey, L.C., & Ustad, K. (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.

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

Wang, E. W., Rogers, R., Giles, C. L., Diamond, P. M., Herrington-Wang, L. E., & Taylor, E. R. (1997). A pilot study of the Personality Assessment Inventory (PAI) in corrections: Assessment of malingering, suicide risk, and

aggression in male inmates. *Behavioral Sciences and the Law, 15,* 469–482.

Weiner, D. (1948). Subtle and obvious keys for the MMPI. *Journal of Consulting Psychology, 12,* 164–170.

Wetter, M. W., Baer, R. A., Berry, D. T., Robinson, 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.

Dorothy D. Blanchard
4507 Edgewood Avenue
Oakland, CA 94602
E-mail: ddblanch@earthlink.net

Received April 4, 2002
Revised September 29, 2002