JOURNAL OF PERSONALITY ASSESSMENT, 1998, 70(3), 551–563

A New Perspective on Gender Orientation Measurement With the MMPI–2: Development of the Masculine–Feminine Pathology Scale

Robert E. McGrath and Elizabeth Sapareto
School of Psychology
Fairleigh Dickinson University

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

A new scale of gender orientation for the MMPI–2 (Minnesota Multiphasic Personality Inventory–II; Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989) called the Masculine–Feminine Pathology Scale, or *Mfp*, was developed as an alternative to the available *Mf*, *GM*, and *GF* scales. It differs from previous scales in its emphasis on symptomatic correlates of gender. Items were included in the new scale if they (a) discriminated between male and female psychiatric patients and (b) were likely to be indicators of psychopathology. Statistical analyses suggested an acceptably reliable but factorially complex scale. When used to predict clinician ratings of global psychopathology, the scale demonstrated incremental validity over both the existing gender-related scales and the traditional clinical scales. Scores at the "feminine" end of the *Mfp* scale seem to reflect distress characterized by high levels of anxiety. Scores at the "masculine" end of the *Mfp* scale suggest a more composed interpersonal presentation, which may reflect an amoral attitude. It is suggested that the new scale may prove superior to the existing gender role scales as a supplement to other clinical scales. Avenues for future research with the *Mfp* scale are discussed.

Among the commonly scored scales of the Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1983), the Masculinity–Femininity (*Mf*) scale is probably the most controversial. Items were originally selected for

the scale because they were endorsed more frequently in the keyed direction by a sample of 13 gay men than by male soldiers. Items were then added that discriminated significantly between men and women. The original purpose of the scale seems to have been the identification of gay men and individuals demonstrating gender role reversal.

Subsequent research has provided little evidence that the scale is capable of achieving these goals. Wong (1984) summarized a series of studies indicating that the differences in mean elevation between gay and heterosexual men, and between men and women, are small. Other criticisms of the scale were noted as well. Extreme diversity in item content means that the interpretation of elevated scores is not straightforward. The markers of masculinity and femininity used for the scale are potentially anachronistic. Contrary to what is generally expected with MMPI scales, extreme scores at either end of the scale are not by themselves indicative of serious emotional difficulties (an issue we will discuss later). Finally, the scale treats masculinity and femininity as endpoints on a single bipolar dimension, rather than as two unipolar dimensions.

The last objection is particularly important in the context of subsequent literature on the nature of masculinity and femininity (e.g., Bem, 1974; Constantinople, 1973). More recent measures of gender roles such as the Bem Sex Role Inventory (BSRI; Bem, 1974) assume that masculinity and femininity are independent dimensions, and the four quadrants within this two-dimensional space represent various styles of gender role functioning: androgynous (high on both dimensions), masculine, feminine, or undifferentiated (low on both dimensions). The one-dimensional model of gender roles underlying *Mf* and similar scales is generally considered anachronistic.

Despite the heterogeneity in item content, studies examining the correlates of the *Mf* scale in a normal population have produced consistent findings. Women with elevated *T* scores on the scale (suggesting a "masculine orientation") are often described as ambitious and unemotional, whereas men with elevated *T* scores (suggesting a "feminine orientation") are seen as sensitive and passive (for a review, see Todd & Gynther, 1988). Clearly, despite obvious problems with the scale, there is reliable information to be gathered using *Mf*. In several cases, Todd and Gynther noted that the *Mf* scale contributes to the interpretation of other elements in the MMPI profile.

Recognizing that the evaluation of masculinity–femininity with the MMPI may be useful, but that the *Mf* scale is problematic, Peterson and Dahlstrom (1992) introduced two new MMPI gender role scales to address the inconsistency between the *Mf* scale and the more popular two-dimensional model of gender roles. Items were selected for the new scales using criteria established by Baucom (1976) in a similar study of the California Psychological Inventory. An item was included on the Gender Role–Masculine (*GM*) scale if it was answered either *true* or *false* by at least 70% of the men in the normative sample for the restandardized MMPI (MMPI–2;

Butcher, Dahlstrom, Graham, Tellegen, & Kaemmer, 1989), and the percentage of women from the normative sample endorsing the item in the same direction was at least 10% less. The same strategy using women as the primary target group was used to identify items for the Gender Role–Feminine (*GF*) scale. It was hypothesized that these scales could be used as a supplement to the *Mf* scale because they would allow classification according to the generally accepted four-quadrant model.

Subsequent research has questioned whether the new scales meet their intended purpose. Johnson, Jones, and Brems (1996) found weak relations between the new MMPI scales and better known measures of the two-dimensional model. Perhaps their most important finding was that the classification of undergraduates as androgynous, masculine, feminine, or undifferentiated using the MMPI scales was essentially unrelated to classification based on the BSRI short form. Although some correlates have been identified for the *GM* and *GF* scales (Butcher et al., 1989), suggesting the scales may have some interpretive value independent of the *Mf* scale, their value as a marker for androgyny is in question.

In this study, we present an attempt at addressing issues associated with the *Mf* scale from a different perspective. Given that the primary purpose of the MMPI is the evaluation of psychopathology, it is unlikely that gender role classification will ever be considered a significant reason for administering the MMPI. Information gathered from the MMPI about gender role will only be considered relevant if it enhances the description of the respondent's emotional difficulties (Todd & Gynther, 1988). To put it another way, we consider the primary problem identified by Wong (1984) not to be that *Mf* is based on a one-dimensional model but that extreme scores do not necessarily indicate anything about pathology.

The purpose of this study was to develop an MMPI scale called the Masculine–Feminine Pathology Scale (*Mfp*) that specifically focuses on gender-related aspects of symptomatology. The goal was to develop a scale that would indicate whether the respondent's pattern of symptoms as reflected in the MMPI are more typical of men or women clinical patients, or whether the symptom pattern reflects a balance between the two.

METHOD

Participants

The initial sample used for item selection included 988 psychiatric inpatients that completed the MMPI–2 during the period of 1993–1995. The questionnaire was administered at the time of hospitalization. Mean age was 36.7 years ($SD = 13.8$), and mean years of education was 13.9 ($SD = 3.1$). The sample included 542 women (54.9%) and 446 men (45.1%). The majority of respondents were either single (48.0%), married (29.0%), or divorced (11.3%).

Measures

In addition to the MMPI–2, in most cases the patient's primary therapist completed the Hopkins Psychiatric Rating Scale (HPRS; Derogatis, 1983) within 72 hr of admission to the inpatient service. The HPRS requests severity ratings on a 7-point scale ranging from 1 (*none*) to 7 (*extreme*) with anchor points for 17 symptom dimensions as well as a global severity rating. The HPRS is very similar in format and content to the more familiar Brief Psychiatric Rating Scale.

Procedure

The following procedure was used to select the items to be included in *Mfp*. All available patients were included in the item selection sample because an initial analysis indicated that the proportion of potentially invalid profiles did not vary as a function of gender. Profiles were identified as potentially invalid on the basis of the following criteria: more than 15 items left blank, *T* score over 64 on scale *L* or *K*, or *T* score over 90 on scale *F*. Table 1 provides the proportion of potentially invalid profiles by gender. The corresponding chi square test was not significant, $\chi^2(1, N=988)=.582, p=.446$. Accordingly, it was concluded that potentially invalid profiles would not differentially influence the item response patterns for the two genders.

The 567 MMPI–2 items were reviewed for inclusion in the *Mfp* scale using Baucom's (1976) criteria. Specifically, if at least 70% of either women or men endorsed the item in a certain direction, and if the proportion of members of the opposite sex endorsing that item in the same direction was at least 10% less, then the item was selected for possible inclusion. This step produced a pool of 77 items.

Because the focus of the scale was on pathology, any item that clearly was unrelated to pathology was then eliminated from the scale. This determination was undertaken independently by Robert E. McGrath and Elizabeth Sapareto. Discrepancies in decisions were reevaluated to make a final determination by consensus. This resulted in elimination of 22 items from the pool. These items referred either to career choice (e.g., enjoying the work of a soldier or building contractor) or gender-stereotyped activities and interests (e.g., enjoying mechanics magazines, dolls, or expensive clothes). Responses to Item 62 (whether one likes or would like

TABLE 1
Proportion of Potentially Invalid Profiles by Gender

	Probably Valid	*Potentially Invalid*
Women	65.9	34.1
Men	68.2	31.8

being a girl) were very strongly related to gender. However, this item varies depending on the gender of the respondent, and therefore was also eliminated because it did not offer a fair comparison between men and women. The final scale consisted of 54 items, of which 37 are included in the first 370 items of the MMPI–2 and may be considered a short form of the *Mfp* scale.

Review of the items led to two important decisions about the structure of the scale. First, in many cases the response alternative indicating potential pathology was not the response that led to selection of the item. For example, Item 2 (*I have a good appetite*) was included in the pool because it was answered true by 78.4% of men, but only 66.7% of women. According to the strategy developed by Baucom (1976), this item should be keyed true as a masculine item. However, the potentially symptomatic response is false. To develop a symptom-related gender scale, the decision was made to key items on the basis of the symptomatic response rather than the response that led to inclusion of the item. For example, Item 2 would be keyed false as an indicator of a female-typical symptom because male patients underreported poor appetite when compared to female patients.

Second, we had originally hoped to develop two scales, one for symptoms more common in men and one for symptoms more common in women, to parallel the accepted two-dimensional model of gender roles. However, this raised a new problem. If either a true or false response could be keyed depending on which response indicated symptomatology, a two-scale model would require distinguishing items that are bipolar indicators of pathology from those that are unipolar. Item 2 mentioned earlier is an example of a unipolar indicator because only a false response indicates pathology. According to the original plan this item would have been included in the female-typical pathology scale. Item 70 (*I am easily downed in an argument*) offers an example of a potentially bipolar item that was endorsed false more frequently by men than women. Either a true or false response to item 70 is potentially suggestive of psychopathology, depending on how the respondent answered other items. If we were to create two scales, an argument could be made for keying this item true on the feminine scale and false on the masculine scale. Rather than making questionable decisions about whether the items were either bipolar or unipolar pathology indicators, we decided to generate a single scale reflecting a relatively high level of endorsement of symptoms more common in female patients in one endpoint and in the other endpoint a relatively high level of endorsement of symptoms more common in male patients.

We also concluded that a four-quadrant classification model is unnecessary when attempting to measure the balance between female-typical and male-typical symptoms. Simultaneous elevation or depression on two scales would add nothing to achieving this goal; it would simply indicate the presence or absence of pathology in general, which can already be determined by examination of the standard clinical and validity scales. Based on these considerations, we felt the decision to develop a single bipolar scale was justified.

To generate a scoring key for the scale, all items were keyed so that higher scores were associated with the feminine endpoint, to be consistent with the *Mf* scale. The scoring key may be found in Table 2.

RESULTS

The remaining analyses were conducted using the 661 profiles that were considered valid according to the criteria listed previously. Demographic data for this subset were similar to those of the original sample. All analyses were conducted using SYSTAT 6.0 (SPSS, 1996).

Descriptive statistics for the inpatient sample used in this study may be found in Table 3. As expected, women demonstrated significantly higher scores on both the short form and the long form, and associated effect sizes were consistently large: short form, $d = 1.37$, $t(668) = 17.7$, $p < .01$; long form, $d = 1.57$, $t(668) = 20.2$, $p < .001$. The new scales were also found to be unrelated to age and education in the sample.

A factor analysis of the *Mfp* long form was conducted. The number of factors to retain was determined using parallel analysis (Horn, 1965), which Zwick and Velicer (1986) found superior to more familiar selection procedures such as the Kaiser or scree test. An unrotated principal components analysis was conducted first, and resulting eigenvalues were compared to values extrapolated from parallel analysis tables provided by Lautenschlager (1989), which provide eigenvalues derived from principal components analyses of random variables. The eigenvalues for the first nine components exceeded the tabled values, suggesting that nine factors should be retained for the factor analysis.

An iterative principal axis factor analysis was then conducted with varimax rotation. All items associated with loadings greater than ±.25 were reviewed to identify the factors. Table 4 presents rationally derived labels for the factors, and items

TABLE 2
Item Composition for the *Mfp* Scale

True											
23	28	30	40	70	73	100	111	121	127	205	263
266	331	351	380	392	395	435	442	457	469	471	532
536											
False											
2	3	8	9	10	20	27	63	68	84	103	139
140	157	164	231	238	239	250	254	269	287	406	431
432	462	510	548	559							

Note. Item numbers below 370 indicate items also included in the short form. *Mfp* = Masculine–Feminine Pathology Scale.

TABLE 3
Descriptive Statistics for the *Mfp* Scales

	Short Form	*Long Form*
Entire sample		
M	20.80	29.59
SD	5.24	7.08
Men		
M	17.60	24.87
SD	4.15	5.71
Women		
M	23.54	33.63
SD	4.15	5.47
Correlations		
Age	.084	.090
Education	.035	.019

Note. None of the correlations were significant. *Mfp* = Masculine–Feminine Pathology Scale.

TABLE 4
Factor Labels and Associated Items

1. Fear of the dark
   Positively related: 395, 435, 471
2. Amorality
   Positively related: 250, 269, 406, 432, 462, 548
   Negatively related: 100, 121, 266, 392
3. Imperturbability
   Positively related: 3, 9, 10, 63, 140, 157, 239
   Negatively related: 73, 127, 331, 469
4. Stress-related physical complaints
   Positively related: 23, 28, 30, 40, 380, 442, 471, 469, 536
   Negatively related: 2, 3, 8, 140, 164
5. Interpersonal hypersensitivity
   Positively related: 127, 205, 287, 331, 532
   Negatively related: 63, 157
6. Stomach problems
   Positively related: 28, 111
7. Interpersonal uncertainty
   Positively related: 70, 73, 380, 457
   Negatively related: 139
8. Stimulation seeking
   Positively related: 27, 103, 406
9. Acting out
   Positively related: 27, 84, 431, 548
   Negatively related: 266

used to identify the factor. The nine factors combined accounted for only 26% of total variance.

Despite the factorial complexity of the new scale, reliability was adequate, and greater than that for many of the more familiar MMPI–2 scales (Butcher et al., 1989, Appendix D). Coefficient alpha for the long form was estimated to be .78, for the short form .76.

Table 5 provides correlations between raw scores on the *Mfp* short and long forms, and standard validity and clinical scale *T* scores from the MMPI–2 as well as the *GM* and *GF* scales. The final column provides the number of items shared with the *Mfp* scale and the total number of items on that scale. Not surprisingly, the new scales showed a fair amount of overlap with the *GM* and *GF* scales (15–16 items). Among the standard scales, overlap was greatest with the first three clinical scales and the *Mf* scale. However, the number of items in common was generally small, never exceeding 9 out of 54 *Mfp* items.

Except for the *Mf* scale, results of correlational analyses for men and women separately were very consistent with those for the total sample. *Mfp* was significantly positively related to every clinical scale except *Mf* and *Ma*. As expected, the new scales were positively related to male *T* scores on scale *Mf* but negatively re-

TABLE 5
Correlations Between *Mfp* Short and Long Forms and Other MMPI Scales

	Entire		*Men*		*Women*		
Scale	*Short*	*Long*	*Short*	*Long*	*Short*	*Long*	*Item Overlap*
L	.119	.090	.101	.084	.110	.061	1/15
F	.279**	.264**	.227**	.205**	.279**	.268**	2/60
K	−.081	−.095*	−.089	−.099	−.094	−.128	2/30
Hs	.532**	.517**	.540**	.544**	.549**	.539**	8/32
D	.635**	.595**	.719**	.674**	.666**	.654**	9/57
Hy	.549**	.522**	.600**	.565**	.502**	.492**	9/60
Pd	.181**	.155**	.163	.128	.252**	.233**	5/50
Mf	−.174**	−.209**	.373**	.367**	−.353**	−.381**	9/56
Pa	.300**	.276**	.380**	.353**	.358**	.359**	3/40
Pt	.534**	.517**	.656**	.648**	.587**	.609**	6/48
Sc	.341**	.340**	.402**	.413**	.418**	.442**	3/78
Ma	−.253**	−.215**	−.367**	−.305**	−.212**	−.202**	7/46
Si	.394**	.380**	.448**	.439**	.482**	.495**	5/70
GM	−.710**	−.792	−.583**	−.673**	−.563**	−.687**	16/47
GF	.578**	.603**	.426**	.442**	.204**	.189**	15/46

Note. *Mfp* = Masculine–Feminine Pathology Scale. MMPI = Minnesota Multiphasic Personality Inventory (Hathaway & McKinley, 1983).

*$p < .05$. **$p < .001$.

lated to female *T* scores. The pattern of results suggests that the *Mfp* scale is not an indicator of any one clinical syndrome tapped by the traditional clinical scales.

Table 6 extends the information in the previous table, providing correlations and item overlap with the Harris and Lingoes (Butcher et al., 1989) subscales. Subscales are ordered according to those that were consistently positively related, consistently negatively related, and inconsistently related or unrelated to the new scales.

As the first MMPI gender scale developed to be specific to the domain of psychopathology, it was hypothesized that the *Mfp* scale should demonstrate some superiority over existing gender scales for the prediction of pathology. A series of analyses was conducted to evaluate this hypothesis using HPRS items as the criterion. The HPRS items provide information about a broad range of issues in psychopathology, but they do not address the issues most likely to be related to a gender-based measure of psychopathology. These issues are outlined in the Discussion section. For this reason, these analyses should be considered preliminary tests of the hypothesis.

To simplify the presentation, only the long form of *Mfp* was used for these analyses. Table 7 provides the correlations between the gender scales and ratings on the HPRS. In general, the new scales were positively related to indicators of neurotic distress (Somatization, Obsessive–Compulsive, Interpersonal Sensitivity, Depression, Anxiety, Phobic Anxiety) and relatively independent of psychotic indicators. Surprisingly, it was independent of the global rating of distress as well. One reasonable interpretation of this finding is that scores on the *Mfp* scale are not indicative of the general level of pathology or adjustment so much as the manner in which pathology manifests itself.

Correlations were on average stronger for the *Mfp* scale than for other gender scales. Except in the case of the *GM* scale for women, the mean gender scale-HPRS item correlation (computed using the root mean squared) was consistently larger for *Mfp* than for the other gender scales. To evaluate this further, each HPRS item was simultaneously regressed onto *Mf*, *GM*, *GF*, and *Mfp* separately for men and women. As indicated in Table 7, in 10 of 36 analyses the regression weight for *Mfp* was significant ($p < .05$) even after partialling out the other three gender scales. The same analyses were then conducted simultaneously regressing each HPRS item onto *Mfp* plus the nine MMPI clinical scales excluding *Mf*. In this case 3 of 36 regression weights for *Mfp* were significant (Somatization and Obsessive–Compulsive for men, and Abjection–Disinterest for women). The results support the tentative conclusion that the *Mfp* scale can predict some general aspects of psychopathology beyond the traditional clinical scales, and demonstrates incremental validity when compared with existing gender-based scales.

TABLE 6
Correlations Between *Mfp* and Harris and Lingoes Scales

	Entire		*Men*		*Women*		
	Short	*Long*	*Short*	*Long*	*Short*	*Long*	*Item Overlap*
Positively related							
Subj Dep (D1)	.629**	.593**	.667**	.621**	.627**	.616**	5/32
Psy Retard (D2)	.418**	.402**	.410**	.412**	.443**	.426**	1/14
Phys Mal (D3)	.499**	.483**	.481**	.469**	.508**	.503**	2/11
Mental Dull (D4)	.529**	.493**	.581**	.531**	.538**	.532**	3/15
Brooding (D5)	.542**	.518**	.558**	.522**	.520**	.521**	1/10
Lass–Mal (Hy3)	.609**	.568**	.627**	.586**	.617**	.584**	4/15
Som Comp (Hy4)	.495**	.506**	.417**	.436**	.458**	.473**	3/17
Soc Alien (Pd4)	.235**	.226**	.247**	.235**	.286**	.296**	1/13
Self-Alien (Pd5)	.329**	.316**	.369**	.372**	.406**	.403**	1/12
Poignancy (Pa2)	.347**	.337**	.336**	.306**	.350**	.368**	1/9
Naivete (Pa3)	.236**	.218**	.225**	.185*	.184**	.184**	0/9
Soc Alien (Sc1)	.232**	.234**	.190*	.195*	.311**	.327**	0/21
Emot Alien (Sc2)	.374**	.368**	.415**	.412**	.390**	.407**	1/11
Cog Ego Mas (Sc3)	.310**	.307**	.351**	.341**	.370**	.402**	0/10
Con Ego Mas (Sc4)	.426**	.402**	.479**	.445**	.455**	.460**	1/14
Def Inhib (Sc5)	.233**	.270**	.192*	.259**	.215**	.248**	1/11
Biz Sens Exp (Sc6)	.202**	.217**	.216**	.252**	.181*	.195**	1/20
Negatively related							
Den Soc Anx (Hy1)	−.191**	−.185**	−.205**	−.207**	−.296**	−.303**	0/6
Author Prob (Pd2)	−.453**	−.451**	−.373**	−.377**	−.302**	−.279**	3/8
Soc Impert (Pd3)	−.294**	−.285**	−.294**	−.286**	−.361**	−.369**	1/6
Amorality (Ma1)	−.492**	−.457**	−.500**	−.457**	−.383**	−.351**	3/6
Imperturb (Ma3)	−.317**	−.311**	−.309**	−.309**	−.321**	−.323**	0/8
Unrelated–inconsistent							
Need Affect (Hy2)	.053	.047	.057	.032	−.013	−.008	1/12
Inhib Agg (Hy5)	.169**	.147**	.228**	.193*	.036	.017	1/7
Famil Disc (Pd1)	.099*	.078*	−.005	−.048	.133*	.116*	0/9
Persec Idea (Pa1)	.042	.037	.100	.115*	.145*	.141*	0/17
Psy Accel (Ma2)	−.154**	−.135**	−.124*	−.079	−.142*	−.136*	1/11
Ego Inflat (Ma4)	−.039	−.015	−.089	−.036	.009	.015	0/9

Note. Full titles of the Harris and Lingoes subscales may be found in Butcher, Dahlstrom, Graham, Tellegen, and Kaemmer (1989). *Mfp* = Masculine–Feminine Pathology Scale.

*$p < .05$. **$p < .001$.

TABLE 7
Correlations Between Gender Scales and Hopkins Psychiatric Rating Scale Items (Derogatis, 1983)

	Men				*Women*			
	Mf	*GM*	*GF*	*Mfp*	*Mf*	*GM*	*GF*	*Mfp*
Somatization	.052	−.031	.141	.149*	.059	−.134*	.076	.116
Obsessive–Compulsive	.283*	−.138	.286*	.292*[a]	.030	−.117	.112	.142*
Interpersonal Sensitivity	.220*	−.207*	.139	.285*[a]	−.099	−.271*	.097	.209*
Depression	.217*	−.244*	.081	.296*[a]	.010	−.239*	.059	.206*
Anxiety	.170*	−.179*	.126	.224*	.041	−.191*	.083	.158*
Hostility	−.090	.072	−.098	−.068	.101	−.062	−.104	−.037
Phobic Anxiety	.041	−.171*	.080	.179*	−.001	−.201*	.160*	.146*
Paranoid Ideation	.007	.052	.102	−.087[a]	.063	−.111	−.028	−.070[a]
Psychoticism	−.016	−.006	.085	−.078[a]	.016	−.040	−.006	−.037
Sleep Disturbance	.175*	−.094	.096	.109	.062	−.110	−.064	.152*[a]
Psychomotor Retardation	.014	−.108	.048	.057	.066	−.083	.019	.026
Hysterical Behavior	.058	−.010	.005	−.059	.014	−.146*	−.019	−.005
Abjection–Disinterest	.058	−.147*	−.025	.091	.046	−.092	−.090	.015
Conceptual Dysfunction	−.109	−.045	.053	−.109[a]	.042	−.115	−.027	.058
Disorientation	−.167*	−.020	.008	−.064	.044	−.063	−.028	.031
Excitement	−.122	.117	.023	−.080	−.030	−.098	−.050	−.051[a]
Euphoria	−.094	.141	.015	−.113	.060	−.156*	−.071	−.050[a]
Global Pathology Index	.005	−.183*	.008	.123	.049	−.115	−.031	.101
Mean r	.133	.130	.104	.159	.053	.143	.074	.110

Note. *Mf* = Masculinity–Feminity scale; *GM* = Gender Role–Masculine; *GF* = Gender Role–Feminine; *Mfp* = Masculine–Feminine Pathology Scale.

[a]Cases where simultaneous regression of the Hopkins Psychiatric Rating Scale item onto the four gender scales resulted in a significant regression weight for the *Mfp* scale.

*$p < .05$

DISCUSSION

The results of the factor analysis suggest the new scale is a factorially complex one. A factor analysis that extracted a relatively large number of factors only accounted for 26% of total variance, and many items were not markedly related to any of the factors. This is consistent with evidence concerning the *Mf* scale (Wong, 1984), suggesting that the empirical correlates of gender are quite diverse in content. Even so, some statements about the general nature of the scale are warranted.

High scores seem to reflect anxious distress, characterized by clearly stress-related somatic complaints, hypersensitivity, and interpersonal sensitivity. It is interesting to note that although no factor emerged in the factor analysis associated with depressive symptoms, the scale correlates as well with depression measures as it does with anxiety measures. This finding is consistent with previous

studies indicating that multitrait correlations between measures of anxiety and depression are about as high as their respective monotrait correlations. This has been interpreted as suggesting that both tend to be heavily influenced by the target individual's general experience of distress (e.g., McGrath & Ratliff, 1993). Low scores seem to indicate a calm facade potentially associated with a willingness to engage in amoral activities. None of the clinician ratings correlated negatively with *Mfp*, probably because of the emphasis on distress indicators in the HPRS. In particular, it should be noted that the scale was unrelated to ratings of hostility, suggesting that low scorers on the scale may not be perceived by clinicians as overtly hostile individuals, despite a possible history of difficulty with authorities.

The scale seems essentially unrelated to psychotic features. Although there were some positive correlations with psychotic scales, these more likely stem from the high loading of distress-related items on those scales. For example, although *Mfp* correlated with the Paranoia scale, it was unrelated to the Persecutory Ideas subscale. Similarly, it correlated significantly with the Schizophrenia scale but actually only shares three items.

Unlike the *GM* and *GF* scales, the *Mfp* scale was developed specifically to enhance the understanding of the respondent's symptom picture. Although the *Mf* scale was originally intended to serve a similar purpose, its capacity to do so is compromised by the inclusion of items that are no longer gender related because of changes in gender roles and of items that are unrelated to pathology. At least three lines of further research with *Mfp* are warranted. First, although it was found that the *Mfp* demonstrated some capacity to predict general aspects of psychopathology over and above existing MMPI scales, as indicated previously, this is not the domain of criterion variables the scale is best suited to evaluate. Future research on the clinical correlates of the scale would be enhanced by collecting ratings of dimensions more likely to reflect both the high and low endpoints based on the descriptions described earlier. Variables to consider would include passivity, introversion–extroversion, imperturbability, motivation for change, passive resistance, motivation for treatment, capacity for guilt, explosiveness and overcontrol, and willingness to accept blame. Second, gender role satisfaction among individuals with gender-inconsistent scores might merit study. Finally, the new scale could prove a useful moderator in the interpretation of existing MMPI code types. We hypothesize, for example, that individuals with elevated scores on *Mfp* would be more overt in their expressions of distress and be more help-seeking, whereas individuals with relatively low scores would seem more comfortable with their disorder, more likely to defy authority, and more resistant to attempts at intervention.

ACKNOWLEDGMENT

We thank Michele Zaccario for her assistance in gathering the data for this study.

REFERENCES

Baucom, D. H. (1976). Independent masculinity and femininity scales on the California Psychological Inventory. *Journal of Consulting and Clinical Psychology, 44,* 876.

Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 47,* 155–162.

Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *MMPI–2: Manual for administering and scoring.* Minneapolis: University of Minnesota Press.

Constantinople, A. (1973). Masculinity–femininity: An exception to a famous dictum. *Psychological Bulletin, 80,* 389–407.

Derogatis, L. R. (1983). *SCL–90–R administration, scoring, and procedures manual.* Towson, MD: Clinical Psychometric Research.

Hathaway, S. R., & McKinley, J. C. (1983). *Manual for the administration and scoring of the MMPI.* Minneapolis, MN: National Computer Systems.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30,* 179–185.

Johnson, M. E., Jones, G., & Brems, C. (1996). Concurrent validity of the MMPI–2 feminine gender role (*GF*) and masculine gender role (*GM*) scales. *Journal of Personality Assessment, 66,* 153–168.

Lautenschlager, G. J. (1989). A comparison of alternatives to conducting Monte Carlo analyses for determining parallel analysis criteria. *Multivariate Behavioral Research, 24,* 365–395.

McGrath, R. E., & Ratliff, K. G. (1993). The use of self-report measures to corroborate theories of depression: The specificity problem. *Journal of Personality Assessment, 61,* 156–168.

Peterson, C. D., & Dahlstrom, W. G. (1992). The derivation of gender-role scales *GM* and *GF* for MMPI–2 and their relationship to scale 5 (*Mf*). *Journal of Personality Assessment, 59,* 486–499.

SPSS. (1996). *SYSTAT for Windows v. 6.0: Command reference.* Chicago: Author.

Todd, A. L., & Gynther, M. L. (1988). Have MMPI *Mf* scale correlates changed in the past 30 years? *Journal of Clinical Psychology, 44,* 505–510.

Wong, M. R. (1984). MMPI scale five: Its meaning, or lack thereof. *Journal of Personality Assessment, 48,* 279–284.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99,* 432–442.

Robert E. McGrath
School of Psychology
T110A
Fairleigh Dickinson University
1000 River Road
Teaneck, NJ 07666

Received September 25, 1996
Revised May 5, 1997