

A Psychometric Analysis of the Personal Attributes Questionnaire¹

**Robert L. Helmreich,² Janet T. Spence,
and John A. Wilhelm**

The University of Texas at Austin

The psychometric properties of the Personal Attributes Questionnaire were examined in independent samples of male and female high school students, college students, and adults. In each of the six samples a two-factor structure (masculinity/instrumentality and femininity/expressivity) paralleling the empirically derived scales was found. Additional factor analyses of negative masculine and feminine traits were reported. Discriminant analyses revealed highly significant differentiation between the sexes. The reliabilities (Cronbach alpha) of the unit-weighted scales in each sample were also satisfactory.

The two major instruments used to assess psychological masculinity and femininity, the Personal Attributes Questionnaire (PAQ; Spence, Helmreich, & Stapp, 1975; Spence & Helmreich, 1978) and the Bem Sex Role Inventory (BSRI; Bem, 1974), each contain separate masculinity (M) and femininity (F) scales that have a low degree of correlation with each other. One of the theoretical assumptions on which these instruments are based is that the masculinity and femininity, as measured by the M and F scales, are each a unidimensional phenomenon. A number of factor analytic studies (almost all on the BSRI) have challenged this assumption (e.g., Pedhazur & Tetenbaum, 1979, Gaudreau, 1977; Moreland, Gulanick, Montague, & Harren, 1978; Edwards, Gaa, & Liberman, Note 1). For

¹ Support for this research was provided by NASA Grant NSG 2065 (Robert L. Helmreich, Principal Investigator) and NSF Grant BNS 78-08911 (Janet T. Spence and Robert L. Helmreich, Principal Investigators). The article was prepared while the second author was a Fellow at the Center for Advanced Study in the Behavioral Sciences, with support from the Spencer Foundation, NIMH Grant 5T32MH14581-03, and the University of Texas Research Institute.

² Correspondence should be sent to Robert L. Helmreich, Department of Psychology, The University of Texas at Austin, Austin, Texas 78712.

example, four rather than two factors have been demonstrated for the BSRI M and F scales by Pedhazur and Tetenbaum (1979), among others. In one of the few studies involving the PAQ, Edwards et al. (Note 1) have also argued for a four-factor solution.

The methodological adequacy of the studies devoted to examining the BSRI's factor structure, as Pedhazur and Tetenbaum (1979) have noted, has not been uniformly high. One methodological flaw from which many of these studies suffer, they suggest, is the practice of combining data from heterogeneous samples for factoring. In particular, factor analyses of combined male and female respondents are clearly inappropriate, since highly significant sex differences are found on both the M and F scales (and probably most items). Several of the studies also suffer from common methodological problems in factor analysis that have been described by Comrey (1978). These include the mechanical rotation and interpretation of all factors with eigenvalues greater than 1.

Despite the methodological flaws in some of the studies, the factorial complexity of the BSRI is beyond dispute, an outcome that has led to criticism of the BSRI and the PAQ on both conceptual and psychometric grounds. However, the PAQ M and F scales are more homogeneous in content than the corresponding BSRI scales and might be factorially simpler. No firm conclusions can be drawn about structure of the PAQ, however, since data have been available only from several small studies. The present study was designed to remedy this deficiency. To be reported are the results of factor analyses and discriminant analyses of the PAQ obtained from males and females in four large independent samples: high school students, two groups of college students, and married individuals who are parents. For one of the college samples, data were available on responses to an expanded version of the PAQ (EPAQ), containing several additional scales. These scales are described below.

METHOD

Measures

PAQ. Interpretation of the analyses of the Personal Attributes Questionnaire may be facilitated by a brief review of the theoretical and empirical properties of the instrument. The version of the PAQ administered to all subjects was the 24-item short form (Spence & Helmreich, 1978). The items, all of which consist of trait descriptions set up on 5-point scales, are assigned to three 8-item scales labeled Masculinity (M), Femininity (F), and Masculinity-Femininity (M-F). Scores on each scale are obtained by summing item scores and can range from 0-32; the M and M-F scales are scored in a masculine direction and the F scale in a feminine direction.

In pilot work leading to the development of the scales, the items on the M scale had all been judged to be stereotypically more characteristic of men, but socially desirable in both sexes. Similarly, the items on the F scale had been judged to be more characteristic of women, but socially desirable in both sexes. The M-F scale contains items on which social desirability ratings for the sexes differed, with the ideal man falling toward the stereotypic masculine pole and the ideal woman toward the stereotypic feminine pole.

In content, the M items reflect goal oriented, instrumental traits and the F items interpersonally oriented, expressive traits. The M-F scale is mixed in content, containing 2 instrumental items ("aggressive" and "dominant") and 6 items suggesting emotional vulnerability.

The M and F scales have been found to be essentially orthogonal in both sexes. The M-F scale (scored in a masculine direction) has more bipolar properties, showing a moderate positive correlation with the M scale and a smaller but significant negative correlation with the F scale. Sex differences have consistently been found on all three scales in groups of widely varying age and socioeconomic origin, males scoring significantly higher than females on M and M-F and lower on F. (See Spence & Helmreich, 1978, for a review).

EPAQ. Theoretical interest in the socially undesirable components of masculine instrumentality and feminine expressiveness has led us to develop additional scales that distinguish the sexes, using an item selection procedure analogous to that used in the development of the PAQ M and F scales (Spence, Helmreich, & Holahan, 1979). The new M scale (labeled negative M or M^- to distinguish it from the original positive M scale) contains 8 trait descriptions, each of which had been judged in pilot work to be stereotypically more characteristic of males than females, but socially undesirable in both sexes (i.e., ratings of the ideal male and female fell toward the pole indicating absence of the trait). Of the traits found in pilot work to meet these criteria, 8 were chosen that reflected negative characteristics that were *instrumental* (e.g., egotistical, arrogant). The M^- scale is thus a conceptual parallel of the M scale. In like fashion, items selected to reflect negative femininity had been judged to be more characteristic of females, but undesirable in both sexes. Although our aim was again to provide a conceptually parallel scale by identifying undesirable *expressive* traits (or, as Bakan, 1966, has labeled it, "unmitigated communion" or selflessness), this goal proved elusive. Two clusters of negative feminine traits were chosen for study because of their potential interest. The first consists of 4 traits that come close to Bakan's "communion": servile, spineless, gullible, subordinates self to others. These items are assigned to a scale labeled F_C^- . The other set of 4 items describes a kind of verbal aggressiveness: whiny, complaining, nagging, and fussy. These items were assigned to a scale labeled F_{VA}^- .

The 16 items making up the M^- , F_C^- and F_{VA}^- scales have been added to the PAQ to form the Extended Personal Attributes Questionnaire (EPAQ). Our initial study with the EPAQ (Spence et al., 1979) indicated significant

differences between the sexes in the expected direction on the three negative scales as well as on the original scales (M, F, M-F). Correlations between same-sex-typed positive and negative scales (e.g., M and M^-) tended to be low but positive, while cross-sex-typed correlations (e.g., M vs. F_{VA}^-) were negative.

Subject Populations

Data for the analyses came from four distinct populations: (1) 674 female and 509 male students from New England high schools, given the PAQ in their classrooms (as described in Spence & Helmreich, 1978); (2) 1,585 female and 1,251 male students of introductory psychology at The University of Texas at Austin, given the PAQ as part of their course experience; (3) 1,465 female and 854 male introductory psychology students at the same university, given the 40-item EPAQ; and (4) 1,028 mothers and 926 fathers, who completed the PAQ as part of a mail survey. The latter sample was composed of parents of college students who had been given the PAQ (Spence & Helmreich, 1978) and parents of children enrolled in Austin public elementary schools. Differences in PAQ scale scores across age ranges tend to be small, although occasionally significant in large samples. Age effects are described in Spence & Helmreich (1979).

RESULTS

Characteristics of the Item Correlation Matrices

Initial analyses consisted of computation of the Kaiser-Meyer-Olkin measure of sampling adequacy and Bartlett's Index of Sphericity, as recommended by Dziuban and Shirkey (1974). These computations were done separately for males and females in the high school, college, and adult samples. Separate computations were made for the 16 items comprising the M and F scales of the PAQ, the full 24-item PAQ, and the 40-item EPAQ. According to Kaiser's (1970) classification scheme, all 16 of the matrices examined were classified as "adequate" or "meritorious" for factor analyses. The values of the measure ranged from a low of .76 to a high of .85 (mean = .82). Bartlett's test of sphericity was significant at better than the .001 level for each of the 16 matrices.

Factor Analyses of the PAQ M and F Items

Because the 16 items forming the M and F scales of the PAQ describe the two conceptual domains explored most frequently in research, initial factor analyses were computed using this matrix as input. Six factor analyses were performed involving males and females in each of the age groups. The Jöreskog fac-

Table 1. Rotated Oblique Factor Pattern Matrices of Masculine and Feminine Items for Male and Female High School, College, and Parent Samples
(Delta = 0)^a

	High school students				College students				Parents			
	Males		Females		Males		Females		Males		Females	
	Factor I Masc	Factor II Fem	Factor I Masc	Factor II Fem	Factor I Masc	Factor II Fem	Factor I Masc	Factor II Fem	Factor I Masc	Factor II Fem	Factor I Masc	Factor II Fem
M scale items												
Independent	.23	.11	.42	.06	.37	.07	.45	-.05	.46	-.03	.58	-.06
Active	.43	.08	.33	.16	.53	.13	.44	.14	.50	.09	.51	.16
Competitive	.54	-.01	.33	.03	.51	-.04	.44	-.03	.54	.01	.51	-.04
Decisive	.33	-.04	.38	-.04	.44	.00	.39	-.12	.50	.03	.44	.06
Never gives up	.49	-.07	.52	.00	.52	-.04	.51	.10	.54	.04	.55	-.01
Self-confident	.56	.09	.71	.04	.71	-.04	.72	.06	.73	-.01	.77	.05
Feels superior	.45	.04	.62	-.01	.63	-.11	.60	-.01	.56	-.09	.61	.02
Stands up under pressure	.54	.00	.47	.01	.54	.05	.46	.05	.60	.11	.38	.09
F scale items												
Emotional	-.08	.33	.16	.24	-.14	.33	-.15	.27	-.06	.23	.12	.24
Devotes self	.04	.51	-.04	.42	.02	.49	-.01	.42	-.02	.60	-.05	.51
Gentle	-.27	.33	.08	.46	-.24	.41	-.08	.43	-.16	.46	.11	.52
Helpful	.13	.53	.13	.56	.16	.60	.11	.52	.18	.64	-.14	.64
Kind	.05	.58	.04	.66	.10	.64	.03	.62	.06	.70	-.04	.71
Aware of feelings	.03	.56	.02	.54	.02	.58	.03	.54	.09	.60	.00	.61
Understanding	.05	.60	.03	.66	.12	.66	.10	.67	.11	.71	-.06	.68
Warm	.13	.48	.10	.53	.20	.52	.13	.57	.12	.58	.20	.57
Factor intercorrelations		.13		-.12		.12		-.08		.16		-.13

^a M scale items are identified by the masculine pole; F scale items are identified by the feminine pole.

tor analysis subroutine in *SPSS* (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975) was employed, using the maximum likelihood solution with oblique rotation and varied degrees of obliqueness. The rotated factor pattern loadings for each group are shown in Table I.³ In each of the six factor analyses, a two-factor solution was optimal. Two large factors emerged with eigenvalues around 4. When a three-factor solution was computed, the third factor proved to be highly correlated with the first factor. Accordingly, the solution was reduced to two factors employing a delta value of zero. As the data in Table I indicate, each analysis reproduces the M and F scales of the PAQ. Across the six analyses the mean loading of M items on the masculinity factor was .51 and the mean loading of the F items on the F scale was .53. The corresponding mean loadings of M items on the Femininity factor and F items on the Masculinity factor were .06 and .09. The mean intercorrelation of the Masculinity and Femininity factors was $-.04$. The results provide strong support for the dualistic conceptualization of masculinity and femininity, and the unitary constructs of instrumentality and expressiveness.

Factor Analyses of the 24-Item PAQ

The same six factor analyses were repeated for the complete 24-item matrix of the PAQ. The rotated factor matrices from these analyses are shown in Table II. Again, two large factors appeared in each sample. Two small factors also appeared, but produced a conceptually unclear structure when included in the rotation. Additionally, the small factors proved to be highly correlated with one of the main factors (r 's in excess of .30). Therefore, a two-factor solution was chosen. The two instrumental items, "aggressive" and "dominant," load strongly and solely on the masculinity factor. "Worldly" (vs. "home-oriented") also loaded on the masculinity factor. The remaining items reflecting the presence or absence of emotional vulnerability, such as "feelings easily hurt," tended to load on both factors — positively on the femininity factor and negatively on the masculinity factor. This outcome appears to reflect the bipolar nature of these attributes. It also stresses the utility of retaining the theoretically distinct M-F items as a separate scale.

Factor Analyses of the EPAQ Negative Items

The intercorrelations of the six empirically derived scales of the EPAQ are shown in Table III for the sample of college students. Inspection indicates that

³In every case the factor pattern and factor structure matrices were nearly identical and the primary factor loading of each item was the same. Thus, for brevity only the pattern loadings are shown.

Table III. Intercorrelations of the Six Scales of the EPAQ for 854 Male and 1,465 Female College Students^a

	M	F	M-F	M ⁻	F _{VA} ⁻	F _C ⁻
M		.16	.46	.12	-.17	-.35
F	.08		-.24	-.37	-.13	.11
M-F	.50	-.28		.16	-.14	-.36
M ⁻	.12	-.44	.20		.42	-.13
F _{VA} ⁻	-.23	-.16	-.22	.37		.16
F _C ⁻	-.33	.17	-.41	-.11	.11	

^aMales are above the diagonal, females are below; $df = 800$, $p_{.01} = .09$.

the negative scales have quite low correlations with their parallel positive scales, but substantial *negative* correlations with the opposite scales. For example, the correlation of M and M⁻ in males is .12, while the correlation of M⁻ with F is $-.37$.

Analyses of the 16 negative items were conducted separately for male and female college students, and the rotated factor pattern matrices are shown in Table IV. Alone among the analyses, a differential factor structure emerged for

Table IV. Rotated Oblique Factor Pattern Matrices of Negative Masculine and Feminine Items (EPAQ) for Female and Male College Students ($\Delta = 0$)^a

	Males		Females		
	Factor I M ⁻ and F _{VA} ⁻	Factor II F _C ⁻	Factor I M ⁻	Factor II F _{VA} ⁻	Factor III F _C ⁻
M⁻ scale items					
Arrogant	.54	-.13	.68	-.09	.02
Boastful	.37	-.02	.42	.07	.06
Egotistical	.51	-.16	.76	-.18	.03
Greedy	.47	.04	.35	.19	-.02
Dictatorial	.50	-.14	.30	.18	-.23
Cynical	.40	.00	.30	.14	-.15
Unprincipled	.39	.02	.30	.10	-.03
Hostile	.50	.04	.29	.26	-.06
F_{VA}⁻ scale items					
Whiny	.25	.34	.00	.42	.10
Complaining	.44	.37	-.02	.69	.01
Nagging	.49	.33	-.02	.64	-.01
Fussy	.37	.23	.08	.41	-.02
F_C⁻ scale items					
Spineless	.06	.44	.15	.09	.37
Subordinates self	-.31	.35	-.04	-.08	.50
Servile	-.20	.36	-.01	-.03	.47
Gullible	-.01	.33	-.07	.13	.26
Factor intercorrelations		.05		.44	-.21
				└.08┐	

^aM⁻ scale items are identified by the masculine pole; F_{VA}⁻ and F_C⁻ items are identified by the feminine pole.

males and females. In the case of males, the negative masculine items (M^-) and the feminine verbal aggression items (F_{VA}^-) formed a single factor. The items assigned to the F_C^- scale formed a separate dimension, which tended also to have positive loadings for the verbal aggression items. In females, a three-factor solution reproduced the three scales (M^- , F_{VA}^- , and F_C^-), with the negative masculine traits yielding the strongest factor. Reflecting the positive relationship between undesirable verbal aggressive and instrumental traits demonstrated in the common factor for males, these two factors were positively correlated in females ($r = .40$). Justification for the retention of the M^- and F_{VA}^- items as discrete scales in men as well as women lies primarily in the differential item endorsements of the two sexes, men reporting themselves significantly higher on negative masculine traits and significantly lower on negative feminine traits than women (Spence et al., 1979).

Discriminant Analyses

While research with the PAQ and EPAQ has overwhelmingly shown significant sex differences on the scales and their constituent trait items (justifying the designations masculinity and femininity), the distributions of each sex on items and scales show substantial overlap. To provide a more quantitative indication of the ability of the instrument to differentiate the sexes, discriminant analyses were conducted using the 24-item PAQ and 40-item EPAQ. Wilks' lambdas for the PAQ were .58, .68, and .58 for the high school student, college student, and parent samples, respectively. Correct gender classification in each of these groups was 80%, 77%, and 81%, with highly significant chi squares ($ps < .0001$).

A parallel analysis of male and female responses to the EPAQ was computed for the sample of college students given this instrument, with a Wilks' lambda of .63 and the correct classification of 79% of the respondents as to gender (again with a highly significant chi square, $p < .0001$). Thirty-nine of the 40 trait descriptions yielded univariate F s significant beyond the .01 level, the largest sex difference being found for the (M-F) item "cries a lot" ($F = 747$).

These results are to be contrasted with those reported by Pedhazur and Tetenbaum (1979) in a discriminant analysis of the BSRI in which the gender of 97% of the respondents was correctly identified. These investigators, however, also reported that 94% correct classification was obtained using only two items, "masculine" and "feminine," and that only these two items produced substantial loadings on the discriminant function. It should also be noted that in factor analyses, "masculine" and "feminine" formed a separate, bipolar factor (i.e., with loadings that were opposite in sign). These findings suggest that self-attributions of global masculinity and femininity are tantamount to inquiries about biological sex. Comparison of the discriminant function analyses of the BSRI and PAQ is therefore not helpful in understanding the more psychological, personality factors they tap.

Table V. Alpha Coefficients for the PAQ and EPAQ Scales^a

	High school students		College students		Parents	
	Males	Females	Males	Females	Males	Females
PAQ scales						
M	.67	.71	.76	.73	.78	.77
F	.72	.73	.76	.73	.80	.79
M-F	.53	.62	.61	.65	.56	.61
EPAQ scales						
M			.74	.74		
F			.75	.75		
M-F			.54	.63		
M ⁻			.69	.70		
F ^{VA} - ^b			.60	.63		
F _C ⁻ - ^b			.46	.41		

^aM, F, and M-F scales are identical for both the PAQ and EPAQ; the EPAQ also contains three negative scales.

^bThese scales contain four items; all other have eight.

Internal Consistency of the PAQ and EPAQ Scales

The final analyses consisted of determining the internal consistency of the unit-weighted scales of the PAQ and EPAQ in each sample. Table V gives the alpha coefficients for each scale in each group. As the data indicate, the reliabilities for these 8- and 4-item scales are satisfactory and quite consistent across samples. Of the PAQ scales, the M-F scale shows the lowest reliability, doubtlessly reflecting its mixed content. On the EPAQ, the F_C⁻ scale, which is the most difficult to conceptualize, shows the lowest reliability of the six scales on that instrument.

DISCUSSION

The development of the 24-item PAQ was guided simultaneously by empirical and theoretical considerations. In the initial PAQ (Spence et al., 1975) items were selected that reflected sex stereotypes and were assigned to scales on the basis of social desirability ratings. In developing the short form, the content of the scales was refined so that the M and F scales were more exclusively measures of instrumentality and expressivity. The results reported above provide reassuring evidence in support of this conceptual structure, indicating that the M and F scales can be reproduced factor analytically, that the factor structure is highly consistent in both sexes in samples representing a considerable portion of the life span, and that the internal consistency of unit-weighted scales reflect-

ing these dimensions is satisfactory. The EPAQ data also justify the separation of socially desirable and undesirable masculine and feminine traits into separate scales. While the factor analyses did not perfectly mirror scale assignments of the M-F items and in males, for the F_{VA}^- items, theoretical considerations, which are not contradicted by the nature of the factor analytic results, suggest the utility of retaining these items as independent scales.

The unidimensionality of the PAQ scale is to be compared with the greater factorial complexity of the BSRI, an outcome that appears to reflect the more heterogeneous content of the latter in terms of social desirability and item content. These results suggest that even on the level of abstract personality traits, "masculinity" and "femininity" are multidimensional phenomena whose components are not strongly related to each other and not necessarily related to criterion variables in the same way or to the same degree.

Considerable evidence is accumulating for the construct and predictive validity of the PAQ scales (e.g., Spence & Helmreich, 1978; Helmreich & Spence, 1978; Helmreich, Spence, & Holahan, 1979; Spence et al., 1979; Klein & Wilerman, 1979) when these scales are regarded narrowly as measures of instrumentality and expressiveness. Since the scales differentiate the sexes, data obtained with them throw light on a number of gender-related phenomena in which the traits measured by the scales are quite directly involved. The PAQ and other personality instruments cannot be regarded, however, as all-purpose measures of masculinity and femininity.

REFERENCE NOTE

1. Edwards, T. A., Gaa, J. P., & Liberman, D. *A factor analysis of the BSRI and the PAQ*. Paper presented at the annual meeting of the American Psychological Association, Toronto, 1978.

REFERENCES

- Bakan, D. *The duality of human existence*. Chicago: Rand McNally, 1966.
- Bem, S. L. The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 1974, 42, 155-162.
- Comrey, A. Common methodological problems in factor analytic studies. *Journal of Consulting and Clinical Psychology*, 1978, 46, 648-659.
- Dziuban, C.D., & Shirkey, E. C. When is a correlation matrix appropriate for factor analysis? Some decision rules. *Psychological Bulletin*, 1974, 81, 358-361.
- Gaudreau, P. Factor analysis of the Bem Sex Role Inventory. *Journal of Consulting and Clinical Psychology*, 1977, 45, 299-302.
- Helmreich, R., & Spence, J. T. The Work and Family Orientation Questionnaire (WFOQ): An objective instrument to assess components of achievement motivation and attitudes toward family and career. *JSAS Catalog of Selected Documents in Psychology*, 1978, 8, 35, (Ms. No. 1677).
- Helmreich, R. L., Spence, J. T., & Holahan, C. K. Psychological androgyny and sex-role flexibility: A test of two hypotheses. *Journal of Personality and Social Psychology*, 1979, 37, 1631-1644.

- Kaiser, H. F. A second generation little jiffy. *Psychometrika*, 1970, 35, 401-416.
- Klein, H. N., & Willerman, L. Psychological masculinity and femininity and typical and maximal dominance expression in women. *Journal of Personality and Social Psychology*, 1979, 37, 2059-2070.
- Moreland, J. R., Gulanick, N., Montague, E. K., & Harren, V. A. Some psychometric properties of the Bem Sex Role Inventory. *Applied Psychological Measurement*, 1978, 2, 249-256.
- Nie, N., Hull, H., Jenkins, J., Steinbrenner, K., & Bent, P. *SPSS: Statistical Package for the Social Sciences*. New York: McGraw-Hill, 1975.
- Pedhazur, E. J., & Tetenbaum, T. J. Bem Sex Role Inventory: A theoretical and methodological critique. *Journal of Personality and Social Psychology*, 1979, 37, 996-1016.
- Spence, J. T., & Helmreich, R. L. *Masculinity and femininity: Their psychological dimensions, correlates and antecedents*. Austin: University of Texas Press, 1978.
- Spence, J. T., & Helmreich, R. L. A comparison of masculine and feminine personality attitudes across age groups. *Developmental Psychology*, 1979, 15, 583-584.
- Spence, J. T., Helmreich, R. L., & Holahan, C. K. Negative and positive components of psychological masculinity and femininity and their relationships to neurotic and acting out behaviors. *Journal of Personality and Social Psychological*, 1979, 37, 1673-1682.
- Spence, J. T., Helmreich, R., & Stapp, J. Ratings of self and peers on sex-role attributes and their relation to self-esteem and conceptions of masculinity and femininity. *Journal of Personality and Social Psychology*, 1975, 32, 29-39.