# Masculinity, Femininity, and Androgyny Viewed and Assessed as Distinct Concepts

## David Lubinski, Auke Tellegen, and James N. Butcher
### University of Minnesota

The objective of this study ($N = 172$) was to evaluate (a) the equivalence of the scales of the short Bem Sex-Role Inventory (BSRI) and the Extended Personal Attributes Questionnaire (EPAQ); (b) the construct validity of the short BSRI and EPAQ as measures of well-being; (c) the validity of the concept of androgyny as an intrinsically interactive (rather than simply additive) concept; and (d) the utility and meaning of two special EPAQ measures: unmitigated agency ($M^-$) and unmitigated communion ($F_C^-$). The short BSRI and EPAQ were found empirically interchangeable when placed in a multitrait–multimethod matrix and two extrinsic convergent validation rectangles. Results of a hierarchical multiple regression analysis with interaction terms obtained with the Differential Personality Questionnaire (DPQ) provided only partial support for masculinity and femininity as measures of psychological well-being and no support for the significance of androgyny treated as an interaction of masculinity and femininity. Unmitigated agency and communion did not show the expected negative correlations with the mutual mitigation (interaction) of masculinity and femininity. Implications of this study for future investigations of androgyny and other integrative concepts defined by interactions are discussed.

As the concept of androgyny has evolved (e.g., Bem, 1974, 1977, 1979b; Lubinski, Tellegen, & Butcher, 1981; Spence & Helmreich, 1979a; Spence, Helmreich, & Strapp, 1975; Strahan, 1975), so has its assessment. For example, the Bem Sex-Role Inventory (BSRI; Bem, 1974) has been refined to the short BSRI (Bem, 1979b), and the Personal Attributes Questionnaire (PAQ; Spence et al., 1975) has been expanded to the Extended PAQ (EPAQ; Spence, Helmreich, & Holahan, 1979). The short BSRI and the EPAQ currently represent the two most refined inventories that purport to measure androgyny and are certainly the most widely used. Both inventories classify individuals into four groups—masculine, feminine, androgynous, and undifferentiated—using a median-split method.

In the early to middle 1970s, androgyny was claimed by some to be an index of adjustment and psychological health (e.g., Bem,

1974; Kaplan & Bean, 1976; Rawlings & Carter, 1977). Bem (1974), one of the advocates of this position, proposed that the BSRI be used to classify individuals with respect to psychological health and that androgynous individuals are exemplars of psychological health. Kaplan and Bean (1976) suggested that psychotherapeutic techniques be developed to change both masculine and feminine sex types to a more androgynous orientation, and a rash of literature sprang from this new psychological "typology," referring to androgyny as an "ideal psychological state" (e.g., Gilbert, 1981; Kaplan, 1976; Osofsky & Osofsky, 1972; Rebecca, Hefner, & Oleshansky, 1976; and Sturdevant, 1980, to mention a few). Given this enthusiasm, it is curious that (with respect to the short BSRI and the EPAQ) no empirical evidence has been provided to justify adding androgyny as a variable to the masculinity–femininity domain, in other words, evidence substantiating the construct validity of androgyny with respect to psychological well-being.

In the most recent theoretical discussion of androgyny (Bem, 1981a, 1981b; Spence & Helmreich, 1981), no mention was made regarding how the concept of androgyny re-

lates to psychological health. Yet, Bem (1981b) writes, "Androgyny was a concept whose time had come, a concept that appeared to provide a liberated and more humane alternative to the traditional, sex-biased standards of mental health" (p. 362). In view of these considerations, the main focus of this article is an attempt to tie down the concept of androgyny empirically and to determine the relation between this popular notion (as measured by the short BSRI and the EPAQ) and several indexes of psychological well-being.

It should be noted that Spence and Helmreich have never claimed that their positively valued EPAQ measures ($M^+$ and $F^+$) should be used to construct a new model of mental health; in fact, they have warned investigators against such extrapolations (Spence & Helmreich, 1979b). Spence and Helmreich (1979b) view their $M^+$ and $F^+$ scales as facets of the more global domain of masculinity and femininity that Bem (1979b) speaks of, namely, instrumental ($M^+$) and expressive ($F^+$) personality traits, respectively. With this in mind, Spence and Helmreich (1979b) have gone on to say that "the relation between androgyny in this personality sense and other types of androgyny . . . [e.g., Bem's, 1974, 1979a] must be determined empirically" (p. 1035). Therefore, most of our discussion of androgyny as a "new model of mental health" is in response to Bem's (1974, 1979a) theoretical considerations.

Spence and Helmreich have, however, investigated the relation between their $M^+$ and $F^+$ measures and indexes of psychological health in normal and clinical populations (e.g., Holahan & Spence, 1980; Spence et al., 1979). Typically, these studies (in concordance with the findings of Antill & Cunningham, 1979) reflect a substantial degree of construct validity for masculinity as an indicator of psychological well-being but little or no support for femininity as such an indicator; moreover, in none of these studies were attempts made to assess the possible emergence of androgyny as an important distinctive dimension.

Therefore, because the overall tenor of the sex role literature suggests that androgyny is an ideal psychological state, we have chosen to inquire into the construct validity of both the short BSRI and the EPAQ with respect to subjective indicators of psychological well-being, even though the theoretical considerations of the authors of the EPAQ are more conservative than Bem's. Furthermore, as we shall see, inclusion of the EPAQ will allow us to examine some of Spence and Helmreich's own distinctive ideas concerning androgyny.[1]

The specific purpose of the present study was to gain a better understanding of the short BSRI and the EPAQ by comparing and contrasting these two instruments to one another and to a more comprehensive inventory of self-view measures—the Differential Personality Questionnaire (DPQ; Tellegen, Note 1). Particular attention is devoted to evaluating the construct validity of masculinity (M), femininity (F), and androgyny (as well as traditional notions of sex role identification) in relation to measures of subjective psychological well-being. In order to assess the utility of androgyny as a concept distinct from F and M, it will be operationalized and evaluated as an intrinsically interactive variable (rather than as a simply additive composite of F and M).

Androgyny was first cast as an interactive concept by Lubinski et al. (1981) using the short BSRI. Our rationale was twofold. First, Bem has suggested that even though M and F are both related to adaptiveness and psychological health, extreme degrees of M and F in isolation become negative and even destructive, whereas in combination, M and F temper each other, one neutralizing the more negative manifestations of the other. We argued that this tempering implies an interaction and can be assessed as such in an analysis of variance (ANOVA) or, as we prefer, a regression model. Second, and more important, if androgyny is to achieve the status of an empirically useful concept (regardless of the domain of behaviors to which it is hypothesized to be related), it must have unique predictive properties relative to masculinity and femininity. Lacking such "surplus" properties,

[1] Although the EPAQ contains six masculinity–femininity measures, when mentioning the EPAQ masculinity and femininity measures, we will always be referring, unless we specify otherwise, to their positively valued ($M^+$ and $F^+$) measures because these two scales are employed to construct the fourfold typology, which is of primary interest to us.

androgyny would be predictively and conceptually redundant, and it would be enough to interpret findings using the concepts of M and F.[2]

Our earlier evaluation of androgyny as an interactive concept produced disappointing results (Lubinski et al., 1981). None of the tested M × F interactions contributed significantly to the prediction of relevant psychological criteria. In the present study, however, we investigate further the utility of two well-known contemporary formulations of androgyny.

By employing the entire DPQ with its particular emphasis on markers of positive and negative affectivity, we hoped to determine the extent to which androgyny (as measured by the short BSRI and the EPAQ) may be viewed as an indicator of psychological health. This would require that the appropriate M × F interaction share variance with these indicators. A second extension of our earlier study is to evaluate Spence et al.'s (1979) new measures: "unmitigated agency" (M⁻) and "unmitigated communion" ($F_C^-$).

Our basic approach to data analysis was to conduct hierarchical multiple regression analyses with interaction terms, using the short BSRI and the EPAQ separately as predictors of seven measures of subjective well-being. Our regression equation is as follows:

$$\hat{Y} = B_1M + B_2F + B_3S + B_4MF$$
$$+ B_5MS + B_6FS + A. \quad (1)$$

It contains the terms that will enable us to test concurrently (a) the construct validity of masculinity (M) and femininity (F) as linear predictors of psychological well-being (Y), namely, $B_1M$ and $B_2F$; (b) the validity of the concept of androgyny conceived as an intrinsically interactive (rather than simply additive) concept, namely, $B_4MF$; and (c) the validity of the traditional assumption that it is the masculine man and feminine woman who typify psychological well-being, namely, $B_5MS$ and $B_6FS$, respectively, where S refers to gender.[3] (For further clarification regarding the specific theoretical significance of each of these interactions, see Lubinski et al., 1981.)

Although our rationale for operationalizing Bem's androgyny as an interactive con-

cept seems clear (cf. Lubinski et al., 1981, pp. 723–724), readers may wonder why we are placing this requirement on the EPAQ, because Spence and Helmreich (e.g., 1978, p. 109) appear to reject the surplus meaning that has accrued to androgyny. In fact, Spence and Helmreich (1979b) state explicitly,

We introduced the term *androgyny* simply as a convenient label to identify individuals who score relatively high on both the M and the F scales of our particular instrument. . . . Androgyny, as we have used the term, has no particular theoretical import, being intended to indicate nothing more than a relatively high degree of both instrumental and expressive personality traits, as defined by the PAQ. (p. 1035)

It would seem, therefore, that Spence and Helmreich view androgyny as simply an additive composite of M and F. However, in some of their discussions at least (Spence et al., 1979, pp. 1674–1675), they appear to advance an androgyny concept that is not simply additive but interactive, that is, distinct from M and F.

---

[2] Interestingly, our operational definition of androgyny is similar to Kaplan and Bean's (1976) conceptual definition. They contend that "androgyny . . . is *not* a simple or even complex union of the totalities of traditionally defined masculinity and femininity. It's a third dimension" (p. 383). By operationalizing androgyny as a partial product, and entering the term representing it (MF) as an independent variable, androgyny becomes a third dimension distinct from masculinity and femininity. This, among other things, enables us to assess the construct validity of androgyny empirically, both convergently and discriminantly (e.g., Campbell & Fiske, 1959; Cronbach & Meehl, 1955). This approach also resolves the problem noted by Spence and Helmreich (1981), namely, that of anchoring a variable empirically to a pair of relatively independent measures.

[3] Our interpretation of these interactions differs drastically from that of Flaherty and Dusek (1980, p. 991). These authors interpret their significant F × S interaction in conjunction with a significant M main effect as indicating that androgyny exists in females but not males. We disagree. Flaherty and Dusek's F × S interaction illustrates what we have called the traditional notion that femininity is adaptive in females but not in males. This type of interaction, however, should be distinguished from those that corroborate the idea of androgyny. To argue that androgyny operates in one gender but not the other amounts to hypothesizing that androgyny itself is moderated by gender in relation to relevant criteria. For its empirical evaluation, this hypothesis would call for adding the product term $B_7MFS$ as an independent variable in the appropriate hierarchical analysis (Cohen, 1968). This three-way interaction was computed for all dependent measures employed in this study for both the EPAQ and the short BSRI and in each case was found not to be significant.

Furthermore, Spence and Helmreich have recently constructed measures of unmitigated agency and communion that appear to reflect an interactive concept ($M^-$ and $F_C^-$, respectively; see Spence et al., 1979). The motivation for the construction of these scales was provided by Bakan (1966). According to Spence et al. (1979, pp. 1674–1675),

He [Bakan, 1966] proposes that a strong sense of agency, unmitigated by a sense of communion, is destructive to the individual and to society. Similarly, communion must be mitigated by agency if the individual is to function effectively. The developmental task of males is thus to learn to balance masculine agency with some degree of communion, and of females, to balance feminine communion with some degree of agency. . . . In expanding our instrument to include the negatively toned masculine and feminine attributes, it was our intent to develop scales that conceptually parallel the [positively valued] M and F scales, containing negatively valued masculine characteristics that reflect Bakan's "unmitigated agency" and negatively valued feminine characteristics that reflect "unmitigated communion."[4]

We believe that the idea of an $M \times F$ interaction is implied by these considerations. Were we not to regard mitigation as an interactive concept, then the extent to which M contributes to predicting relevant psychological criteria would not change as a function of F, and vice versa. But in order to say that the tendencies associated with high standings on M and F mitigate one another—supposedly allowing full expression of positive masculine and feminine attributes but inhibiting the more negative expressions that are present when one has high standing on only M or F—we would have to posit an interaction. Without it, the very ideas of unmitigated agency and communion, along with their measures, $F_C^-$ and $M^-$, would be meaningless. In fact, to be valid, $F_C^-$ and $M^-$ should behave as measures of mitigation in reverse; that is, they should show a clear and negative correlation with the interactive integration of F and M, assessed directly with the multiple regression procedures described in Lubinski et al. (1981).

Specifically, the validity of $F_C^-$ and $M^-$ may be assessed by computing the partial correlations between $M^-$ and $F_C^-$ with the appropriate $M \times F$ interaction, after controlling for M, F, and sex of subject.[5] If $M^-$ (reversed) and $F_C^-$ (reversed) "tap" mitiga-

tion, then we would expect to find an appreciable *negative* partial correlation between these measures and the $M \times F$ interaction variable.

## Method

### Subjects

A total of 172 college students (87 men and 85 women) were recruited from a number of introductory psychology classes at the University of Minnesota. Only students under 30 years of age were included in the sample.

### Procedure

In groups of about 85, students were administered a test battery consisting of three personality inventories: the short Bem Sex-Role Inventory, the Extended Personal Attributes Questionnaire, and the Differential Personality Questionnaire (Tellegen, Note 1). A brief description of the three instruments follows:

*The short BSRI.* The short Bem Sex-Role Inventory (Bem, 1978, 1979b) consists of two 10-item scales—Masculinity and Femininity—and 10 filler items. The BSRI classifies individuals, using a median-split method, as androgynous (high M, high F), masculine (high M, low F), feminine (low M, high F), and undifferentiated (low M, low F).

*The EPAQ.* The Extended Personal Attributes Questionnaire consists of six (three positive and three negative) masculinity–femininity scales. The four scales analyzed in this study may be described as follows: Masculinity ($M^+$) is designed to measure a facet of positively valued masculinity, namely, instrumentality; femininity ($F^+$) is designed to measure a facet of positively valued femininity, namely, expressiveness (the EPAQ $F^+$ and $M^+$ scales are the old PAQ F and M scales); negative masculinity ($M^-$) is intended to reflect unmitigated agency; and negative femininity ($F_C^-$) is intended to reflect unmitigated communion. The latter two are measures of negatively valued characteristics specifically designed to be counterparts to the former two (see Spence et al., 1979, particularly pp. 1674–1675). Like Bem, Spence et al. (1979) use the median-split method, on their positively valued measures ($M^+$ and $F^+$), to classify individuals as masculine, feminine, androgynous, or undifferentiated.

As described in some detail in Lubinski et al. (1981), subjective well-being involves two broad and relatively

---

[4] The concept of unmitigated communion is an extension of Bakan's (1966) ideas introduced by Spence, Helmreich, and Holahan (1979). Bakan himself speaks only of unmitigated agency not unmitigated communion.

[5] For both of these analyses, we use a partial correlation because the critical $M \times F$ interaction can only be captured after partialing out its constituents. Also, gender was added as a third main effect to circumvent the possibility of the $M \times F$ interaction's absorbing some of the variance due to sex differences, which would compromise its interpretability.

independent dimensions, namely, positive and negative affectivity (see Costa & McCrae, 1980, for a similar view). In this study, factor scores representing these two higher order dimensions and derived from the Differential Personality Questionnaire (DPQ) are our primary criteria.

*The DPQ.* The 11 primary scales of the DPQ were developed factor analytically in a series of studies aimed at converging, through a process of successive approximations, on major self-view dimensions (Lubinski et al., 1981; Tellegen, 1981, Note 1). Jointly, the 11 scales mark three higher order factors that can be interpreted as broad self-view dimensions, namely, Positive Affectivity, Negative Affectivity, and Constraint. Of the primary scales, Well-Being, Social Potency, and Achievement are primarily associated with Positive Affectivity; Stress Reaction, Alienation, and Aggression, with Negative Affectivity; and Control, Authoritarianism, Harm Avoidance, and Aggression (negatively), with Constraint. These higher order dimensions can also be discerned in other broad-range personality inventories, but from an interpretive point of view, they emerge especially clearly from the primary DPQ scales. Factor scores representing all three higher order dimensions were computed.

Students were given as much time as they needed to complete the questionnaires and spent, on the average, approximately 55 minutes on the task.

## Results

The results are reported in three parts. First, in order to determine the extent to which corresponding scales of the short BSRI and the EPAQ may be considered conceptually equivalent, a multitrait–multimethod matrix is presented and discussed. Second, the correlational profiles between these scales are also inspected to assess, in the context of all 11 DPQ measures, the amount of *extrinsic convergent validity* shared by these two inventories. These profiles further define masculinity and femininity through *construct explication*. Third, multiple regression analyses are reported, using first the short BSRI and then the EPAQ to predict seven measures of subjective psychological health. The dependent measures chosen for this analysis are two factor scores—Positive Affectivity and Negative Affectivity—and the major first-order markers of these higher order dimensions, namely, Well-Being, Social Potency, and Achievement for Positive Affectivity and Stress Reaction and Alienation for Negative Affectivity. Finally, the view of the EPAQ scales ($M^-$ and $F_C^-$) as measures of unmitigated agency and communion, respectively, is tested using partial correlations.

The top of Table 1 presents a multitrait–multimethod matrix (Campbell & Fiske, 1959), organized around traits rather than methods. Four coefficient alpha reliability estimates (Cronbach, 1951) were computed, two in each of the heteromethod–monotrait triangles. Although all of the reliability estimates were encouragingly high, the short BSRI scales appeared more reliable than the EPAQ measures (BSRI–M = .85, BSRI–F = .89, EPAQ–$M^+$ = .77, EPAQ–$F^+$; = .77). The two validity coefficient estimates (femininity = .75, masculinity = .72) were also impressive (in fact, when corrected for attenuation, these two coefficients rose to .89 and .90 for masculinity and femininity, respectively). In addition, discriminant validity is also displayed in that the heterotrait–monomethod entries (.29 and .16; $M$ = .23) were significantly smaller than the convergent validity coefficients.

The lower part of Table 1 contains two extrinsic convergent validation rectangles, an extension of the multitrait–multimethod matrix proposed by Fiske (1971). According to Fiske (1971, p. 245), "as a rule, measures of the same construct are not interchangeable empirically and must not be considered conceptually equivalent until a high degree of convergence in correlational patterns with other variables, or extrinsic convergent validation, has been demonstrated empirically." The two extrinsic convergent validation rectangles compare the correlational profiles of the four masculinity–femininity measures with all 11 scales of the DPQ and its three factor scores. Inspection of these immediately reveals the substantial degree of overlap of these two inventories. One may, in fact, conclude that their corresponding measures are conceptually near equivalent.

The two rectangles (in Table 1) also contribute to the construct explication of masculinity and femininity by allowing us to define these M–F measures in the context of a fairly comprehensive set of major self-view dimensions. (Because the amount of extrinsic convergent validation for these two inventories was high, the corresponding M–F scales can be given the same interpretations.)

Measures of masculinity seem to be primarily associated with markers of positive affectivity and may be defined, specifically, as indicators of subjective well-being, social potency, achievement, and stress reaction

(reversed). Thus, high masculinity scores reflect a view of oneself as interpersonally effective and dominant, whereas absence of such self-attributions is associated with low scores on masculinity. Indeed, an inspection of the content of the masculinity items indicates that the domain sampled could quite adequately be labeled dominance–poise (e.g., short BSRI–M: "dominant," "forceful," "strong personality"; EPAQ–M+: "feels superior," "self-confident," "stands up well under pressure," "competitive"). Therefore, some association between these masculinity measures, and markers of positive affectivity, is expected. The femininity measures, on the other hand, tap a different region of the self-view domain and may be defined as markers of well-being, social closeness, and aggression (reversed). The association of femininity with these three measures indicates that, unlike

masculinity, femininity is primarily related to "nurturance and accommodating warmth" (Lubinski et al., 1981). The item content of the two femininity scales also seems to warrant this conclusion (e.g., short BSRI–F: "warm," "understanding," "eager to sooth hurt feelings," "sympathetic"; EPAQ–F+: "understanding of others," "warm in relation with others," "helpful to others," "able to devote self completely to others").

It appears, then, that the attempt at construct explication is illuminating. In the context of 11 self-view dimensions, the measures in question—masculinity and femininity—seem to share most of their variance with two quite distinct subsets.

Turning now to the multiple regression analyses, recall that these were performed in order to evaluate interactive relations between the F and M scales in relation to spe-

Table 1

*A Multitrait–Multimethod Matrix Between Corresponding M–F Measures of the Short BSRI and the EPAQ Followed by Two Extrinsic Convergent Validation Rectangles*

| Measure | Masculinity | | Femininity | |
|---|---|---|---|---|
| | BSRI–M | EPAQ–M | BSRI–F | EPAQ–F |
| BSRI–M | (.85) | | | |
| EPAQ–M | .72 | (.77) | | |
| BSRI–F | .29 | .25 | (.89) | |
| EPAQ–F | .14 | .16 | .75 | (.77) |
| DPQ | | | | |
| Well-Being | .33 | .46 | .33 | .30 |
| Social Potency | .63 | .55 | .10 | .13 |
| Achievement | .36 | .40 | .15 | .12 |
| Social Closeness | .21 | .20 | .42 | .45 |
| Stress Reaction | −.42 | −.53 | −.17 | −.07 |
| Alienation | −.14 | −.11 | −.26 | −.28 |
| Aggression | .03 | .02 | −.39 | −.40 |
| Control | .03 | .04 | .22 | .15 |
| Harm Avoidance | −.14 | −.14 | .19 | .20 |
| Traditionalism | −.06 | .11 | .29 | .27 |
| Absorption | −.02 | −.03 | .26 | .31 |
| Positive Affectivity | .53 | .59 | .36 | .35 |
| Negative Affectivity | −.21 | −.24 | −.24 | −.20 |
| Constraint | −.10 | −.01 | .30 | .26 |

*Note.* Each heteromethod–monotrait triangle is enclosed by a solid line. Both triangles contain an italicized validity coefficient and two alpha coefficients (in parentheses). Below the multitrait–multimethod matrix are two extrinsic convergent validity rectangles. For all correlations above, $r_{.05} = .13$, $r_{.01} = .18$, $r_{.001} = .23$. BSRI = the short Bem Sex-Role Inventory. EPAQ = Extended Personality Attributes Questionnaire. DPQ = Differential Personality Questionnaire.

cific dependent variables. The results, reported in Table 2, were obtained by first entering the three main effects (masculinity, femininity, and gender) in an incremental stepwise fashion and then entering the three interaction terms (M × F, M × S, F × S) using the same procedure. Results with the short BSRI are reported first.[6]

### Positive Affectivity

Masculinity accounted for 28% of the variance, and when femininity was added, the amount of variance accounted for increased to 33% ($R = .57$). None of the remaining variables increased the $R^2$ significantly.

### Negative Affectivity

Femininity accounted for 6% of the variance, and when masculinity was added, the amount of variance accounted for increased to 8%. The only other significant entry is the (M × F) interaction, increasing the variance accounted for to 10% ($R = .32$), but the direction of the partial correlation was opposite to what was expected ($r_{Y4.123} = .16$).

The same multiple regression analysis was conducted using the EPAQ masculinity and femininity measures.

### Positive Affectivity

Masculinity was entered first, accounting for 35% of the variance. When femininity was added, the amount of variance accounted for increased to 41% ($R = .64$).

### Negative Affectivity

Masculinity was entered first, accounting for 6% of the variance. When femininity was added, the amount of variance accounted for increased to 9%. One other variable, the (M × F) interaction term, accounted for a significant portion of the variance. However, the interaction again was in the wrong direction ($r_{Y4.123} = 18$).

To scrutinize these data further, the same type of regression analysis was performed for both the short BSRI and the EPAQ, using the major markers of Positive and Negative Affectivity separately as criterion variables (with Well-Being, Social Potency, and Achievement primarily defining Positive Affectivity and Stress Reaction and Alienation defining Neg-

ative Affectivity). As expected, the results of these analyses for the main effects correspond closely to those reported in Table 1, and no detailed discussion is presented. To be sure, in some of these analyses the (M × F) interaction term shared a significant portion of variance with the criterion. However, as in the case of Negative Affectivity, all of these significant interactions were in the wrong direction. (The only significant M × F interaction effects involved the following scales: For the short BSRI, Well-Being, $R^2$ change = .03, $p < .05$; Alienation, $R^2$ change = .07, $p < .001$. For the EPAQ, Alienation, $R^2$ change = .09, $p < .001$.)[7]

---

[6] Before reporting the results of our regression analyses, the three hypotheses about to be tested, along with their empirical requirements for corroboration, should be discussed further. For both inventories, measures of M and F must, in addition to accounting for a significant amount of the criterion variance, be in the predicted direction. This stipulation also pertains to the partial correlations as they are entered in the stepwise procedure. (This is an important requirement because even if the simple correlations of M and F with a measure of psychological health are positive, one of the predictors may become negatively correlated with the criterion when the other predictor is controlled.) The same logic applies to the M × F interaction term; it too must represent a relation that is in the predicted direction. Thus, M, F, and M × F should have positive partial correlations with Positive Affectivity and negative partial correlations with Negative Affectivity. The final hypothesis to be tested is the traditional notion that the masculine man and feminine woman are the ones who typify psychological health. This hypothesis will receive empirical support (gender is coded: females = 2, males = 1) if the F × S partial is positively correlated with Positive Affectivity and the M × S partial is negatively correlated with Positive Affectivity, whereas the inverse of these directions would be expected for Negative Affectivity. Here, too, the direction of these interactions may be evaluated by means of partial correlations. On the other hand, under the traditional view, the M × F interaction should not be significant. (The partial correlations discussed above are given in Table 2 under the column labeled Partial r.)

[7] We recognize that the linear-by-linear interaction tested by us is only one of several possible patternings of masculinity and femininity. However, this form of interaction appears to capture the current interaction conception of androgyny (i.e., M and F "temper" [Bem, 1979a] or "mitigate" [Spence et al., 1979] each other). An alternative idea, one discussed in Bem's (1974) earlier work, namely, the "balance model," would require a different analysis in which the discrepancy between M and F (e.g., |M − F|) would be evaluated as an independent variable. Although apparently not currently advocated, this alternative model was also tested for each of the eight dependent measures discussed above using both the short BSRI and the EPAQ with uniformly negative results.

In our final analysis, we assessed the validity of M− and $F_C^-$ as measures of unmitigated agency and communion, respectively. Partial correlations showed that M− (or unmitigated agency) is essentially independent of "mitigation," or the M × F interaction (i.e., $r_{AM-.M,F,S}$ = .05, where A stands for the "androgynous" M × F variable). $F_C^-$ displayed a low correlation with mitigation; however, this correlation indicated a positive relation between maladaptive femininity and androgyny (i.e., $r_{AF_C-.M,F,S}$ = .26, $p < .001$).

## Discussion

The purpose of this study was to collect and evaluate evidence that would (a) determine the extent to which corresponding scales of the short BSRI and the EPAQ may be considered empirically near equivalent; (b) clarify the nature of masculinity and femininity and androgyny in the context of several general self-view dimensions, particularly in reference to subjective indicators of psychological well-being, possibly in interaction with the individual's gender; and (c)

assess the utility and meaning of Spence and Helmreich's new measures, unmitigated agency (M−) and unmitigated communion ($F_C^-$).

In our first analysis, a substantial degree of overlap was found between corresponding scales of the short BSRI and the EPAQ. They showed an impressive convergent and discriminant pattern when placed in a multitrait–multimethod matrix. An even more striking display of similarity was obtained by placing pairs of corresponding scales in two extrinsic convergent validation rectangles. Our findings suggest that researchers should be able to use these inventories interchangeably.

Construct explication shows that masculinity and femininity, as measured by the short BSRI and the EPAQ, have an affinity to different groups of self-view dimensions. The correlational profile for femininity defines it as being primarily related to a cluster of variables that can be labeled as "nurturance–warmth." Masculinity, on the other hand, was defined by a cluster of variables reflecting "dominance–poise."

Table 2

*Regressions of Masculinity, Femininity, Gender, and All Possible Pairwise Interactions on Positive Affectivity and Negative Affectivity*

| Step | Positive affectivity | | | | Negative affectivity | | | |
|---|---|---|---|---|---|---|---|---|
| | Variable | Partial $r$ | $R$ | $R^2$ change | Variable | Partial $r$ | $R$ | $R^2$ change |
| | BSRI | | | | | | | |
| 1 | M | .53 | .53 | .28*** | F | −.24 | .24 | .06** |
| 2 | F | .25 | .57 | .05*** | M | −.15 | .28 | .02* |
| 3 | S | −.02 | .57 | .00 | S | −.02 | .29 | .00 |
| 4 | M × F | −.09 | .58 | .01 | M × F | .16 | .32 | .02* |
| 5 | M × S | −.04 | .58 | .00 | M × S | −.06 | .32 | .00 |
| 6 | F × S | .01 | .58 | .00 | F × S | .01 | .32 | .00 |
| | EPAQ | | | | | | | |
| 1 | M | .59 | .59 | .35*** | M | −.24 | .24 | .06*** |
| 2 | F | .32 | .64 | .07*** | F | −.17 | .29 | .03** |
| 3 | S | −.05 | .64 | .00 | S | −.03 | .30 | .00 |
| 4 | M × S | −.07 | .65 | .00 | M × F | .18 | .34 | .03* |
| 5 | F × S | .03 | .65 | .00 | F × S | −.07 | .35 | .00 |
| 6 | M × F | .01 | .65 | .00 | M × S | .01 | .35 | .00 |

*Note.* These data were obtained by entering the three main effects in an incremental stepwise fashion and then entering the three interaction terms using the same procedure. The column labeled Partial $r$ represents the partial correlation for that variable and the criterion variable after controlling for all of the preceding variables in the equation. Thus, the first entry in every column is merely the simple correlation. $N = 172$. M = masculinity; F = femininity; S = gender. BSRI = the short Bem Sex-Role Inventory. EPAQ = Extended Personality Attributes Questionnaire.
* $p < .05$. ** $p < .01$. *** $p < .001$.

This result is consistent with our earlier findings using the short BSRI and other studies using the BSRI "old version" (e.g., Bernard, 1980; DeGregorio & Carver, 1980; Wiggins, & Holzmuller, 1978). Moreover, Antill and Cunningham (1979) and Spence et al. (1979) showed that $M^+$ shares much variance with markers of Positive Affectivity, whereas the relation between $F^+$ and these markers was quite weak. (For a brief summary of the above studies, see Lubinski et al., 1981.)

The masculinity scales of both inventories in fact received a good deal of construct validation in view of their pattern of positive relations to markers of Positive Affectivity (i.e., Well-Being, Social Potency, and Achievement) and negative relation to Stress Reaction. Femininity did not fare as well, showing a corroborative but less impressive pattern of relations with indicators of Well-Being, Social Closeness, and Alienation. In the broad context of these personality dimensions, a fairly consistent definition of masculinity and femininity emerges. Masculinity appears to be primarily characterized by a dominant and "assimilative" style, and femininity by qualities of nurturance and warmth, an "accomodative" style. (Consistent with this interpretation are the recent findings of Harrington and Andersen, 1981, who report a substantial relation between masculinity and several measures of creativity but negative or negligible correlations between femininity and these measures.) However, in view of our regression analyses, it does not appear that in relation to indicators of subjective well-being the combination of these two distinct and relatively independent dimensions interact synergistically to produce an important third dimension of androgyny.

Our failure to "triangulate" androgyny as a significant distinct concept was disappointing. Each time androgyny appeared to emerge with predictive properties distinct from masculinity and femininity, its association with the criterion was in the wrong direction. These findings are, of course, mostly relevant to Bem's ideas because Spence and Helmreich do not share Bem's view that androgyny epitomizes psychological health. However, similarity of the results obtained with the EPAQ increases the generalizability of our findings and indicates that claims for androgyny, however measured, as a constellation specifically indicative of psychological health so far have not received empirical support.

Not only were the relations found in this study in the wrong direction but the evidence was also negative in our first study (Lubinski et al., 1981). Recognizing that one disconfirmation is generally more decisive than one corroboration in evaluating the verisimilitude of a theory (Meehl, 1979), androgyny seems to be in some difficulty, at least with respect to subjective psychological well-being.

On the whole, and in the light of the findings reported above, the fourfold typology described and advocated by Bem (but not Spence and Helmreich) for classifying individuals with respect to psychological health seems to have little utility. Furthermore, the results clearly indicate that the suggestion to develop psychotherapeutic techniques within this framework is unwarranted (e.g., Kaplan & Bean, 1976). For the same reasons, we are skeptical about the heuristic merits of a classification of counselor personalities based on the fourfold typology (e.g., Highlen & Russell, 1980). Such a classification, at the very least, is premature because there is simply no evidence to warrant the expansion of masculinity and femininity into a functional typology at this point. Moreover, in view of the nonsignificance of the M × S and F × S interaction terms, we conclude that with respect to subjective well-being, neither the contemporary formulations of androgyny nor the traditional assumptions regarding sex role identification are accurate.

If traditional formulations of sex role identification are inaccurate and androgyny as an interactive variable does not share variance with indicators of psychological well-being, can anything of a prescriptive nature be concluded from the contemporary sex role literature? The most relevant and provocative findings are the strong relation or overlap of instrumentality (or positively valued masculinity as assessed by the short BSRI–M and EPAQ–$M^+$ scales) to indicators of psychological well-being and the fact that the strength of this relation is the same in men and women.

This may suggest the straightforward conclusion that "masculinity" in the sense of dominance–poise should be encouraged and reinforced in women and men alike. Moreover, the class of "feminine" behaviors, such as nurturance and warmth (as assessed by the short BSRI-F and the EPAQ-F$^+$), may be cultivated independently of dominance–poise. But the available evidence indicates that nurturant qualities are not as strongly associated with a sense of well-being as dominant qualities are.

On the other hand, there is no evidence that with respect to subjectve well-being, masculinity and femininity become dysfunctional when an individual has a high standing on one variable but not on the other. As assessed by the short BSRI and the EPAQ (i.e., M$^+$ and F$^+$) scales, masculinity and femininity do not temper (Bem, 1979a) or mitigate (Spence et al., 1979) one another to enhance an individual's overall subjective well-being.

The possibility remains, of course, that M and F interact in relation to a more restricted range of variables than would be required by the more encompassing concept of androgyny advanced by Bem. For example, Spence and Helmreich (1979a) report what appears to be an M × F interaction in a study of attitudes toward women. Harrington and Andersen also report some significant M × F interactions in the prediction of certain indexes of creativity. (However, this finding is offset by the negative correlations of femininity with these measures of creativity, which are substantial compared to the variance accounted for by the M × F interaction.)[8]

Because no evidence was found indicating that the single M$^-$ (reversed) and Fc$^-$ (reversed) scales measure the mitigation of positively valued masculinity and femininity, we suggest that researchers who wish to determine the extent to which masculinity and femininity mitigate one another, or assess androgyny, continue the study of F × M interactions, using the multiple regression procedure discussed in this study.

One attractive feature of this technique, not illustrated in the present article, is that it enables investigators to determine whether masculinity or femininity function as moderator variables even when not correlated

with the criterion under consideration. For example, although most studies indicate that femininity shares little variance with measures of psychological health, femininity may have moderated relationships with masculinity to produce a significant M × F interaction (Saunders, 1956; Zedeck, 1971). Thus, although neither masculinity nor femininity seems to function as a moderator variable with respect to subjective indicators of psychological well-being, the possibility remains an empirical question for other relevant psychological criteria. Also, by adding the product term M × F × S as an independent variable, researchers may determine whether androgyny (i.e., M × F) operates in one gender but not in the other.

Finally, we suggest that the multiple regression techniques discussed in this article be viewed as a general data-analytic device for investigating the construct validity of those higher order psychological concepts that are defined by nonadditive combinations of certain first-order personal characteristics. (See Heilbrun, 1978, p. 106, and Diamond, Royce, & Voorhees, 1981, p. 188, for two contemporary theoretical formulations that might benefit from this type of analysis.) The surplus meaning of these higher order concepts may be captured empirically in a regression analysis by entering their components and their relevant combinations in the proper sequence. For example, we have suggested that androgyny is only one example of a broader class of "fulfillment" or "self-actualization" concepts (Lubinski et al., 1981). All such concepts appear to imply that a "fully functioning" person has integrated seemingly contradictory or unrelated qualities as synergistic constituents of an organized whole that consequently "is more than the sum of its parts." We again suggest that this model can also be used to assess other higher order concepts (fulfillment or otherwise) that imply nonadditive combinations of first-order components.

---

[8] Stokes, Childs, and Fuehrer (1981) also recently reported a significant M × F interaction in the prediction of "disclosure to intimates." However, instead of entering the product term, M × F, in the appropriate hierarchical fashion, they entered it first in their stepwise regression. Consequently, the results of their analysis are uninterpretable (cf. Lubinski, 1983).

## Reference Note

1. Tellegen, A. *Brief manual for the Differential Personality Questionnaire.* Unpublished manuscript, University of Minnesota, 1982.

## References

Antill, J. K., & Cunningham, J. D. Self-esteem as a function of masculinity in both sexes. *Journal of Consulting and Clinical Psychology,* 1979, *47,* 783–785.

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.

Bem, S. L. On the utility of alternative procedures for assessing psychological androgyny. *Journal of Consulting and Clinical Psychology,* 1977, *45,* 196–205.

Bem, S. L. *The short Bem Sex-Role Inventory.* Palo Alto, Calif.: Consulting Psychologists Press, 1978.

Bem, S. L. Beyond androgyny: Some presumptuous prescriptions for a liberated sexual identity. In J. Sherman & F. Denmark (Eds.), *Psychology of women: Issues in psychology.* New York: Psychological Dimensions, 1979. (a)

Bem, S. L. Theory and measurement of androgyny: A reply to the Pedhazur–Tetenbaum and Locksley–Colten critiques. *Journal of Personality and Social Psychology,* 1979, *37,* 1047–1054. (b)

Bem, S. L. The BSRI and gender schema theory: A reply to Spence and Helmreich. *Psychological Review,* 1981, *88,* 369–371. (a)

Bem, S. L. Gender schema theory: A cognitive account of sex typing. *Psychological Review,* 1981, *88,* 354–364. (b)

Bernard, L. C. Multivariate analysis of new sex role formulations and personality. *Journal of Personality and Social Psychology,* 1980, *38,* 323–336.

Campbell, D. T., & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin,* 1959, *56,* 81–105.

Cohen, J. Multiple regression as a general data-analytic system. *Psychological Bulletin,* 1968, *70,* 426–443.

Costa, P. T., & McCrae, R. R. Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology,* 1980, *38,* 668–678.

Cronbach, L. F. Coefficient alpha and the internal structure of tests. *Psychometrika,* 1951, *16,* 297–334.

Cronbach, L. F., & Meehl, P. E. Construct validity in psychological tests. *Psychological Bulletin,* 1955, *52,* 281–302.

DeGregorio, E., & Carver, C. S. Type A behavior pattern, sex role orientation, and psychological adjustment. *Journal of Personality and Social Psychology,* 1980, *39,* 286–293.

Diamond, S. R., Royce, J. R., & Voorhees, B. The factor "fulfillment model" and the concept of multivariate personality type. *Personality and Individual Differences,* 1981, *2,* 181–189.

Fiske, D. W. *Measuring the concepts of personality.* Chicago: Aldine-Atherton, 1971.

Flaherty, J. F., & Dusek, J. B. An investigation of the relationships between psychological androgyny and components of self-concept. *Journal of Personality and Social Psychology,* 1980, *38,* 984–992.

Gilbert, L. A. Toward mental health: The benefits of psychological androgyny. *Professional Psychology,* 1981, *12,* 29–38.

Harrington, D. M., & Andersen, S. M. Creativity, masculinity, femininity, and three models of psychological androgyny. *Journal of Personality and Social Psychology,* 1981, *41,* 744–757.

Heilbrun, A. B. An exploration of antecedents and attributes of androgynous and undifferentiated sex roles. *Journal of Genetic Psychology,* 1978, *132,* 97–107.

Highlen, P. S., & Russell, B. Effects of counselor gender and counselor and client sex role on females' counselor preference. *Journal of Counseling Psychology,* 1980, *27,* 157–165.

Holahan, C. K., & Spence, J. T. Desirability and undesirable masculine and feminine traits in counseling clients and unselected students. *Journal of Consulting and Clinical Psychology,* 1980, *48,* 300–302.

Kaplan, A. G., Androgyny as a model of mental health for women: From theory to therapy. In A. G. Kaplan & J. P. Bean (Eds.), *Beyond sex-role stereotypes: Readings toward a psychology of androgyny.* Boston: Little, Brown, 1976.

Kaplan, A. G., & Bean, J. P. (Eds.). *Beyond sex-role stereotypes: Readings toward a psychology of androgyny.* Boston: Little, Brown, 1976.

Lubinski, D. The androgyny dimension: A comment on Stokes, Childs, and Fuehrer. *Journal of Counseling Psychology,* 1983, *30,* 130–133.

Lubinski, D., Tellegen, A., & Butcher, J. N. The relationship between androgyny and subjective indicators of emotional well-being. *Journal of Personality and Social Psychology,* 1981, *40,* 722–730.

Meehl, P. E. Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology,* 1979, *46,* 806–834.

Osofsky, J. D., & Osofsky, H. J. Androgyny as a life style. *Family Coordinator,* 1972, *21,* 411–419.

Rawlings, E. I., & Carter, D. K. *Psychotherapy for women: Toward equality.* Chicago: Charles C Thomas, 1977.

Rebecca, M., Hefner, R., & Oleshansky, B. A model of sex-role transcendence. *Journal of Social Issues,* 1976, *32,* 197–206.

Saunders, D. R. Moderator variables in prediction. *Educational and Psychological Measurement,* 1956, *16,* 209–222.

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. On assessing "androgyny." *Sex Roles,* 1979, *5,* 721–738. (a)

Spence, J. T., & Helmreich, R. L. The many faces of androgyny: A reply to Locksley and Colten. *Journal of Personality and Social Psychology,* 1979, *37,* 1032–1046. (b)

Spence, J. T., & Helmreich, R. L. Androgyny versus gender schema: A comment on Bem's gender schema theory. *Psychological Review,* 1981, *88,* 365–368.

Spence, J. T., Helmreich, R. L., & Holahan, C. K. Negative and positive components of psychological mas-

culinity and femininity and their relationships to neurotic and acting out behaviors. *Journal of Personality and Social Psychology*, 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.

Stokes, J., Childs, L., & Fuehrer, A. Gender and sex roles as predictors of self-disclosure. *Journal of Counseling Psychology*, 1981, *28*, 510–514.

Strahan, R. F. Remarks on Bem's measurement of psychological androgyny: Alternative methods and a supplementary analysis. *Journal of Consulting and Clinical Psychology*, 1975, *43*, 568–571.

Sturdevant, S. *Therapy with women: A feminist philosophy of treatment.* New York: Springer Publishing Co., 1980.

Tellegen, A. Practicing the two disciplines for relaxation and enlightenment: Comment on "Role of the feedback signal in electromyograph biofeedback: The relevance of attention" by Qualls and Sheehan. *Journal of Experimental Psychology: General*, 1981, *110*, 217–226.

Wiggins, J. S., & Holzmuller, A. Psychological androgyny and interpersonal behavior. *Journal of Consulting and Clinical Psychology*, 1978, *46*, 40–52.

Zedeck, S. Problems with the use of "moderator" variables. *Psychological Bulletin*, 1971, *76*, 295–310.