

Some Methodological Comments on Labels, Traits, Interaction, and Types in the Study of "Femininity" and "Masculinity": Reply to Spence

Auke Tellegen and David Lubinski
University of Minnesota

We concur with Spence that the labels "femininity" and "masculinity" refer to a multidimensional domain and have been used inappropriately for scales that measure more circumscribed gender-related traits such as dominance and nurturance. It is nevertheless possible to investigate more encompassing constructs, such as androgyny and gender identity, through appropriate multivariate methods that incorporate, for example, interactions between dominance, nurturance, and other gender-related variables. In contrast, the common practice of exclusively relying on "feminine," "masculine," and other types or patterns, and treating these as purely nominal categories, hinders rather than aids recognition of underlying simple and complex (for example, interactive) relations. Furthermore, the study of interactions among measures of individual differences should be distinguished from a search for "real types." Interactions do not imply types, types do not imply interactions, although empirically the two phenomena may co-occur. The study of types does have a place in the study of gender-related characteristics and individual differences generally; it bears directly on questions of continuity and discontinuity important in personality and psychopathology and requires its own methodology.

Spence's (1983) comments concern two issues: the correct interpretation of the "femininity" and "masculinity" measures of the Bem Sex-Role Inventory (BSRI) and the Personality Attributes Questionnaire (PAQ), and our interpretation (Lubinski, Tellegen, & Butcher, 1983) of the construct of androgyny as we encountered it in the literature. With respect to the first issue, Spence faults us for uncritically accepting overinclusive concepts of femininity and masculinity and blithely treating the BSRI and the PAQ scales as measures of these "dubious constructs." In regard to the second issue, she credits us with a "neat way of thinking about androgyny" (p. 445) but stresses the legitimacy of alternative conceptions and doubts that our interactive androgyny concept has actually been advocated by any other investigator.

In this reply we will deal with the specific points Spence has raised and with some important closely related issues. Our purposes

are (a) to explain our views of femininity and masculinity as labels and concepts, (b) to document further our claim that an interactive androgyny concept is clearly implied by certain current formulations and deserves commensurate attention, (c) to show that currently popular "typological" analyses hinder rather than help the discovery of underlying relations and are no substitute for appropriate multivariate dimensional approaches, and (d) to stress the distinction between interactional and typological analyses. We hope that this exchange will contribute to the conceptual and methodological quality of future studies of gender-related personality characteristics.

Labeling Scales

Spence is displeased with what she believes to be our unquestioning acceptance of femininity and masculinity as highly generalized dispositional constructs and of the BSRI and the PAQ F (Femininity) and M (Masculinity) scales as measures of these constructs. She cites the empirical evidence indicating the multidimensionality of gender-related attri-

Requests for reprints should be sent to Auke Tellegen, Department of Psychology, University of Minnesota, Elliot Hall, 75 East River Road, Minneapolis, Minnesota 55455.

butes and points to the implication that the PAQ and BSRI F and M scales provide only a very partial representation of this broad domain.

As a case in point, Spence and Sawin (in press) have recently made a persuasive case for the importance of a bipolar concept of gender identity—the “image” of oneself as feminine or masculine—which they consider to be reflected in self-ratings of femininity and masculinity. These two descriptors form a bipolar dimension distinct from the Masculinity and Femininity factors formed by the items of the PAQ and BSRI F and M scales.

We can only concede Spence's point that Femininity and Masculinity are misleading and overinclusive names for what the BSRI and the PAQ measure. The problem is that the authors of these scales happen to have chosen these labels and continue to use them. In the words of Helmreich, Spence, and Holahan (1979), “The scales have been demonstrated to discriminate between the sexes in diverse populations, varying widely in age, ethnicity, and social class . . . thus justifying use of the labels *masculinity* and *femininity*” (p. 1632).¹ Elsewhere, Spence (1979) describes the characteristics measured by these scales as “desirable components of psychological masculinity and femininity” (p. 170) or, more briefly, as “psychological masculinity and femininity” (p. 181–182); the latter description recurs in the title of Spence, Helmreich, and Holahan's (1979) article. It seemed simplest for us to adopt these labels for the BSRI and the PAQ scales and for the personality dispositions they represent. We regret not to have stated explicitly that to us these titles were purely conventional and nominal and do not reflect our own substantive views.

Our contrition is made tolerable by knowing that confusion could have been avoided altogether if Spence and Helmreich (as well as Bem), in the light of the empirical evidence available at the time indicating the multidimensionality of the F–M domain (cf. Constantinople, 1973), had acted in accordance with Spence's own current advice and had launched their scales under different names to begin with. Even now we suspect confusion will continue unless the scales in question are formally renamed and corresponding changes

in the authors' theoretical views are explicitly cited.

In fact, in our own article we express the view that the M and F scales in question are measures of dominance and nurturance–warmth, respectively. In our concluding sections we speak qualifyingly of “‘masculinity’ in the sense of dominance–poise” (as measured by the short BSRI–M and the EPAQ–M⁺ [Extended Personality Attributes Questionnaire]) and of “‘feminine’ behaviors, such as nurturance and warmth (as assessed by the short BSRI–F and the EPAQ–F⁺)” (p. 437).

Indeed, our preference is for scale names that inform us, within limits of conciseness, of the specific manifest (“phenotypic”) item content. Phenotypic scale names, being neither overinclusive nor overexclusive, not only lead to more accurate predictions of correlations with other variables but they also leave it clearly up to the theorist to spell out reasons for any predicted connections that go beyond phenotypic overlap such as are implied by more inferential (“genotypic”) labels.

Some personality measures may be so heterogeneous in content that a simple phenotypic label eludes us. This is often true for “empirically keyed” measures, particularly if they are developed against a single and psychologically complex external criterion. For example, some existing femininity–masculinity scales are essentially heterogeneous collections of items whose one ensured common feature is the capacity to differentiate to some extent, if in varying ways, between females and males, gender being the item-selection criterion. The internal and external correlational networks involving the total score derived from such heterogeneous scales may be so poorly structured and haphazard that they cannot be claimed to embody a coherent psychological construct. Sweeping genotypic labels, such as psychological femininity–masculinity, for mixed collections of gender-differentiating items are too pretentious. Even

¹ In response to a preprint of this article, Spence, in a personal communication, stresses that the choice of the labels femininity and masculinity for the Personality Attributes Questionnaire scales was based on the gender-differentiating property of these measures and not on any assumptions of psychological surplus meaning.

if a heterogeneous M-F scale represents a certain pool of important gender-differentiating attributes, the aggregate scale score does not necessarily measure anything as systematic as a personality trait.

Occasionally we encounter a scale that not only covers a range of content that defies simple phenotypic description but at the same time represents a systematic structure that justifies a genotypic label. Loevinger and Wessler's (1970) measure of ego development is perhaps the best known recent example. Some well-substantiated "higher order" factor scales fall in the same category. In general, however, genotypic labels are hazardous, much as the hope of finding genotypic connections may motivate research. We see good reasons to prefer phenotypic names like "dominance," "honesty," "orderliness," and "sociability" (if they are descriptively apt) over such genotypic counterparts as "masculine protest," "superego strength," "anal-ity," and "strength of nervous system."

In the case of the BSRI and the PAQ scales, Spence prefers "instrumentality" and "expressiveness" over "masculinity" and "femininity," respectively. She adopted the former two terms from Parsons and Bales (1955), who postulated a universal division of labor between the sexes according to which man's "instrumental" tasks maintain the family in the larger community setting, whereas woman's "expressive" roles promote well-being and harmony within the family.

Their merits as global sociological constructs aside, as phenotypic descriptors, "instrumentality" and "expressiveness" do not on reflection seem to us much more justified than "masculinity" and "femininity." "Instrumental" covers a great deal more than is contained in the BSRI and PAQ M scales, and "expressive" as a psychological term is hardly coextensive with the specific nurturant qualities described in the F scales. Our own preference is to see "femininity" replaced by "nurturance-warmth" and "masculinity" by "dominance-poise" as labels for the BSRI and PAQ scales. Although those interpretations are more restrictive, it should be kept in mind that in the traditional cultural view, dominance and nurturance-warmth are especially important traits, supposedly requiring differential reward in the

two sexes and leading to mental health (only) if sex appropriate.

In the meantime, we neither accept nor reject potentially important genotypic concepts such as femininity, masculinity, androgyny, or gender identity. The data available bearing on these constructs are not, as yet, compellingly confirmatory or disconfirmatory and by the same token neither are the methods suggested for their assessment. We are not in a position today to designate individuals confidently as, for example, "psychologically androgynous" or "weakly identified with own gender" on the basis of a formal and objective assessment procedure let alone proceed from there with treatment recommendations as some have suggested (e.g., Gilbert, 1981; Kaplan, 1976; Osofsky & Osofsky, 1972; Rebecca, Hefner, & Oleshansky, 1976; Sturdevant, 1980).

Current Status of the Interactive Androgyny Construct

Spence finds something of interest in the view of androgyny as a trait-interactive construct. She does not believe, however, that this idea has in so many words been advanced by other investigators. We agree. Had we encountered it in the literature we would not have troubled ourselves to formulate it once more. We concur with Harrington and Andersen's (1981) recent characterization of the multiplicative model as "barely articulated and rarely examined" (p. 745).

On the other hand, we believe that an interactive concept is inherent in certain existing formulations, including Bem's and Spence's writings, as we have tried to show in our discussion of the concepts of tempering (Lubinski, Tellegen, & Butcher, 1981) and mitigation (Lubinski et al., 1983). Our self-assigned task was to explicate, that is, to put in clearer and more testable form, what we believe to be implied in their formulations. The reader will have to decide how cogent a case we have made.

Although Spence acknowledges the affinity of Kaplan and Bean's (1976) interpretation of androgyny as more than "a sum of masculinity and femininity" to our interactive concept, she does not really evaluate our attempt at explication. We are especially dis-

appointed that she does not discuss our observation that an interactive conception allows for the assessment of androgyny as an emergent phenomenon on the basis of two (or more) relatively independent measures of femininity and masculinity. As we pointed out (Lubinski et al., 1983, Footnote 2), this interpretation completely resolves the "logical contradiction" Spence and Helmreich (1981) attributed to the use of Bem's two essentially independent BSRI scales for assessing a coherent trait of "gender-schematic processing" (the low end of which would be occupied by androgynous individuals). Its ability to impart logical consistency to Bem's approach is precisely what supports our interpretation of her concept of gender-schematic processing as implicitly interactive. (For a different interactive conception, see the discussion below.) Spence also, unfortunately, does not discuss our interactive view of "unmitigated" trait expressions that her F_C^- and M^- scales are meant to assess.

In the interim, an explicitly interactive view of androgyny, in relation to creativity, has been independently proposed by Hargreaves, Stoll, Farnworth, and Morgan (1981) and in the earlier mentioned paper by Harrington and Anderson (1981). Hargreaves et al. conceive of androgyny as an adaptive interdependence of feminine and masculine characteristics, testable as a Femininity \times Masculinity interaction. This idea was borne out by their results, although these were not presented in enough detail to permit a substantive appraisal.

In addition, Harrington and Andersen (1981), to illustrate their multiplicative model, described the hypothetical case of a synergistic relation between feminine aesthetic sensitivity and masculine engineering skills in contributing to the creativity of architects. They interpret the results of their Femininity \times Masculinity interaction analysis as tentatively compatible with the multiplicative model. Evidently, interactive androgyny is no longer an implicit notion.

Quantitative Traits and Nominal Classes

Within sexes, measures of gender-related characteristics typically show quasi-continuous distributions and estimate the varying

amounts of a given trait different individuals possess. In our regression analyses, the original scores are preserved so that no information is lost. As Spence points out, she and her co-workers (Helmreich et al., 1979) have also performed regression analyses. However, in the study just cited, the $M \times F$ interaction was not appropriately assessed because a different scale, the Attitude Toward Women Scale, was entered as a predictor intermixed with M and F, *before* the $M \times F$ variable.

In contrast, procedures that replace the original scale scores by dichotomies, and substitute a fourfold "typology" of feminine, masculine, androgynous, and undifferentiated individuals for the original F and M scores of the BSRI or the PAQ, result in considerable loss of information. This much we (and others) have said before. Unfortunately, the typological analyses reported in the literature often do not amount to merely a coarse way of measuring individual differences on certain dimensions but represent abandonment of the dimensional structure itself. Specifically, findings are often not displayed and analyzed, as one might expect, in a two-dimensional framework that would directly reveal F and M main and interaction effects (for example, in a figure containing, say, one graph representing the low-F subjects, permitting a comparison of the low-F subjects with low scores on M with those receiving high scores on M, and a second graph allowing the same comparison among high-F subjects). On the contrary, the results are often displayed and compared in a way that juxtaposes the four subgroups as purely nominal categories without underlying dimensionality (e.g., S. L. Bem, 1981; Helmreich et al., 1979; Markus, Crane, Bernstein, & Siladi, 1982; Shaw, 1982). Sometimes the structural picture is made still more fragmentary by the omission of one or two subgroups (Clarey & Sanford, 1982; Lamke & Bell, 1982; Welch & Huston, 1982). In all these cases, one actually has to replot the results to perceive the dimensional and interactive pattern or make meaningful comparisons between different studies. If that is done, published findings become more revealing.

As an example, let us consider a particularly interesting study by Moran (reported in S. L. Bem, 1981) on the tendency in free re-

call to cluster according to gender. Subjects were presented with a heterogeneous list of words that had been consistently rated by judges as either masculine, feminine, or neutral. The dependent variable was the percentage of sequential word pairs produced during recall that belonged to the same "gender." This measure was chosen as a subtle index of preference for gender-based schematic processing.

For the female half of the sample, the analysis consisted of the following three planned comparisons among the four familiar types derived from the BSRI F and M scales: (a) sex-typed (high-F-low-M) subjects versus the other three groups combined, (b) cross-sex-typed (low-F-high-M) subjects versus androgynous (high-F-high-M) and undifferentiated (low-F-low-M) subjects combined, and (c) androgynous versus undifferentiated subjects. For the male subjects, the comparisons were the same except that the criteria for assigning subjects to the sex-typed and cross-sex-typed groups were, of course, interchanged. With these assignments made prior to the actual analysis, gender itself was not found to affect the clustering measure either directly or in interaction with the F or M variables. At this point, attention necessarily focused on the above mentioned planned comparisons. Only the first contrast produced significant results and indicated that sex-typed subjects (high-F-low-M females and low-F-high-M males) engaged in more gender-based clustering than the rest of the sample. Accordingly, Bem's main conclusion was that sex-typed individuals engage in gender-schematic processing.

However, the chosen three contrasts had in effect destroyed the dimensionality of the individual differences underlying these findings. In order to reconstruct it, separate sets of dimensional plots were derived for females and males from the results reported by S. L. Bem (1981, Figure 1) for the sample as a whole. Sex differences were assumed to be negligible following the opposite assignments of females and males to sex-typed and cross-sex-typed cells.

The new plots revealed that among females, F is correlated positively with gender-schematic processing, whereas M is negatively correlated. For males, the results are

in the opposite direction: F is negatively correlated with gender-schematic processing, and M positively correlated. In other words, in both sexes both scales were found to predict gender-schematic processing, but these predictive relations were in opposite directions for the two scales as well as for the two sexes.

To the extent that the results reported by Bem represent all major trends, an appropriate "hierarchical" regression analysis, with F, M, and gender and their interactions as independent variables, should show no appreciable F, M, or gender main effects and no androgynous F \times M interaction but conspicuous Gender \times F and Gender \times M interactions, demonstrating the role of gender itself as a general moderator variable!

Given this dimensional perspective, one can now recognize sex-typed subjects (high-F-low-M females and low-F-high-M males) as having the strongest gender-schematizing tendencies, which Bem indeed found to be the case. On the other hand, androgynous and undifferentiated subjects can now be seen to occupy an intermediate position, which is contrary to the familiar view of the androgynous person as especially free of sex-typed thinking. Finally, it is the cross-sex-typed subjects (low-F-high-M females and high-F-low-M males) who now naturally emerge as least likely to use gender schemata.

On the basis of this dimensional pattern of results, one could even venture the following generalization: A person will tend to rely on gender as an informative distinctive feature of things in general to the extent that her or his own perceived characteristics fit the sexual stereotype, that is, fit the relevant prototypical gender schema. Which of the two gender schemata is the relevant one depends, of course, on the individual's gender or, alternatively, her or his sense of gender identity (the distinction between the two becomes important when they diverge as in the case of transsexuality).

If one were to include a larger number of personality measures in one's investigation (in other words, if one were not limited to measures of dominance and nurturance), then jointly powerful relations between these variables and gender-schematic processing might emerge, with gender or gender identity

playing the pivotal role of controlling the direction of these relations. Given Spence's current emphasis on gender identity, it is interesting that Moran's findings indeed appear to indicate the special importance of gender for cognitive functioning.

From a methodological point of view, the main implication of this example is that we cannot discard the ordering provided by measures such as current F and M scales without considerable loss of structure and perspective. Replacing dimensional data by nominal categories makes psychometrically, and therefore psychologically, no sense.

Nominal Classes and Real Types

Our main focus has been on potentially important interactions between gender-related individual-difference variables. We have argued that they be attended to and analyzed properly. It is in that context that we have been critical of data analyses in which subject groupings are created by collapsing quasi-continuous variables into two-point scales.

Our criticism is twofold. First, the practice discards information about the strength and precise form of the relation being studied. It is, at best, an inefficient approach, even if it does not preclude the discovery of some interactions. Second, and more serious, certain analyses ignore and in effect destroy the underlying dimensionality of the original quantitative measurements by treating certain classes based on score patterns as purely nominal categories.

Our concern with studies relying on placing subjects in classes or types has not led us so far to consider typological issues proper. However, in a discussion of gender-related traits and interactions, a few comments on types seems called for, even if they are in part not in direct response to Spence.

First, we offer a definitional clarification. Following Meehl (1982, p. 127), we define *type* loosely and generally as "nonarbitrary class," emphasizing that it is a theoretical construct whose particular meaning consequently will become clear only as one's chosen overall conceptual and empirical context develops more fully. As Meehl points out (p. 127), all "taxa" (his preferred term for types) are classes, but not all classes are taxa. The

difference between taxa or *real types* (the term we will use) and classes that may or may not be real types is an important one. In a recent exchange in this journal about some of the same issues as concern us here, Block and Ozer (1982) make a distinction between type as distinctive form, for which they reserve the term *type* (and which we have called real type), and type as label. Weiss, Mendelsohn, and Feimer (1982), in the same series of articles, concur with the distinction.

Had our main concern been of a typological nature, we might have raised questions such as the following: Do individuals of the *same* gender belong to feminine and masculine types? Can one be a member of both types? Do additional types exist? What are biological or social sources of these types? Do we find types among men but not among women, or vice versa? Do we find any evidence of types at all within sexes in the domain of gender-related traits?

Interesting as these questions are, we viewed problems of interaction, our chosen topic, as conceptually separable from typological issues. Even so, some current writings may create the impression that an inherent conceptual connection does exist between real types and interactions. The following remarks deal primarily with this issue.

In an article by Mendelsohn, Weiss, and Feimer (1982), the first of the series mentioned earlier, it is asserted (pp. 1158-1159) that the distinction between being a member of a real type and merely occupying a particular location on a personality dimension requires some empirical evidence of "discontinuity." The issue of continuity versus discontinuity is, of course, a very important one in personality and psychopathology, and in the present context it is central. But Mendelsohn et al. (1982) go on to specify that discontinuity can be established in one of three ways: (a) evidence of a multimodal score distribution, (b) evidence of a nonlinear (e.g., U-shaped) regression of external variables on the dimension in question, and (c) evidence of interactions between certain dimensions in relation to external variables. It is the third criterion that is especially relevant to our discussion. Actually, our position is that these specific standards are not suitable for real types, although we do accept the gen-

eral criterion of underlying discontinuity (which in the case of continuous dimensions corresponds to the idea of class boundary).

We begin by pointing out that the requirement of a multimodal distribution is commonly regarded as too stringent. This is easily seen to be the case if one interprets real types as hypothetical entities whose presence is to be inferred from fallible indicators. The values of such indicators are expected to show some variation even among individuals belonging to the same real type. For example, the measures of dominance and nurturance discussed earlier are fallible indicators of gender (and as such, of course, are indicators of real types). One typological model assumes that fallible quantitative indicators will show a normal distribution within any given type. Under this model, the "mixed" distribution of a population consisting of distinct types is not necessarily multimodal. For example, if two types do not have markedly different means, or if they are present in markedly differing proportions, they could jointly form a unimodal distribution. The basic problem is inferring a mixture not observing bimodality (e.g., Everitt, 1981).

On the other hand, the existence of nonlinear relations, including U-shaped ones, mentioned by Mendelsohn et al. (1982), is too lenient a criterion for the presence of real types. The discontinuity of a quasi-step function may suggest a typological situation, but many other nonlinear functions could reflect processes whose variations are intrinsically continuous. Nonlinearity does not necessarily mean discontinuity.

Finally, and most relevant to our discussion, interactions between individual-difference variables again do not necessarily reflect the presence of real types. For example, some interaction effects may be in the nature of a relation between two variables that changes continuously in strength as a function of continuous changes on a third variable. Such effects would not in themselves provide the evidence of discontinuity that suggests real types.

Because of these considerations, we do not see our discussion of interaction as directly related to truly typological issues. Gender-related measures, such as the BSRI and PAQ scales, could be indicators of real types.

Membership of a real type, in turn, could interact with other variables (i.e., it could be a moderator variable). Tellegen, Kamp, and Watson (1982), in their examination of analyses by D. J. Bem and Allen (1974), elaborated a hypothetical example contrasting a "trait" subpopulation, consisting of individuals among whom a set of trait indicators were positively correlated, with a "no trait" subpopulation, in which the same indicators were uncorrelated. In other words, subpopulation membership was a moderator variable. But, again, the presence of interaction effects ("configurality") is not in itself evidence for the existence of types, unless the particular form of these effects indicates discontinuity, as was the case in the contrived example of Tellegen et al. (1982). Furthermore, types do not imply interactions any more than interactions imply types: The non-implication holds in both directions.

On the other hand, grouping subjects in test-profile types could be rationalized on the grounds that any set of reasonably homogeneous profile classes, whether real types or not, will capture whatever interactions happen to exist between the profile components (Butcher & Tellegen, 1978). This was essentially the topic of our first two papers (Lubinski et al., 1981, 1983), in which we stressed the desirability of testing these interactions empirically and doing this through regression methods without recourse to rather crude and arbitrary typologies. Even if the presence of scale interactions would explain and partially justify the creation and use of profile classes, it would in no way imply an inherent link between interactions and real types.

These observations apply to any kind of classification, including Minnesota Multiphasic Personality Inventory profile types (Dahlstrom, Welsh, & Dahlstrom, 1972); Personality Disorders, according to the most recent psychiatric diagnostic manual known as DSM-III (*Diagnostic and Statistical Manual of Mental Disorders*; American Psychiatric Association, 1980); and other current classification systems. Each of these may capture important interaction effects and may be descriptively and predictively useful for that reason. They may also reflect certain variations in underlying personality structure. Yet

these features are in themselves no evidence that the profile classes in question represent discontinuities indicative of real types. They could instead be arbitrary divisions of a multidimensional space in which covariational changes do occur, and motivate the creation of profile types mentioned earlier, but are continuous.

Conclusion

We agree with Spence that the domain of gender-differentiating characteristics is multidimensional, not one- or two-dimensional. Any clearly one-dimensional femininity or masculinity scale can therefore represent only one facet of this domain. Generalizations based on findings with such a scale will have to be correspondingly restrained.

This realization does not imply that the study of gender-related phenomena is fated to be fragmentary. Interaction models, described earlier, of androgyny and gender identity illustrate one possible kind of functional link between seemingly unrelated phenomena.

Furthermore, the creation of feminine, masculine, and other subject categories has not been productive. Worse, their treatment as nominal categories has tended to obscure underlying relations. We suggest that analyses of this kind be abandoned in favor of appropriate multivariate methods.

Although we have obviously been concerned with certain uses of subject classifications or types, our initial topic, interactions, is not inherently linked to the issue of real types. On the other hand, we agree with Dahlstrom (1972) and Block and Ozer (1982) that typological concepts are important in psychology. As we pointed out earlier, they involve issues of continuity and discontinuity in personality and psychopathology. Of course, we share the view, most recently expressed by Mendelsohn et al. (1982), that the contribution of real types to personality description be demonstrated empirically. As in the case of interactions, this requires methods appropriate to that specific purpose.

References

American Psychiatric Association. *Diagnostic and statistical manual of mental disorders* (3rd ed.). Washington, D.C.: American Psychiatric Association, 1980.

- Bem, D. J., & Allen, A. On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. *Psychological Review*, 1974, 81, 506-520.
- Bem, S. L. Gender schema theory: A cognitive account of sex typing. *Psychological Review*, 1981, 88, 354-364.
- Block, J., & Ozer, D. J. Two types of psychologists: Remarks on the Mendelsohn, Weiss, and Feimer contribution. *Journal of Personality and Social Psychology*, 1982, 42, 1171-1181.
- Butcher, J. N., & Tellegen, A. Common methodological problems in MMPI research. *Journal of Consulting and Clinical Psychology*, 1978, 46, 620-628.
- Clarey, J. H., & Sanford, A. Female career preference and androgyny. *Vocational Guidance Quarterly*, 1982, 30, 258-264.
- Constantinople, A. Masculinity-femininity: An exception to a famous dictum? *Psychological Bulletin*, 1973, 80, 389-407.
- Dahlstrom, W. G. *Personality systematics and the problem of types*. Morristown, N.J.: General Learning Press, 1972.
- Dahlstrom, W. G., Welsh, G. S., & Dahlstrom, L. E. *An MMPI handbook: Vol. 1. Clinical interpretation*. Minneapolis: University of Minnesota Press, 1972.
- Everitt, B. S. Bimodality and the nature of depression. *British Journal of Psychiatry*, 1981, 138, 336-339.
- Gilbert, L. A. Toward mental health: The benefits of psychological androgyny. *Professional Psychology*, 1981, 12, 29-38.
- Hargreaves, D., Stoll, L., Farnworth, S., & Morgan, S. Psychological androgyny and ideational fluency. *British Journal of Social Psychology*, 1981, 20, 53-55.
- 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.
- 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.
- 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. Conclusion from sex stereotypes to androgyny: Considerations of societal and individual change. In A. G. Kaplan & J. P. Bean (Eds.), *Beyond sex-role stereotypes: Readings toward a psychology of androgyny*. Boston: Little, Brown, 1976.
- Lamke, L. K., & Bell, N. J. Sex-role orientation and relationship development in same-sex dyads. *Journal of Research in Personality*, 1982, 16, 343-354.
- Loevinger, J., & Wessler, R. *Measuring ego development I: Construction and use of a sentence completion test*. San Francisco: Jossey-Bass, 1970.
- 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.
- Lubinski, D., Tellegen, A., & Butcher, J. N. Masculinity, femininity, and androgyny viewed and assessed as dis-

- tinct concepts. *Journal of Personality and Social Psychology*, 1983, 44, 428-439.
- Markus, H., Crane, M., Bernstein, S., & Siladi, M. Self-schemas and gender. *Journal of Personality and Social Psychology*, 1982, 42, 38-50.
- Meehl, P. E. Taxometric methods. In P. C. Kendall & J. N. Butcher (Eds.), *Handbook of research methods in clinical psychology*. New York: Wiley, 1982.
- Mendelsohn, G. A., Weiss, D. S., & Feimer, N. R. Conceptual and empirical analysis of the typological implications of patterns of socialization and femininity. *Journal of Personality and Social Psychology*, 1982, 42, 1157-1170.
- Osofsky, J. D., & Osofsky, H. J. Androgyny as a life cycle. *Family Coordinator*, 1972, 21, 411-419.
- Parsons, T., & Bales, R. F. *Family socialization and interaction process*. Glencoe, Ill.: Free Press, 1955.
- Rebecca, M., Hefner, R., & Oleshansky, B. A model of sex-role transcendence. *Journal of Social Issues*, 1976, 32, 197-206.
- Shaw, J. S. Psychological androgyny and stressful life events. *Journal of Personality and Social Psychology*, 1982, 43, 145-153.
- Spence, J. T. Traits, roles, and the concept of androgyny. In J. E. Gullahorn (Ed.), *Psychology and women: In transition*. Washington, D.C.: V. H. Winston, 1979.
- Spence, J. T. Comment on Lubinski, Tellegen, and Butcher's "Masculinity, femininity, and androgyny viewed and assessed as distinct concepts." *Journal of Personality and Social Psychology*, 1983, 44, 440-446.
- 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 masculinity and femininity and their relationships to neurotic and acting out behaviors. *Journal of Personality and Social Psychology*, 1979, 37, 1673-1682.
- Spence, J. T., & Sawin, L. L. Images of masculinity and femininity: A reconceptualization. In V. O'Leary, R. Unger, & B. Wallston (Eds.), *Sex, gender and social psychology*. Hillsdale, N.J.: Erlbaum, in press.
- Sturdevant, S. *Therapy with women: A feminist philosophy of treatment*. New York: Springer, 1980.
- Tellegen, A., Kamp, J., & Watson, D. Recognizing individual differences in predictive structure. *Psychological Review*, 1982, 89, 95-105.
- Weiss, D. S., Mendelsohn, G. A., & Feimer, N. R. Reply to the comments of Block and Ozer. *Journal of Personality and Social Psychology*, 1982, 42, 1182-1189.
- Welch, R. L., & Huston, A. C. Effects of induced success/failure and attributions on the problem-solving behavior of psychologically androgynous and feminine women. *Journal of Personality*, 1982, 50, 81-97.

Received August 31, 1982 ■