

ILL: 170305438



ILLiad TN: 1047267



Borrower: ORU

Ship via: Odyssey

Service Level: GWLA

Odyssey: 128.223.84.211

Email: ill@uoregon.edu

Lending String:

*AFU,AZS,COD,CSL,IAY,ILU,IWA,IXA,IYU,KKU

ISSN: 0360-0025

OCLC: 2243426

ILL - AFU
UNIVERSITY OF ARKANSAS
UNIVERSITY LIBRARIES
365 N MCILROY AVE
FAYETTEVILLE AR 72701-4002

RETURNED Postage Guaranteed

University of Oregon
Knight Library, ILL
1501 Kincaid St.
Eugene, Oregon 97403-1299
United States

Date: 8/31/2016 9:19:34 AM

Call #: HQ768 .S4 v.17 Oc-D

1987

Location: **MAIN**

Volume: 17

Issue: October

Year: 1987

Pages: 375-400

Journal Title: Sex roles.

Article Author:

Article Title: : measuring gender differences

**Notice: This material may be protected by
Copyright Law (Title 17 U.S. Code)**

Maxcost charge: 35.00IFM

Patron:

Initials: _____
Shelf: _____ Per: _____
Sort: _____ ILL: _____

Bad Cite: _____

Years checked _____

Table of Contents / Index _____

Measuring Gender Differences: The Expressive Dimension and Critique of Androgyny Scales¹

Sandra Gill

Gettysburg College, Gettysburg, Pennsylvania

Jean Stockard, Miriam Johnson, and Suzanne Williams

University of Oregon, Eugene, Oregon

This paper calls for a new approach to the measurement of gender similarities and differences in personality. Critiques of current measures that are based on the measurement of stereotypical masculinity and femininity are reviewed, and an alternative measure is presented. It avoids many of the problems in the other scales by measuring expressiveness and instrumentality in such a way that they are not confounded with variables such as independence or autonomy. Results with this measure in five different samples are presented. These confirm our theoretical definitions of expressiveness and instrumentality, and suggest that only the expressive or relational dimension consistently differentiates samples of women and men.

The masculinity and femininity scales that were widely used in the 1930s through the 1960s tended to treat personality characteristics associated with women and men as bipolar. These scales also implicitly assumed that it is desirable for women to be feminine and for men to be masculine. Androgyny scales, designed to allow individuals to score high on both femininity and masculinity and thus be "androgynous," were developed in the 1970s. Initial

¹We thank Mary Rothbart for her insightful and constructive comments on an earlier draft; Leisha Sanders, Marcia DeCaro, Vickie Van Nortwick, and Lyn Cogswell for their clerical assistance; the Center for the Study of Women in Society for financial support; and Liza Kuecker and Patricia Gwartney-Gibbs for providing data used in the analysis. An earlier version of this paper was presented at the 1984 annual meeting of the Pacific Sociological Association.

research using such scales suggested that it was "healthy" to be both masculine and feminine. These new scales were not, however, based on theoretically derived personality dimensions but relied instead on a hodgepodge of stereotypical ideas, primarily of college students, about personality differences between males and females. More recently, in the 1980s, androgyny scales have been criticized on both methodological and theoretical grounds, yet they continue to be relied upon for measuring the degree to which an individual combines presumably masculine and feminine traits. This reliance upon androgyny scales is particularly problematic at a time when increasing academic attention is focused upon positive aspects of the female orientations and capacities that are not easily amenable to measurement by these instruments.

In this paper we argue that the time is ripe for a third wave of research on gender differences in personality. This work should begin to develop scales based on theoretically meaningful, unidimensional variables that will allow researchers to examine the ways in which the personalities of men and women do and do not differ. In support of this goal we present findings from self-ratings on an instrument developed by two of the present authors (Johnson, Stockard, Acker, & Naffziger, 1975). It is designed to measure theoretical dimensions thought to be associated with gender differences in personality and avoids many of the methodological problems attributed to androgyny scales. We describe replications of this scale on various populations, and then discuss the implications of our findings for the theoretical understanding and empirical measurement of gender differences in self-concept in contemporary American society.

THEORY AND RELATED LITERATURE

Below we briefly review previous attempts to theoretically and empirically distinguish masculinity and femininity, examine the two most commonly used androgyny scales, and discuss our own attempt to link theoretical distinctions with empirical measures of male-female personality differences.

Distinguishing Masculinity and Femininity

While theoretical discussions of personality differences typically characterizing males and females have long occupied social scientific attention, the most important and influential of these theories have not been based on the concepts of masculinity and femininity per se. Talcott Parsons associated expressive orientations with female roles and instrumental orientations with

male roles (Johnson, 1955, 1963; Parsons & Shils, 1952; Parsons, Bales, & Shils, 1954). David Gutmann (1970) described autocentric and allocentric ego styles, and David Bakan (1966) developed a distinction between communion and agency. More recently, Nancy Chodorow (1978) has connected women's more relational orientation and men's less relational orientation to women's mothering, and Carol Gilligan (1982) has described women's views of morality in terms of responsibility and interdependence, and men's views in terms of rights and noninterference in the rights of others. All of these distinctions have a common theme that emphasizes the orientation of women toward social integration and men toward more impersonal or individualistic goals.

Social psychologists who have tried to develop empirical measures of masculinity and femininity have generally paid little attention to these more theoretical approaches. Beginning in 1936, a number of measures were developed to place individuals on a single continuum ranging from very masculine to very feminine (e.g., Constantino & Heilbrun, 1964; Hathaway & McKinley, 1943; Gough, 1952; Gough & Heilbrun, 1965; Terman & Miles, 1936). These scales have been criticized on a number of grounds, but the criticism that proved to be most influential on future research was that they assumed that masculinity and femininity constituted a single, bipolar dimension (Constantinople, 1973).

During the 1970s several scales were developed that have separate masculinity and femininity scales. The first developed and by far the most widely used scales at this time are Bem's Sex Role Inventory (BSRI; Bem, 1974) and the Personal Attributes Questionnaire (PAQ; Spence, Helmreich, & Stapp, 1974, 1975; Spence & Helmreich, 1978; for other scales see also Berzins, Welling, & Wetter, 1978, and Heilbrun, 1976). Both of these instruments allow individuals to be categorized as either masculine, feminine, undifferentiated, or if high in both masculinity and femininity, androgynous (e.g., Bem, 1975; Spence & Helmreich, 1978). Research findings based on these scales, however, are questionable, since the instruments themselves, particularly the widely used BSRI, have come under serious attack on theoretical, methodological, and philosophical grounds (see Heerboth & Ramanaiah, 1985; Jones, Chernovetz, and Hansson, 1978; Kelly & Worell, 1977; Locksley & Colten, 1979; Lott, 1981; Morgan, 1982; Myers & Sugar, 1979; Myers & Gonda, 1982; Pedhazur & Tetenbaum, 1979; Stark-Adamec, Graham, & Pyke, 1980).

That these scales are designed to measure androgyny in no way eliminates the necessity of measuring the concepts of masculinity and femininity. The term *androgyny* itself, however it is defined operationally, does involve combining "masculinity" and "femininity" in some way. If the initial measures of masculinity and femininity, viewed as separate dimensions, are

unsatisfactory, combining them by changing the scoring system does not eliminate the measurement problem. In other words, the same basic measurement problems apply to androgyny scales as to conventional masculinity-femininity (M-F) scales (Taylor & Hall, 1982). The interest in the idea of androgyny obscured this fundamental underlying difficulty for a long time.

Development of the BSRI and the PAQ

To develop their measures of masculinity and femininity, both Bem (1974) and Spence, et al. (1974, 1975) attempted to ascertain the personality characteristics stereotypically associated with males and females. Bem presented Stanford undergraduates with a list of 400 personality traits, and asked them to evaluate each on its desirability for a woman or for a man in American society. On the basis of *t* tests, 20 traits such as "aggressive," "willing to take risks," and "self-sufficient" were judged as significantly more desirable for a man than for a woman, and thus determined to reflect masculinity; while 20 traits, including "cheerful," "flatterable," and "gullible," were in a similar manner determined to reflect femininity.

Spence, et al., using a pool of items nominated by college students as differentiating men and women (Spence & Helmreich, 1978, p.32), selected 55 items on which both sexes reported consistent stereotypes. Items seen as socially desirable for both sexes but believed to occur to a greater degree in males were placed on the Masculine scale. Items seen as socially desirable for both sexes but believed to occur to a greater degree in females were included on the Feminine scale. A third scale, designated the Masculinity-Femininity scale, is composed of items seen as socially desirable for one sex group, but not for the other. Spence and Helmreich report that "conceptual justification for the division of the items . . . was found by inspecting item content" (1978, p. 33).

The BSRI and PAQ are empirically developed in the sense that both scales are deliberately based on students' perceptions of stereotypically perceived differences between the sexes. The linkage to theory about actual differences between the sexes in self-concept was developed only later by inspecting the selected items (see Bem, 1974; Spence & Helmreich, 1978, p. 33). Both Bem and Spence et al. claim that they are measuring instrumental and expressive or agentic and communal personality traits, and loosely equate these with masculinity and femininity. Spence and Helmreich state without theoretical discussion that "the items on the M scale refer largely to instrumental, agentic characteristics, while items on the F scale refer largely to expressive, communal attributes" (Spence & Helmreich, 1978, p. 33). Bem states that her measure of "psychological androgyny" allows individuals to

"be both masculine and feminine, both assertive and yielding, both instrumental and expressive" (Bem, 1974, p. 157; see also Bem, Martyna, & Watson, 1976, p.1016).. Yet beyond this brief equating of masculinity with assertiveness and instrumentality, and femininity with yielding and expressiveness, neither Bem nor Spence and Helmreich discuss the dimensions implicit in measuring masculinity and femininity in our culture. Our own argument, made first in 1975 and presented in more detail below, is that measures of expressiveness and instrumentality must conform to theoretical descriptions of these terms. We specifically maintain that it is essential to separate the idea of dependency and emotionality from expressiveness, a distinction not made in the BSRI or PAQ.

Moreover, in the original development of their measures, neither Bem nor Spence and Helmreich assessed the internal consistency of each scale through techniques such as a factor analysis. If they had done so it would have become clear that the scales draw upon a number of different dimensions. Several factor analyses of results using Bem's scale have demonstrated that it contains more dimensions than those suggested in the original development (see Collins, Waters, & Waters, 1979; Feather, 1978; Gaudreau, 1977; Moreland, Gulanick, Montague, & Harren, 1978; Pedhauzer & Tetenbaum, 1979; Ruch, 1984; Waters, Waters, & Pincus, 1977; and discussion below). Studies focusing on the PAQ also indicate that the scales may contain more dimensions than originally suggested (see Gaa, Liberman, & Edwards, 1979). In our view a scale is needed that has a better theoretical/conceptual rationale and whose assumptions can then be examined empirically.

The Expressive-Instrumental Distinction

The terms *instrumental* and *expressive* were originally used by the sociologist Talcott Parsons in the 1950s. In spite of much criticism by feminists (most recently by Miriam Lewin, 1984), the terms remain in general use. Our interest is not to discredit these terms, but rather to define them more clearly than Parsons did and argue that they can be of use if more narrowly specified.

Drawing on distinctions running through several of Parsons' theoretical works (e.g., Parsons, 1951; Parsons & Shils, 1951; Parsons et al., 1954),² we define an *instrumental* orientation as concern with the attainment of goals external to the interaction process, while an *expressive* orientation gives prima-

²Parsons himself dropped the terms *expressive* and *instrumental* when he adopted his "four function paradigm" outlined in Parsons and Smelser (1956).

cy to facilitating the interaction process itself (see Johnson et al., 1975). An instrumental orientation involves manipulating objects, the environment, and even people to attain goals and accomplish tasks external to the interactive system itself, while expressiveness involves understanding and dealing with emotions in self and others. Expressive actions involve tension management and motivational control among individuals; instrumental actions more often involve formal authority and technical control. Expressive action is oriented toward the system of interaction itself and is typically rewarded by affective attitudes such as love and friendship. Instrumental action is oriented toward objective ends and is typically sanctioned with more "affectively neutral" or impersonal attitudes such as approval, respect, and esteem.

It is crucial to note that in specifying this definition of expressiveness we have rejected aspects of the stereotypical perception of females as emotional, passive, and dependent, but retained the idea that expressiveness involves what is now called "interdependence" and "relationality." First we maintain that expressiveness is not the opposite of instrumentality; instead the concepts constitute two separate dimensions, each with a positive and a negative pole. Specifically, this means that expressiveness does not imply a lack of instrumental competence, nor does competence in instrumental activities preclude the ability to relate expressively.

Further, expressiveness is not the same as emotionality or emotional acting out. Although expressiveness does engage socioemotional skills, it is misleading to view it as simply being emotional or emotionally labile. In our culture, women are expected to resonate with, cope with, and even define emotions for self and others, but this is hardly the same as *being* emotional. Sometimes the expression of emotion needs to be suppressed in order to emotionally help another person, and this too could properly be considered an expressive act.

Finally, we argue that passivity and dependence are *not* components of expressiveness (and concomitantly that activity and independence are not the same as instrumentality). Both expressiveness and instrumentality are active orientations. Stereotypical thinking tends to confound expressiveness with dependence and passivity, assuming that the expressive individual is dependent on others to a greater extent than the instrumental individual. In fact, however, expressiveness involves a sense of interdependence rather than dependence and can certainly also involve taking initiative as opposed to simply responding to others.³ In making these distinctions we assume that it is

³From these definitions it should also be clear that the instrumental-expressive distinction is narrower than and different from Bakan's agentic-communal distinction. Bakan's concepts are so global and all inclusive that their specific meaning is difficult to pin down. For example, the term *agentic* describes highly disciplined, self-oriented striving in the occupational world as well as the impulse-ridden self-assertion and self-extension of the infant. The instrumental-expressive distinction is also not the same thing as a self-oriented-other-oriented distinction.

an empirical question as to how males and females actually differ on these characteristics.

When the items in the PAQ and BSRI are compared to these theoretical specifications, it is apparent that they both include many items that do not properly belong in the instrumental or expressive dimensions. For instance, the PAQ's Masculinity scale contains items that clearly reflect autonomy rather than instrumentality (e.g., not at all independent to very independent, very passive to very active). The PAQ Femininity scale contains items suggesting positive expressive personality traits (e.g., not at all kind to very kind, not at all aware of the feelings of others to very aware of the feelings of others, very cold in relations with others to very warm in relations with others, etc.), but also includes items reflecting emotionality (not at all emotional to very emotional). The BSRI Femininity scale confounds expressiveness, emotionality, and dependence. For example, such words as *childlike*, *shy*, and *yielding*, which are included on Bem's scale, are not indicative of expressiveness. The Masculinity scale includes few items properly identified as instrumental, but many that could measure autonomy.

METHODOLOGY

The complete details on the original development of our instrument are given in an earlier article (Johnson et al., 1975). Below we summarize this earlier work and then report the methods used in replicating the study.

The Original Study

Work to develop our original instrument began in the early 1970s. After reviewing and discussing the definitions given above and the theoretical work of Johnson (1955, 1963), four women and three men, who were either Ph.D.s or advanced graduate students, independently selected adjectives from the Gough Adjective Checklist to reflect each of the following categories: positive expressive, negative expressive, positive instrumental, negative instrumental, active and/or independent, and passive and/or dependent. Thus items were specifically selected to reflect both the positive and negative components of three separate dimensions: expressiveness, instrumentality, and independence or autonomy.

Items placed in the same theoretical dimensions by at least five of the seven judges were then presented in a questionnaire to male and female college students enrolled in introductory sociology courses at the University of Oregon. These students were asked to rate themselves on each adjective with one of four possible answers: *very true of me*, *somewhat true of me*, *some-*

what untrue of me, and very untrue of me. One hundred thirty men and 135 women completed each item on the questionnaire.

Both factor and cluster analyses were used to determine whether variation in students' responses could be accounted for by underlying factors consistent with our theoretical contentions regarding the distinctions between expressiveness, instrumentality, and autonomy. Positive expressive, positive instrumental, and autonomy dimensions were clearly identified for both men and women. Smaller groupings, which conformed to judges' expectations of negative expressive, negative instrumental, and dependent traits also emerged from the multivariate analyses (see Johnson et al., 1975). There were only two deviations from the judges' expectations: First, within the theoretically derived concepts of instrumentality and independence, more groups of adjectives were identified by empirical analysis than predicted. Second, even though a small number of negative expressive and negative instrumental personality traits were clearly identified in the factor analyses, fewer adjectives were associated with the factors tapping negative expressive, negative instrumental, and dependent aspects of personality than were expected by the judges. Consistent with our theoretical expectations, the positive expressive traits formed a dimension separate from traits associated with emotionality and the acting out of feelings. The major finding from this analysis was that only the adjectives we had selected to represent the positive aspects of expressiveness seemed to form a unified group.

Scale scores were computed by summing an individual's self-ratings on the items included in each dimension (see below for details on scoring). The largest gender difference in these scores was on the expressive dimension, with men seeing themselves as less positively expressive than women. The differences on the other dimensions were much smaller.

Replications

To examine the consistency and generalizability of our findings we conducted four replications of this study. In 1982 we gave a questionnaire containing a list of adjectives and instructions identical to those used in our earlier study to 340 students enrolled in introductory level sociology courses at the University of Oregon. In 1982 the items were also administered to 250 high school students at a rural, largely working-class school in Oregon. In both cases the students completed the forms in class. In 1983 a subset of the items was administered to a representative sample of 430 male and female nurses in Oregon through a mail-out, mail-back survey (Kuecker, 1985). The items were also included on a mail-out, mail-back survey given to a random sample of 460 University of Oregon undergraduates in 1984. All samples con-

tained approximately equal numbers of males and females. (The sampling design called for an overrepresentation of male nurses so that male-female comparisons could be made.)

The samples differed in age, socioeconomic status, work status, and knowledge about gender role issues. They also differed in manner of administration (in an intact group vs mail-out, mail-back). In all cases the subjects were asked to mark the extent to which a given adjective "is true of you" on the 4-point scale used in the original study. Identical items were used in all administrations with only two exceptions. First, on the schedule given to the nurses, *daring*, *stern*, and *obliging* were omitted, and *hardworking* was substituted for *industrious*. Second, only items representing the positive pole of each dimension were used in the replications done in 1983 and 1984. This simplifies the presentation, and more closely parallels the versions of the BSRI and PAQ currently in use. Only the results for positively worded traits and dimensions are reported in this analysis. To the extent that similar results appear with each sample we may gain confidence in our conclusions.

To identify the underlying dimensions, separate principal component factor analyses with a varimax orthogonal rotation were performed on the self-ratings of the men and women in each of the sample groups. The results of these ten analyses were used to confirm or, when deemed necessary, to modify the scales slightly from the form presented in the 1975 article.

Scores on each item in the scales were summed and averaged, resulting in scale scores which could vary from 1 to 4. A higher score indicates that subjects were more likely to say that a given set of traits were "true" of them. The internal consistency of the scales was measured using Cronbach's alpha, a measure of reliability, and correlations among the scale scores were also examined. Finally, *t* tests were used to examine differences between the sex groups in average scale scores, and Cohen's *d* (the standardized mean difference or effect size) (Cohen, 1977) was used to measure the relative size of these differences.

RESULTS

The results with the replications generally confirm those found in the original study.

Testing the Theoretical Dimensions

Tables I and II report results of the factor analyses for each of the ten groups. The items determined by the judges to fall in each theoretical dimen-

Table I. Continued

Item	Sample											
	1972 College		1982 College		1982 High school		1983 Nurses		1984 College			
	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men		
Daring ^f	5 (.66)	5 (.78)	3 (.63)	1 (.68)	1 (.40)	2 (.75)	2 (.36)	2 (.33)	2 (.75)	4 (.80)	3 (.80)	
Adventurous ^f	5 (.82)	5 (.64)	3 (.81)	1 (.71)	1 (.30)	2 (.68)	2 (.31)	2 (.46)	2 (.36)	4 (.80)	3 (.79)	

^aThe numbers outside of parentheses represent the factors (in order of appearance) on which the item had a large loading. The figure in parentheses is the factor loading. Only loadings greater than .30 are included. A number of negatively worded items were included on the 1972 questionnaire. The results for these items have been omitted from this table for clarity of presentation.

^bItems included in expressiveness scale.

^cItems included in industrious dimension of instrumental scale.

^dItems included in analytical dimension of instrumental scale.

^eItems included in forceful dimension of autonomy.

^fItems included in adventurous dimension of autonomy.

^gAdjective not included on the nurses' questionnaire.

Table II. Percentage of Common Variance Accounted for by Each Factor in Five Different Samples

Factor	Sample											
	1972 College		1982 College		1982 High school		1983 Nurses		1984 College			
	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men		
1	28	24	35	34	29	30	37	35	32	40		
2	^a	19	19	18	15	15	27	20	27	23		
3	12	15	14	13	13	12	21	16	17	17		
4	10	8	10	10	10	10	15	16	8	10		
5	7	8	8	9	9	9	10	13	8	10		
6	^a	^a	7	8	9	8	8	8	8	8		
7	5	4			8	8						
8					7	7						

^aFactor associated with negatively worded items (see Table I footnote *d*).

sion are listed in the left-hand column of Table I. Factor loadings of .30 or greater were treated as meaningful and are reported in parentheses following the number indicating the order in which the factor emerged in each solution. The percentage of variance accounted for by each factor in each sample is given in Table II. The number of interpretable factors differed for the various groups analyzed, as we continued to include factors as long as they could be meaningfully identified. *All loadings of .30 or larger are reported in the table.* It is clear that the dimensions found in the factor analyses are very clear-cut and very similar to those theoretically determined. They are also very similar in each of the sample groups.

A strong expressiveness factor (including the items sympathetic, understanding, pleasant, considerate, good-natured, warm, and obliging⁴ underlies the self-ratings in each of the diverse groups. Only among college men in 1972 and 1982 was the expressiveness factor not the first to emerge, and only among these men did the adjectives associated with expressiveness not represent a single dimension.

Items selected to indicate instrumental personality traits emerged on identifiable factors in all sampled groups, but in none of the sampled groups were they unidimensional. As in 1975 we found that the adjectives selected as positive instrumental by the judges tend to reflect two distinct dimensions of self-ratings—one that we call the “industrious” dimension (thorough, efficient, industrious, planful) and another we refer to as the “analytical” dimension (analytical, foresighted, rational).

Items judged to reflect autonomy are identifiable in all groups. While the factor structures were less consistent from one sample to the next, two separable dimensions seem to appear. We refer to these as “forceful” (stern, forceful, aggressive, outgoing, assertive, independent, and active), and “adventurous” (daring and adventurous).⁵

⁴The term *obliging* may seem to connote a lack of independence but perhaps only from an instrumental point of view. Being helpful does not necessarily conflict with autonomy. Generally, when one looks at intercorrelations, it is men—not women—who are more likely to see expressive traits as associated with passivity and dependence (see Johnson et al., 1975, pp. 638–640). While certainly one may rationally pursue expressive goals, the connotations of rationality in stereotypic usage are predominantly instrumental. An underlying problem is that in an instrumental culture such as ours, expressiveness tends to be defined as the residual, the “other,” and develops connotations not of being a different dimension but an opposite dimension.

⁵This distinction differs from the one reported in 1975 when “independent” and “active” were seen as comprising a third separate dimension within the autonomy scale. We could find no consistent support for this distinction across the samples studied here, and thus these terms are now combined with the forceful dimension.

Table III. Reliability Coefficients of Expressive, Instrumental, and Autonomy Scales in Four Samples^a

Scales	1982 College		1982 High School		1983 Nurses		1984 College	
	Women	Men	Women	Men	Women	Men	Women	Men
Expressiveness	0.75	0.75	0.74	0.75	0.83	0.82	0.82	0.81
Instrumental								
Industrious	0.66	0.71	0.59	0.54	0.68 ^c	0.75	0.77	0.77
Analytical	0.39	0.52	0.44	0.44	0.69	0.69	0.59	0.59
Autonomy ^b								
Forceful	0.75	0.67	0.63	0.69	0.76	0.72	0.74	0.74

^aAll coefficients are Cronbach's alpha.

^bCoefficient alpha is not reported for the adventurous dimension because there are only 2 items on that dimension.

^cIf *hardworking* is included to replace *industrious*, 0.64.

Reliability

Coefficients alpha for the scales for each sex group in the 1982, 1983, and 1984 samples are shown in Table III. These results suggest that the scales are internally consistent. All of the coefficients are relatively high, although the instrumental scales show the lowest internal consistency and the most variability from sample to sample.⁶

The intercorrelations among the scales are usually positive and moderate in size (see Table IV). These results support the contention that the scales represent distinct dimensions and subdimensions. In addition, the moderate correlations between the expressive, instrumental, and autonomous scales support our theoretical contention that the measured attributes are not opposites, but that individuals may see themselves as expressive as well as instrumental and autonomous. Again, with only a few exceptions, similar results appear with each sample.

⁶In some of her early articles Bem reported relatively high levels of coefficient alpha (1974), yet later articles (e.g., Gaudreau, 1977; Moreland et al., 1977; Pedhazuer & Tetenbaum, 1979; Ruch, 1984) noted that her scales were far from unidimensional. When coefficient alpha was computed for the combined or total scales for the instrumental and autonomous dimensions, the reliability coefficients were somewhat higher than the coefficients for the subdimensions, even though the factor analyses usually distinguished the subdimensions fairly clearly. This occurred because the total scales have more component items than the subscales and illustrates the difficulties of using a reliability measure such as coefficient alpha apart from a test of the unidimensionality of a measure through some type of multivariate analysis (see Zeller & Carmine, 1980, p. 60; also Arm6r, 1974).

Table IV. Intercorrelations of Scale Scores in Four Samples^a

Scale and sample	Scale				
	Expressiveness	Industrious	Analytical	Forceful	Daring
Women					
Industrious					
1972 College		1.00			
1982 College	.31	1.00			
1982 High school	.33	1.00			
1984 College	.25	1.00			
Analytical	.32	1.00			
1972 College					
1982 College	.16	.30	1.00		
1982 High school	.21	.37	1.00		
1984 College	.32	.30	1.00		
Forceful	.14	.44	1.00		
1972 College					
1982 College	.18 ^b	.20 ^b	.21 ^b	1.00	
1982 High school	.17	.41	.12	1.00	
1984 College	.17	.23	.32	1.00	
Daring	-.08	.23	.30	1.00	
1972 College					
1982 College	.13	-.05	.08	.36 ^b	1.00
1982 High school	.15	.22	-.12	.46	1.00
1984 College	.24	.28	.09	.45	1.00
	-.03	.00	.18	.50	1.00
Men					
Industrious					
1972 College		1.00			
1982 College	.23	1.00			
1982 High school	.19	1.00			
1984 College	.26	1.00			
Analytical	.20	1.00			
1972 College					
1982 College	.24	.43	1.00		
1982 High school	.13	.35	1.00		
1984 College	.24	.24	1.00		
Forceful	.27	.47	1.00		
1972 College					
1982 College	.10 ^b	.30 ^b	.24 ^b	1.00	
1982 High school	.18	.45	.27	1.00	
1984 College	.30	.37	.09	1.00	
Daring	.17	.35	.26	1.00	
1972 College					
1982 College	.19	.05	.09	.36 ^b	1.00
1982 High school	.22	.29	.17	.55	1.00
1984 College	.28	.19	.19	.57	1.00
	.14	.25	.20	.53	1.00

^aAll measures are Pearson product moment correlations.

^bIn the 1972 study the term *active* and *independent* formed a third subdimension of autonomous (see footnote 5 in text). The correlations reported with the forceful dimensions for the 1972 sample are an average of those with the "forceful" and "active-independent" subdimension in the 1972 sample.

Gender Differences in the Dimensions

Table V shows, for each of the samples, the means and standard deviations for the men and women, the results for *t* tests between the mean self-ratings, and the effect size of this difference, as measured by Cohen's *d*.

In every group and time period, women rate themselves significantly higher than men on the expressiveness scale. The effect sizes range from .26 to .75, with an average value of .47, generally considered as indicating a moderate effect (Cohen, 1977).

With regard to the two instrumental scales, women in both the 1982 and the 1984 samples of college students and the sample of nurses rated themselves higher on the industrious scale than the comparable men. These differences do not appear in the 1972 college sample or among high school students, although the trend is in the same direction. The effect sizes range from .09 to .36, with an average of .22, generally considered small. While college men in 1972 and 1982 and male nurses reported significantly higher levels of analytic characteristics, high school females had significantly higher scores than high school males on this dimension and there were no gender differences in the 1984 college group. The effect sizes range from $-.38$ to $+.38$, with an average value of $-.09$, considered very small.

With regard to autonomy, there were no significant differences in any of the groups on the forceful dimension and the average effect size of $-.09$ is very small. Among high school students and the nurses, women had lower scores on the adventurous scale.⁷ There were no significant differences for the other samples and the mean of the effect sizes is $-.11$, considered small.

DISCUSSION AND SUMMARY

The results described above confirm our theoretical conceptions of the expressive-instrumental distinction. In addition, the largest and only consistently measured difference between the self-concepts of men and women was not that women saw themselves as less autonomous or as less instrumental than men, but that women reported significantly higher levels of expressive personality traits.

Expressiveness

We have argued that expressiveness, or the concern with the relations among individuals, is a dimension of personality distinct from instrumen-

⁷"Daring" was not included on the nurses' questionnaire, so the difference involves only the item "adventurous."

Table V. Average Scale Scores of Men and Women in Each Sample^a

Scales	Sample											
	1972 College		1982 College		1982 High school		1983 Nurses ^b		1984 College		Effect size ^c	
	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men		
Expressiveness												
<i>X</i>	3.35	3.26	3.45	3.25	3.60	3.33	3.64	3.47	3.49	3.33		
<i>s</i>	0.36	0.34	0.36	0.35	0.35	0.37	0.38	0.43	0.41	0.42		
<i>N</i>	215	195	152	161	117	125	240	186	250	215		
<i>t(p)</i>	2.75	(.006)	4.95	(<.001)	5.79	(<.001)	4.30	(<.001)	4.04	(<.001)	.38	(<.001)
Effect size ^c	.26		.56		.75		.42					
Instrumental												
Industrious												
<i>X</i>	3.11	3.05	3.16	3.02	3.07	3.03	3.58	3.43	3.16	3.04		
<i>s</i>	0.53	0.46	0.48	0.48	0.50	0.44	0.38	0.46	0.55	0.52		
<i>N</i>	213	194	154	170	118	126	240	186	249	216		
<i>t(p)</i>	1.22	(.224)	2.65	(.009)	0.57	(.569)	3.57	(<.001)	2.13	(0.21)		
Effect size	.12		.29		.09		.36					
Analytic												
<i>X</i>	3.04	3.16	3.03	3.17	3.20	3.01	3.32	3.45	3.21	3.24		
<i>s</i>	0.55	0.48	0.47	0.45	0.50	0.50	0.50	0.50	0.48	0.49		
<i>N</i>	209	192	145	173	113	128	237	186	250	216		
<i>t(p)</i>	-2.34	(.020)	-2.64	(.009)	2.94	(.003)	-2.51	(.012)	-0.81	(.419)		
Effect size	-.23		-.30		.38		-.26		-.06			
Autonomy												
Forceful												
<i>X</i>	2.84	2.90	2.89	2.93	2.97	3.05	3.02	3.03	2.86	2.89		
<i>s</i>	-	-	0.40	0.41	0.44	0.42	0.51	0.50	0.46	0.50		
<i>N</i>	196	213	163	167	117	123	237	184	247	208		
<i>t(p)</i>	-	-	-0.81	(.416)	-1.41	(.161)	-0.09	(.929)	-0.66	(.511)		
Effect size	-	-	-.09		-.19		-.02		-.06			
Adventurous												
<i>X</i>	2.90	2.83	3.06	3.05	2.98	3.21	2.88	3.08	3.01	3.04		
<i>s</i>	0.70	0.62	0.61	0.58	0.66	0.58	0.86	0.81	0.64	0.68		
<i>N</i>	219	195	166	175	121	132	242	189	249	216		
<i>t(p)</i>	0.24	(.810)	0.02	(.984)	-2.87	(.004)	-2.40	(.017)	-0.45	(.656)		
Effect size	.11		.02		-.37		-.24		-.05			

^aScores on each scale have been summed and averaged. An average score of 4.0 would mean all respondents had indicated items on the scale were *very true of me*; an average score of 1.0 would mean that all respondents had indicated the items were *very untrue of me*.

^bIn this sample, the expressiveness scale does not include *obliging*; the industrious subscale of instrumental includes *hardwork-ing* instead of *industrious*; the forceful subdimension of the autonomy scale does not include *stern*; and the adventurous sub-dimension does not include *daring*.

^cThe effect size *d* was computed with the formula $\bar{X}_1 - \bar{X}_2/S$ where \bar{X}_1 is the mean in one sample, \bar{X}_2 the mean of the other sample, and *S* is the within-group standard deviation.

tality and autonomy. Factor analyses of males and females in five samples indicate that underlying the self-concept of individuals in all these groups is a dimension that closely conforms to our conceptualization of expressiveness. Our analyses also show that this personality dimension is distinct from instrumentality and autonomy. Research reporting factor analyses of the BSRI and PAQ provides further empirical evidence for the meaningfulness of this dimension. At least ten separate studies have found a factor in these scales that is very similar to our expressiveness dimension (Collins, et al., 1979; Feather 1978; Gaa et al., 1979; Gaudreau, 1977; Kimlicka, Wakefield, & Friedman, 1980; Moreland et al., 1977; Pearson, 1980; Pedhazur & Tetenbaum, 1979; Ruch, 1984; Waters, et al., 1977). For example, Pedhazur and Tetenbaum found a central factor in the BSRI to be made up of the items "affectionate, sympathetic, sensitive to the needs of others, understanding, compassionate, eager to soothe hurt feelings, warm, tender and gentle" (Pedhazur & Tetenbaum, 1979, pp. 1010-1011). Items connoting emotionality or dependency, which are included in the PAQ and BSRI femininity scales (e.g., soft-spoken, gullible, childlike, and does not use harsh language on the BSRI; emotional on the PAQ), generally are not a part of this dimension.⁸ It seems that, in attempting to measure femininity, Bem and Spence et al., have included items related to expressiveness, but a true measure of this concept is distorted by the inclusion of other variables.

In all the groups we have studied, women reported significantly higher levels of expressive traits than men, which suggests that expressiveness is more centrally associated with the female gender role. These findings could be explained by the established theoretical argument that expressiveness, or a relational orientation, is the first interactional mode learned by all children and is thus basic to personality development. Later, as part of learning what it means to be male, boys tend to repress this relational aspect of their personalities and define masculinity negatively as not being feminine. Various versions of this explanation can be found in both the work of psychoanalytic theorists such as Horney (1967a, b), Deutsch (1944-1945), and Chodorow (1974, 1978), as well as that of more general social theorists such as Parsons (1954a, b, 1955) and Hartley (1976). (See Stockard & Johnson, 1979 and 1980, for summaries.) From this perspective it is not surprising that the most consistent sex difference across our samples involves expressiveness. This could reflect males' repression of the relational orientation first learned by both male and female children.

⁸Helmreich, Spence, and Wilhelm (1981) claim that a two-factor oblique solution is the optimal one in their analysis of the masculine and feminine items of the PAQ, but the item "emotional" has a much lower loading on their "femininity" factor than any of the other items (see Table I, p. 1101). In general, it may be that their two-factor solution was to some extent forced upon the data. While they note that a three-factor solution was also obtained (p. 1102), they do not report the results of that analysis.

Instrumentality and Autonomy

Our results, as well as the multivariate analyses of the PAQ and BSRI, suggest that instrumentality and autonomy involve subdimensions and thus are not unitary concepts. Yet it is important to note that we did *not* find that women consistently rate themselves lower on any dimension of these two scales (cf. Frodi, Macauley, & Thome, 1977). For example, in measuring autonomy we found no differences between males and females in any group on the forceful dimension. On the industrious subdimension of instrumental, females had higher scores than males in most of the samples. Except for this subdimension, the mean effect sizes on the instrumental and autonomy subscales are very small. These findings are especially important as instrumentality and autonomy are associated in the public mind with the male gender role and tend to be seen as antithetical to expressiveness. Further, these instrumental and autonomous traits have been incorporated into scales of masculinity developed by both Bem and Spence et al.

It is also important to remember that three of the five samples used here involve college students. The results reported in Table IV are overly affected by this portion of the sample, and it may be important to examine variations in results by age and student status. Consistent results appear across all five groups with the expressive dimension, with females always rating themselves higher than males, and with the forceful subdimension of autonomy, with no gender differences appearing. But sharp differences appear in the results of adults and high school students on the two subdimensions of instrumental and in those of college students and noncollege students on the adventurous subscale of autonomy. Adult females (nurses and the college students) are more likely to rate themselves as industrious (average effect size = .25), and adult males are more likely to see themselves as analytical (average effect size = -.21). Among the high school students there is no significant difference on the industrious scale and girls rate themselves as *more* analytical (see Table IV). On the adventurous subdimension of autonomy no sex differences appear for the college students (average effect size = .03), but in the noncollege group (nurses and high school students) males rate themselves significantly higher than females (average effect size = .30).

Since males appear to see themselves as more autonomous and instrumental at some points in the life cycle, and women see themselves as more autonomous and instrumental at other times, it may be that instrumental and autonomous behaviors are not central aspects of a man's perception of himself as male. The differences we find between males and females may be a result of the social demands placed upon individuals. For instance, adult males may be expected to be analytical and rational, while adolescent males are not. Most families tend to supervise more closely the activities of teenage girls, and adventurousness tends to be especially prized by male peer groups

in adolescence. Adult working women (here represented by nurses) may be especially constrained in the area because of family responsibilities. Both college men and women may have more freedom for "adventurous" behavior. While consistent distinctions between males and females on instrumentality and autonomy may indeed appear in later work, a central point is that such work is needed to assess whether differences exist.

Thus, while our research indicates that a central difference in men's and women's self-concepts is that women typically see themselves as more relational or expressive than men, this does not imply that women are necessarily less instrumental or less independent than men. This finding is exceedingly important in view of the tendency among both researchers and the general public to misunderstand and misdefine the nature of expressiveness. Perhaps because we live in a male-dominant society, the term *feminine* seems to connote deference to males, emotionality, and lack of assertiveness, rather than the more positive qualities that facilitate interpersonal interaction. Previous measures of instrumental-expressiveness are not unidimensional and are based on the very stereotypes that the women's movement has sought to overcome. We believe more is to be learned from measuring actual differences in the self-definitions of men and women than from measuring the individual's conformity to the sex role stereotypes of masculinity and femininity.

Further, it is becoming increasingly clear that the masculinity scales of the BSRI and the PAQ are more highly correlated with measures of psychological health and adjustment than are the corresponding femininity scales or the interactive effects of these scales (i.e., androgyny). Taylor and Hall (1982) summarize these studies by suggesting that it is "primarily masculinity, not androgyny, that yields positive outcomes for individuals in American society" (p. 362). Thus, uncritically examined, the current state of research on gender and personality implies that individual women would maximize their health and happiness by being masculine and forsaking a relational orientation. It may be that it is the conceptualization of the scales that intermixes negative instrumental, dependent, and expressive items in the measure of femininity that accounts for the inability of the femininity scales to correlate with positive outcomes such as mental health and happiness. Moreover, our results would suggest that masculinity, as measured by these scales, may not be a unidimensional concept nor always more descriptive of males than of females.

Measuring Gender Differences

While androgyny and masculinity/femininity scales have been widely critiqued, there seems to have been little effort to develop alternative measures of gender-related characteristics. Thus researchers have continued to

rely on measures that are questionable in their theoretical and empirical grounding. For example, during a five-month period in 1986, 15 of 40 (37.5%) articles in *Sex Roles* relied on the BSRI, PAQ, or some other measure of masculinity and femininity based on stereotyped images of males and females.⁹

Marsh and Myers (1986, p. 424) have noted that a new approach for measuring gender-related personality characteristics requires that we "adamantly reject the atheoretical, empirical approach often used to develop M and F scales. Instead, an explicit theoretical model should be the starting point for instrument construction." This is what we have done in our research. The theoretical work presented above suggest that males and females differ in relationality or expressiveness, and in instrumentality, and that it is important to distinguish these two dimensions from autonomy or independence. Our empirical analyses suggest that our scales represent these three dimensions, but that only the one tapping expressiveness or relationality is unidimensional. Moreover, our findings suggest that the basic, and perhaps only single dimension along which groups of males and females consistently differ, through time and across groups, is expressiveness, or a relational orientation, with men less relational than women. *We suggest that this dimension is the one researchers should focus upon in examining gender difference.*

Clearly the further development and validation of our instrument, or the development of alternative ways of measuring these constructs, is needed. While there may well be better ways to measure the relational dimension, at this time researchers might rely on the items on our expressiveness scales, or on the items on the factor from the BSRI identified by Pedhauzer and Tetenbaum as reflecting "interpersonal sensitivity" (compassionate, gentle, warm, tender, affectionate, sympathetic, sensitive to the needs of others eager to soothe hurt feelings, and understanding), or the relevant items on the PAQ's femininity scale (able to devote self completely to others, gentle, helpful, kind, aware of feelings, understanding, warm).¹⁰ Essentially we advocate a shift in the measurement of gender differences, which would focus on relationality or expressiveness. We believe this could lead to a redefinition of the research agenda in gender and personality, allowing us to more accurately measure correlates of traits that theoretically and empirically dis-

⁹Of the 40 articles in the January–May 1986 issues of *Sex Roles*, 8 used the BSRI, 3 used the PAQ, and 4 used other measures that are based on stereotypes of males and females.

¹⁰An instrumental factor is typically not found in Bem's instrument, probably because it includes few items that reflect this dimension of self-concept. In Parsons's and Bales's (Parsons et al., 1954) terminology, there might be instrumental leaders and expressive leaders, thus leadership, the other area tapped by her masculinity scale, is a quality not correlated with either the instrumental or expressive dimension per se and it is thus understandable that it would load on an "assertive" factor.

tinguish males and females. It would also allow researchers to address questions concerning the consequences for individuals and society of people defining themselves in terms of more positive aspects of female traits than those measured by traditional femininity scales.¹¹

REFERENCES

- Armor, D. J. Theta reliability and factor scaling. In H. L. Costner (Ed.), *Sociological methodology*, 1973-74. San Francisco: Jossey-Bass, 1974.
- 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. Sex-role adaptability: One consequence of psychological androgyny. *Journal of Personality and Social Psychology*, 1975, 31, 634-643.
- Bem, S. L., Martynne, W., & Watson, C. Sex typing and androgyny: Further explorations of the expressive domain. *Journal of Personality and Social Psychology*, 1976, 34, 1016-1023.
- Berzins, J. I., Welling, M. A., & Wetter, R. E. A new measure of psychological androgyny based on the Personality Research Form. *Journal of Consulting and Clinical Psychology*, 1978, 46, 126-138.
- Chodorow, N. Being and doing: A cross-cultural examination of the socialization of males and females. In V. Gornick & B. K. Moran (Eds.), *Woman in sexist society*. New York: New American Library, 1974.
- Chodorow, N. *The reproduction of mothering*. Berkeley: University of California Press, 1978.
- Cohen, J. *Statistical power analysis for the behavioral sciences*. New York: Academic Press, 1977.
- Collins, M. Waters, C. W., & Waters, L. K. Factor analysis of sex-typed items from the Bem Sex-Role Inventory: A replication. *Psychological Reports*, 1979, 44, 517-518.
- Constantinople, A. Masculinity-femininity: An exception to the famous dictum? *Psychological Bulletin*, 1973, 80, 389-407.
- Costantino, F., & Heilbrun, A. B. Anxiety correlates of sex-role identity in college students. *Psychological Reports*, 1964, 14, 729-730.
- Deutsch, H. *The psychology of women: A Psychoanalytic interpretation. Vol. 1. Girlhood*. New York: Bantam, 1944-1945.
- Eisler, R., Hersen, M., & Miller, P. Situational determinants of assertive behaviors. *Journal of Consulting and Clinical Psychology*, 1975, 91, 31-37.
- Feather, N. T. Factor structure of the Bem Sex-Role Inventory: Implications for the study of masculinity, femininity and androgyny. *Australian Journal of Psychology*, 1978, 30, 241-254.
- Frodi, A., Macaulay, J., & Thome, P. Are women always less assertive than men? A review of the experimental literature. *Psychological Bulletin*, 1977, 84, 634-660.
- Gaa, J. P., Liberman, D., & Edwards, T. A. A comparative factor analysis of the Bem Sex-Role Inventory and the Personality Attributes Questionnaire. *Journal of Clinical Psychology*, 1979, 35, 592-598.
- Gaudreau, P. Factor analysis of the Bem Sex-Role Inventory. *Journal of Consulting and Clinical Psychology*, 1977, 45, 299-302.
- Gilligan, C. *In a different voice*. Cambridge: Harvard University Press, 1982.
- Gough, H. G. Identifying psychological femininity. *Educational and Psychological Measurement*, 1952, 12, 427-439.
- Gough, H. G., & Heilbrun, A. B. *Joint manual for the Adjective Check List and the Need Scales for the ACL*. Palo Alto, CA: Consulting Psychologists Press, 1965.

¹¹Volmer (1986) provides an example of the use of the PAQ to measure instrumentality and expressiveness.

- Gutmann, D. Female ego styles and generational conflict. In J. Bardwick et al. (Eds.), *Feminine personality and conflict*. Belmont, CA: Brooks/Cole, 1970.
- Hartley, R. E. Sex-role pressures and the socialization of the male child. In D. S. David & R. Brannon (Eds.), *The forty-nine percent majority: The male sex role*. Reading, MA: Addison-Wesley, 1976. [Originally published 1959.]
- Hathaway, S. R., & McKinley, J. C. *The Minnesota Multiphasic Personality Inventory*. Minneapolis: University of Minnesota Press, 1943.
- Heerboth, J. R., & Ramanaiah, N. V. Evaluation of the BSRI masculine and feminine items using desirability and stereotype ratings. *Journal of Personality Assessment*, 1985, 49, 264-270.
- Heilbrun, A. B., Jr. Measurement of masculine and feminine sex role identities as independent dimensions. *Journal of Consulting and Clinical Psychology*, 1976, 44, 183-190.
- Helmrich, R. L., Spence, J. T., & Wilhelm, J. A. A psychometric analysis of the Personal Attributes Questionnaire. *Sex Roles*, 1981, 7(11), 1097-1108.
- Horney, K. The dread of women: Observations on a specific difference in the dread felt by men and by women respectively for the opposite sex. In H. Kelman (Ed.), *Feminine psychology*. New York: W. W. Norton, 1967. (a) [Originally published 1932.]
- Horney, K. The flight from womanhood: The masculinity complex in women as viewed by men and by women. In H. Kelman (Ed.), *Feminine psychology*. New York: W. W. Norton, 1967. (b) [Originally published 1926.]
- Johnson, M. M. *Instrumental and expressive components in the personalities of women*. Dissertation, Radcliffe College, 1955.
- Johnson, M. M. Sex role learning in the nuclear family. *Child Development*, 1963, 34, 319-333.
- Johnson, M. M., Stockard, J., Acker, J., & Naffziger, C. Expressiveness reevaluated. *School Review*, 1975, 83, 617-644.
- Jones, W. H., Chernovetz, M. E., & Hansson, R. O. The enigma of androgyny: Differential implications for males and females? *Journal of Consulting and Clinical Psychology*, 1978, 46, 298-313.
- Kelly, J. A., & Worell, J. New formulations of sex roles and androgyny: A critical review. *Journal of Consulting and Clinical Psychology*, 1977, 45, 1101-1115.
- Kimlicka, T. M., Wakefield, J. A., Jr., & Friedman, A. F. Comparison of factors from the Bem Sex Role Inventory for male and female college students. *Psychological Reports*, 1980, 46, 1011-1017.
- Kuecker, L. *Men in nursing*. Dissertation, University of Oregon, 1985.
- Lewin, M. Psychological measures of femininity and masculinity, 2. In M. Lewin (Ed.), *In the shadow of the past: Psychology portrays the sexes*. New York: Columbia University Press, 1984.
- Locksley, A., & Colten, M. E. Psychological androgyny: A case of mistaken identity. *Journal of Personality and Social Psychology*, 1979, 37, 1017-1031.
- Lott, B. A feminist critique of androgyny: Toward the elimination of gender attributions for learned behavior. In C. Mayo & N. M. Henley (Eds.), *Gender and nonverbal behavior*. New York: Springer, 1981.
- Marsh, H. W., & Myers, M. Masculinity, femininity, and androgyny: A methodological and theoretical critique. *Sex Roles*, 1986, 14, 397-430.
- Moreland, J. R., Gulanick, N., Montague, E., & Harren, V. Some psychometric properties of the Bem Sex-Role Inventory. *Applied Psychological Measurement*, 1978, 2, 249-256.
- Morgan, K. P. Androgyny: A conceptual critique. *Social Theory and Practice*, 1982, 8, 245-283.
- Myers, A., & Sugar, J. A. A critical analysis of scoring the BSRI: Implications for conceptualization. *JSAS Catalog of Selected Documents in Psychology*, 1979, 9, 24. (Ms. No. 1833.)
- Myers, A. M., & Gonda, G. Empirical validation of the Bem Sex-Role Inventory. *Journal of Personality and Social Psychology*, 1982, 43, 304-318.
- Parsons, T. *The social system*. Glencoe, IL: Free Press, 1951.
- Parsons, T. Age and sex in the social structure of the United States. In T. Parsons (Ed.), *Essays in sociological theory*, rev. ed. Glencoe, IL: Free Press, 1954. (a)
- Parsons, T. Certain primary sources and patterns of aggression in the social structure of the western world. In T. Parsons (Ed.), *Essays in sociological theory*, rev. ed. Glencoe, IL: Free Press, 1954. (b)

- Parsons, T. Family structure and the socialization of the child. In T. Parsons & R. F. Bales, (Eds.), *Family socialization and interaction process*. Glencoe, IL: Free Press, 1955.
- Parsons, T., & Shils, E. A. (Eds.), *Toward a general theory of action*. Cambridge: Harvard University Press, 1952.
- Parsons, T., & Smelser, N. *Economy and society*. New York: Oxford University Press, 1956.
- Parsons, T., Bales, R. F., & Shills, E. *Working papers in the theory of action*. Glencoe, IL: Free Press, 1954.
- Pearson, J. C. A factor analytic study of the items in three selected sex-role instruments. *Psychological Reports*, 1980, 41, 1119-1126.
- Pedhazur, E., & Tetenbaum, T. Bem Sex Role Inventory: A theoretical and methodological critique. *Journal of Personality and Social Psychology*, 1979, 37, 996-1016.
- Ruch, L. O. Dimensionality of the Bem Sex-Role Inventory: A multidimensional analysis. *Sex Roles*, 1984, 10, 99-117.
- Spence, J. T., & Helmreich, R. L. *Masculinity*. Austin: University of Texas Press, 1978.
- Spence, J. T., Helmreich, R. L., & Stapp, J. The Personal Attributes Questionnaire: A measure of sex-role stereotypes and masculinity-femininity. *JSAS Catalog of Selected Documents in Psychology*, 1974, 4, 43.
- Spence, J. T., Helmreich, R. L., & 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.
- Stark-Adamec, C., Graham, M., & Pyke, S. Androgyny and mental health: The need for a critical evaluation of the theoretical equation. *International Journal of Women's Studies*, 1980, 3, 490-507.
- Stockard, J., & Johnson, M. M. The social origin of male dominance. *Sex Roles*, 1979, 5, 199-218.
- Stockard, J., & Johnson, M. M. *Sex roles*, Englewood Cliffs, NJ. Prentice-Hall, 1980.
- Taylor, M. C., & Hall, J. A. Psychological androgyny: Theories, methods, and conclusions. *Psychological Bulletins*, 1982, 92, 347-366.
- Terman, L. L., & Miles, C. C. *Sex and personality*. New York: McGraw-Hill, 1936.
- Vollmer, F. Why do men have higher expectancy than women? *Sex Roles*, 1986, 14, 351-362.
- Waters, C. W., Waters, L. K., & Pincus, S. Factor analysis of masculine and feminine sex-typed items from the Bem Sex-Role Inventory. *Psychological Reports*, 1977, 1977, 40, 567-570.
- Zeller, R. A., & Carmines, E. G. *Measurement in the social sciences: The link between theory and data*. Cambridge: Cambridge University Press, 1980.