



8-1-1979

Self-Esteem Effects in the Measurement of Sex-Role Orientation

Steven H. Weaver

Follow this and additional works at: <https://commons.und.edu/theses>

Recommended Citation

Weaver, Steven H., "Self-Esteem Effects in the Measurement of Sex-Role Orientation" (1979). *Theses and Dissertations*. 2519.
<https://commons.und.edu/theses/2519>

This Thesis is brought to you for free and open access by the Theses, Dissertations, and Senior Projects at UND Scholarly Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UND Scholarly Commons. For more information, please contact zeinebyousif@library.und.edu.

SELF-ESTEEM EFFECTS IN THE MEASUREMENT
OF SEX-ROLE ORIENTATION

by
Steven H. Weaver

Bachelor of Arts, University of Kansas, 1972

A Thesis
Submitted to the Graduate Faculty
of the
University of North Dakota
in partial fulfillment of the requirements
for the degree of
Master of Arts

Grand Forks, North Dakota

August
1979

This Thesis submitted by Steven H. Weaver in partial fulfillment of the requirements for the Degree of Master of Arts from the University of North Dakota is hereby approved by the Faculty Advisory Committee under whom the work has been done.

Paul H. Wright
(Chairman)

James A. Carl
John P. Rebo

This Thesis meets the standards for appearance and conforms to the style and format requirements of the Graduate School of the University of North Dakota, and is hereby approved.

A. William Johnson
Dean of the Graduate School

Permission

Title Self-Esteem Effects in the Measurement of Sex-Role
Orientation

Department Psychology

Degree Master of Arts

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the Library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work or, in his absence, by the Chairman of the Department or the Dean of the Graduate School. It is understood that any copying or publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

Signature Steven H. Wear

Date July 26, 1979

TABLE OF CONTENTS

| | |
|--|-----|
| LIST OF TABLES..... | v |
| ACKNOWLEDGMENTS..... | vi |
| ABSTRACT..... | vii |
| CHAPTER I. INTRODUCTION..... | 1 |
| CHAPTER II. BACKGROUND AND HYPOTHESES..... | 4 |
| CHAPTER III. METHODS..... | 49 |
| CHAPTER IV. RESULTS..... | 56 |
| CHAPTER V. DISCUSSION AND CONCLUSIONS..... | 77 |
| APPENDIX..... | 84 |
| APPENDIX: RAW SCORES ON THE BEM SEX-ROLE INVENTORY AND THE TEXAS SOCIAL BEHAVIOR INVENTORY FOR 212 FEMALE AND 165 MALE SUBJECTS..... | 85 |
| REFERENCES..... | 95 |

LIST OF TABLES

| Table | Page |
|---|------|
| 1. Correlation Coefficients among Groups Rating 46 Cartoons used in Non-critical Trials of the Conformity Induction..... | 59 |
| 2. Distribution of 165 Male and 212 Female Undergraduates By Sex-Role Category and Level of Self-Esteem.... | 61 |
| 3. Distribution of Conformity Induction Participants by Sex-Role Category and Level of Self-Esteem..... | 62 |
| 4. Correlation Matrices of the Questionnaire Variables for Males and Females..... | 63 |
| 5. Multiple Regression: Females; Predicting Self-Esteem from Masculinity, Femininity, Social Desirability, and the Masculinity-Femininity Product.... | 65 |
| 6. Multiple Regression: Females; Predicting Self-Esteem from Masculinity and Social Desirability..... | 66 |
| 7. Multiple Regression: Males; Predicting Self-Esteem from Masculinity, Femininity, Social Desirability, and the Masculinity-Femininity Product..... | 67 |
| 8. Multiple Regression: Males; Predicting Self-Esteem from Masculinity and Social Desirability..... | 68 |
| 9. Multiple Regression: Females; Predicting Independence from Masculinity, Femininity, M x F, and Self-Esteem..... | 71 |
| 10. Multiple Regression: Males; Predicting Independence from Masculinity, Femininity, M x F, Self-Esteem, and Tape Number..... | 73 |
| 11. Multiple Regression: Males and Females; Predicting Independence from Sex and Tape Number..... | 75 |

ACKNOWLEDGMENTS

I would like to express my gratitude to the students of Psychology who participated as subjects in this research. For their encouragement and supervision, I am indebted to the members of my advisory committee: Dr. Paul Wright, Dr. James Clark, and Dr. Lila Tabor. Each made a singular and invaluable contribution to the culmination of this research.

ABSTRACT

In a replication of a study which demonstrated a relationship between Masculinity and Independence, the nature of the relationship between Self-esteem and orthogonal Masculinity and Femininity dimensions was examined. Regression analyses for 212 female and 165 male undergraduates indicated that for both sexes, Masculinity was moderately correlated to Self-esteem, but that Femininity did not add to this prediction significantly. Scores for Independence were obtained from a subset of 24 females and 29 males by placing them in a conformity paradigm situation. The major hypothesis was that Masculinity would not contribute uniquely to the prediction of Independence made on the basis of Self-esteem alone. Regression analyses indicated that neither Self-esteem nor Masculinity were predictive of Independence scores for either sex. No conclusions could therefore be drawn about the precise nature of the relationship between Masculinity and Self-esteem. Consideration is given to methodological departures in the replication and the advisability of using a conformity paradigm as a criterion behavior for masculinity.

CHAPTER I

INTRODUCTION

Supposedly, everyone knows what masculine means and what feminine means. Given that, the process of identifying these qualities systematically so that their impact on behavior can be studied has not proved to be as easy as might have been expected.

The importance of continuing to try to identify and describe sex-role orientation and its impact on the social behavior of individuals cannot be underestimated. Society is undergoing a persistent and large-scale re-examination of its underlying assumptions, prescriptions, and prejudices regarding the behavioral norms which are subtly imposed on its male and female members. In the face of traditional pressures to adopt polarized traditional role values, ever-increasing numbers of people are seeking out and enacting alternative forms of sex-role behavior. Many feel these more flexible behaviors are actually more congruent with their internal dispositions. At the same time, employment, social activities, and heterosexual dating patterns serve as examples of other spheres where old traditions die hard and in some cases are becoming more entrenched. Paradoxes occur; it has been argued, for example, that the increasing number of females obtaining Ph.D's signals not an advance for women but a decline of the importance and usefulness of advanced academic credentials. In this mass of contradiction and conflict it is difficult for the

informed observer to be complacent about the extent of our current knowledge of the forms and the uses of sex-role related behaviors.

The role of psychology in this process of societal self-scrutiny should be a fundamental one. Admittedly, the questions are large ones. Nevertheless, whether reformer or conservative, the psychological researcher has the opportunity to participate in this debate by providing first, some understanding of the significance of sex-role behavior, and second, informed predictions about the possible outcomes implicit in the choices which society will make, with or without scientific input.

Fundamental to the effort to understand concepts like masculinity, femininity, and sex-role orientation in general, is the development of appropriate measures. Considerable attention has been focused recently on the measurement of masculinity and femininity with particular emphasis on the reconceptualization of these two sex-role variables as separate and uncorrelated dimensions. As a result, a number of investigators have developed questionnaire measures of sex-role orientation incorporating this orthogonal conception (Kelly and Worell 1977).

Using a variety of sex-role and self-esteem measures, several different investigators have reported moderate correlations between Masculinity scores and Self-esteem scores (Spence, Helmreich, & Stapp 1975; Bem 1977; Fay & Brown, 1979; Wetter 1975). The dearth of behavioral validation for these sex-role inventories and the structural similarities between them and the self-esteem measures invite speculation that the strength of this relationship is an artifact of the construction of the sex-role inventories themselves rather than a

true description of the relationships between the theoretical constructs. Underscoring this speculation is the counterintuitive finding that femininity is less highly correlated to self-esteem than masculinity in both sexes, and in some studies has shown no relationship to self-esteem whatsoever.

The study reported here specifically attempted to answer the question of the relative importance of Masculinity scores to Self-esteem scores in the prediction of performance within a particular experimental situation. The situation--an Asch-type conformity experiment--was chosen to represent a behavioral example of masculinity by Bem (1975). The present study attempted to replicate in all essential regards the earlier study made by Bem. The object of the replication was to find out if in fact scores on the Masculinity scale of the Bem Sex-Role Inventory (Bem 1974) were better overall indicators of masculine performance than Self-esteem scores. If they were not, support would be lent to the interpretation that the Masculinity scale was little more than a poorly constructed measure of self-esteem, and the importance and validity of this and other similar measures of masculinity would have been seriously jeopardized.

CHAPTER II

BACKGROUND AND HYPOTHESES

Historical Overview

The Critique of the Unidimensional Paradigm in M-F Measurement

In a tough-minded review of the major Masculinity-Femininity (M-F) measures, Constantinople (1973) posed the critical issues which set the stage for a revolution in the methodology employed in this area. These issues ranged from conceptual to empirical. Six of the major criticisms bear review in this context because they give shape to the arguments favoring the development of the major sex-role orientation measures in the 1970's. These six issues were: a) the assumption that M-F conforms to a bipolar, unidimensional model, b) the use of atheoretical methodologies to study M-F, c) the use of sex differences as the criterion of M-F discrimination, d) the confounding of M-F with gender identification, or with sex-role identification, e) the impact of item subtlety, perceived social desirability, and stereotyping on M-F response sets, and f) the inability of M-F measures to reliably predict behavior. These issues are considered separately below.

The Bipolar, Unidimensional Model

Constantinople treated the issues of unidimensionality and bipolarity as conceptually distinct although in practical terms they are

intertwined. Taken together, they describe a model in which masculinity and femininity are considered as polar opposites defining the endpoints of a continuum of values. A single value on that continuum is assumed to represent the degree to which a particular respondent is masculine/feminine relative to all other individuals.

It was suggested that the question of unidimensionality could be approached in two ways. The first of these incorporates the concern for the validity of the bipolarity assumption by asking if Masculinity and Femininity might not be two separable dimensions which vary independently. Implicit in this alternative view is a criticism of the form of the items in M-F measurement where often a single item will be scored either masculine or feminine depending upon a binary response. The evidence cited by Constantinople for the possibility of two separate dimensions is suggestive at best. She notes that as of that time no attempts had been made to measure M-F as separate dimensions. However, it will be seen that considerable evidence has accumulated since the time of her review in support of the consideration of Masculinity and Femininity as separate theoretical constructs.

The conceptual assumption of unidimensionality could also be questioned, according to Constantinople, on the basis of the content of the items or procedures used to measure M-F. There was no substantive evidence that any of the existing measures of M-F could be said to measure a unitary trait, and as a result, the estimate of M-F for any particular individual could vary depending upon the behaviors sampled. This argument was clearly supported by the factor-analytic work of Lunneborg (Lunneborg & Lunneborg 1970; Lunneborg 1972).

Although the newer, non-bipolar measures of sex-role orientation are much "cleaner" in factor-analytic terms (Whetton & Swindells 1977; Gaudreau 1977), a similar argument concerning the content of the instrument may be leveled at them. For example, although the Bem Sex-Role Inventory (BSRI; Bem 1974) is more limited in the behaviors it samples and is therefore psychometrically more sound than its predecessors, it may be sampling too narrow a range to be of empirical or practical use in the prediction of behavior. This is an issue which awaits further experimental investigation.

Parallel criticisms of the assumption of unidimensionality are to be found in discussions of the issue by Pleck (1975), Tyler (1968), and Edwards and Abbott (1973, p. 248).

Just as Constantinople's discussion of unidimensionality was interpenetrated by the assumption of bipolarity, the discussion of bipolarity alluded back to the problem of unidimensionality. Constantinople described three basic aspects of the bipolarity assumption which characterized the measures of M-F under analysis. The first of these aspects was the use of a single score to represent the total of an individual's responses which placed that individual somewhere on a continuum ranging from Feminine at one endpoint through a neutral zero-point to Masculine at the other endpoint. The second aspect was the use of the ability of an item to discriminate the biological sexes as a method of item selection. This implies that a continuous variable, M-F, is validated by a dichotomous variable, sex of respondent, which seems methodologically questionable. Third, the bipolarity assumption was seen in the use of logical reversal or inverse correlation between masculine and feminine. What was not

masculine was therefore feminine and vice versa. Again, this was especially true where a single item requiring a dichotomous response was scored in one direction or the other.

Constantinople contended that the two assumptions of bipolarity and unidimensionality had not been tested for validity before they were applied in the measurement of M-F. The criticism of this type of model for use in the study of psychological sex-roles was echoed by Bem (1974; 1976; 1972) Spence, Helmreich, & Stapp (1975), and Spence and Helmreich (1978, p. 17). The revision of this model in favor of a dualistic approach has since received widespread support and acceptance in the literature.

Atheoretical Methodologies

The second issue advanced by Constantinople to be considered here is that of the lack of a theoretical base for the construction of measures of M-F. Noting that the terms masculinity and femininity "seem to be among the muddiest concepts in the psychologist's vocabulary" (1973, p. 390), she observed further that they seemed to have been taken over from the public domain for scientific purposes without systematic definition or explication.

The most generalized definitions of the terms as they are used by those developing tests of M-F would seem to be that they are relatively enduring traits which are more or less rooted in anatomy, physiology, and early experience, and which generally serve to distinguish males from females in appearance, attitudes, and behavior (1973, p. 390).

Empirical definitions varied widely with the exception of the use or partial use of sex differences as a criterion for item selection.

Most objectionable were those approaches which were strictly empirical and therefore unrelated to any abstract definition.

Sex Differences as the Criterion for M-F Measurement

The third issue raised by Constantinople was subsidiary to the previous issues of bipolarity and lack of a theoretical base. This issue concerns the problem of using biological sex as the basis for Masculinity-Femininity discrimination. This point is mentioned separately to re-emphasize its possible implication that item selection may be arbitrarily tied to differences existing in a particular culture at a particular time.

Confounds with Masculinity-Femininity

A fourth issue is the confounding of a construct called Masculinity-Femininity with related constructs such as sex-role identification and gender identification. In some cases, assumptions have even been made about a relationship between M-F scores and sexual preference, which is insupportable given the weakness of M-F measures in regard to reliability and inter-measure correlations. Clearly, this issue is also related to the lack of a guiding definition for the M-F construct.

Stereotypy, Social Desirability, and Item Subtlety

Citing the work of Nichols (1962) and Lunneborg (1970), Constantinople voices a concern with the effects of stereotypy, social desirability, and item subtlety on M-F measurement. Any of these can have an impact on the accuracy of M-F measurement or appraisal: (a) the degree to which an item represents a widely held stereotyped view of sex differences; (b) the degree to which a respondent desires or is willing to acknowledge a sex-role-related behavior or trait as applying

to him or herself; or (c) the degree to which an item is obvious. It is not clear, however, how extensive or important the interactions of these variables may be.

The Prediction of Behavior from M-F Scores

Finally, Constantinople alleges that there is not a ". . . body of data which indicates that M-F, or M, or F alone, consistently is related to other variables in predictable ways (except whether or not the subject is male or female!)." (1973, p. 389). This echoes the criticism of Lunneborg that

[M]uch of the MF literature, i.e., studies that have explored the relationships of masculinity-femininity to school achievement, sex-role identification, occupations, homosexuality, field independence, creativity, etc., must be interpreted very guardedly, for what exactly did the measure of MF that was employed actually measure? (Lunneborg 1972, p. 316).

Clearly it is incumbent on the advocates of a measure of M-F or sex-role orientation to provide substantive evidence of the empirical effectiveness of the measure.

The six criticisms which have been presented form the foundation upon which sex-role measures of a much different sort have been constructed. Any defensible attempt at sex-role measurement must provide a theoretical understanding of the concepts of masculinity and femininity and show how the operations employed are related in a systematic way to that definition. In addition, procedural issues such as the influences of stereotyping and item subtlety must be given consideration. Finally, no measure can be considered adequate that has not been shown to predict behavior in some reliable and

theoretically satisfying fashion. Until these conditions are satisfied any measure can at best be considered tentative and hypothetical in nature.

Current Measurement Approaches

Four major questionnaires have been developed in the 1970's which incorporate a dualistic view of masculinity and femininity. Two are primarily original and two were extracted from larger, established personality instruments. The two original approaches are the Bem Sex-Role Inventory (BSRI) developed by Bem (1974) and the Personal Attributes Questionnaire (PAQ) developed by Spence, Helmreich, and Stapp (1974, 1975). The two measures which were outgrowths of other scales are the ANDRO scale developed by Berzins, Welling and Wetter (1978) from the items of the Personality Research Form (Jackson 1967), and separate Masculinity and Femininity subscales taken from the Adjective Check List (Heilbrun 1976). These last two measures present evidence of their psychometric soundness and external validation. However, the discussion to follow will focus on the two original measures, since these are the most sophisticated pure measures of sex-role orientation that are currently available, and are the standards to which the derivative scales are compared.

A comment needs to be made about the relationship of the PAQ and the BSRI. These are parallel but not identical measures. They are treated together for the purpose of the discussion that follows but the relevant similarities and differences of the two approaches will be described as they pertain to the specific issues under consideration. At present, the use of one instrument rather than the

other may be largely a matter of the personal preferences of the investigator rather than a clear-cut discrimination arising from the accumulation of empirical support for the use of one instead of the other. The view presented here is that the approach taken toward rectifying the mistakes of past M-F measures is essentially the same in both cases. The differences that do exist do not preclude the consideration of the literature pertaining to one as relevant to the other.

The most obvious way in which both the BSRI and the PAQ are distinguished from earlier measures of M-F is that they employ an underlying assumption that masculinity and femininity are orthogonal psychological constructs. In each case Masculinity and Femininity are completely separate and independent scales composed of two completely different pools of items. Both instruments also contain a third scale. The nature of this third scale is quite different for the two instruments. In its earliest form, the BSRI produced three scores--Masculinity, Femininity, and Social Desirability. A fourth score called "Androgyny" which represented the difference between the M and F scores could also be computed. The use of this Androgyny score has since been discontinued (Bem 1977). On the BSRI, the M and F scales are each composed of twenty positive adjectives or adjective phrases. The so-called Social Desirability scale is composed of ten positive and ten negative sex-neutral adjectives.

The third scale on the PAQ, in contrast to the BSRI, is an empirically derived bipolar Masculinity-Femininity subscale, which therefore consolidates in the PAQ both the bipolar and dualistic models

of M-F with the expectation that the bipolar scale contributes to the predictive ability of the orthogonal subscales (Spence & Helmreich 1978, p. 20). The discovery that the independence of Masculinity and Femininity was not only theoretically plausible but empirically possible was an important step. It would seem to clearly satisfy the objections to the lack of true unidimensionality in the earlier measures of M-F.

A second important aspect of the development of these two scales was the major shift in conceptual definition which was utilized by their authors. The implicit notion of M and F as sex-linked traits attributed by Constantinople to earlier M-F measures has been discarded in favor of what Spence and Helmreich have called "a theoretical approach to psychological phenomena, as opposed to a purely empirical one." (1978, p. 14) This approach is congruent with current trends in social psychology away from the expectation that a psychological disposition, such as an attitude, will correspond highly to a specific behavior, and toward the view that such psychological dispositions are important variables insofar as they are predictive across a number of behavioral situations. The proposition that the sex-role variables measured by the PAQ cannot be expected to demonstrate more than very weak correlations with real-world or laboratory behaviors would almost seem to insulate these psychological constructs from empirical verification. At the outset of their book, Spence and Helmreich note, "[W]e will argue that, at least in contemporary society, these psychological dimensions are only weakly related within each sex to the broad spectrum of sex-role behaviors."

(Spence & Helmreich 1978, p. 3).

What is the nature of the constructs "masculinity" and "femininity" if they are not traits and are not closely related to behavior? They are variously referred to by Spence and Helmreich as "self-variables," "attributes," and "inner characteristics of the individual," among other descriptors. The crucial distinction is between role-related behaviors and masculine/feminine personality characteristics; between the sex-role behaviors which may vary with the situation, and the properties of the behaving individual (1978, pp. 14-15). The following definition is provided: Masculinity and Femininity are ". . . clusters of socially desirable attributes stereotypically considered to differentiate males and females and thus to define the psychological core of masculine and feminine personalities" (Spence & Helmreich 1978, p. 3).

Since in experimental psychology a concept is defined by its operations, it is instructive to describe the items of the PAQ. The PAQ is composed of 24 items. The 24 items on the PAQ are a refinement of the 55-item original form. Subjects are asked to describe themselves on a five-point scale, the endpoints of which are bipolar descriptions of personality characteristics: for example, a masculine-valued item is "Very independent" v. "Not at all independent." Items were chosen first on the basis that raters of both sexes judged the typical male and the typical female to differ significantly on that dimension. Assignment to the individual subscales depended on ratings by separate judges regarding the attribution of each descriptor to the ideal male or ideal female. For example, where the mean

rating of the ideal female and the mean rating of the ideal male lay on the same side of the scale midpoint as the mean rating of the typical female, the item was assigned to the feminine-valued (F) subscale.

Consequently, masculine items were those describing characteristics which are desirable for either sex (ideal ratings on the same side of the midpoint) but are felt by college raters to be more typical of males than females. Feminine items were those describing characteristics which are desirable for either sex but are felt to be more typical of females. Finally, items were assigned to the M-F subscale when the ratings for the ideal male and ideal female fell on opposite sides of the scale midpoint.

In addition to the desirability ratings, self-ratings were also of importance in the creation of the PAQ so that items on the M scale, for example, represent not only attributes that both males and females agree are more typical of males, but also only those attributes which were significantly more highly endorsed by male self-raters than by female self-raters.

Bem apparently did not agree with the assertion made by Spence and Helmreich that the psychological nature of sex-role orientation made it unlikely that behavior could be accurately predicted from M or F scores, since she attempted to demonstrate a linkage between sex-role orientation and behavior in a series of experiments. Although she has not articulated a theoretical position as elaborately as Spence and Helmreich, it is clear that she regards psychological masculinity and femininity as aspects of self-concept, therefore as cognitive entities or beliefs rather than as traits inherent in the

personality structure. According to Bem, ". . . [T]he BSRI was founded on a conception of the sex-typed person as someone who has internalized society's sex-typed standards of desirable behavior for men and women" (1974, p. 155). The BSRI was created by asking respondents to rate a set of personality characteristics for their appropriateness for a man in our society or for a woman. Items were selected which both males and females agreed were more appropriate for one sex than the other. This differs from the rationale of the PAQ and yet the contents of the two scales noticeably overlap.

Though the rationale for item inclusion differs between the PAQ and the BSRI, in both cases an attempt is made to define M and F in terms of the conceptions of those entities that are shared by both males and females. This clearly distinguishes these instruments from others based merely on empirically observed differences between males and females in our particular society. What emerged in the process of developing the BSRI and the PAQ was that the nature of these more or less universal beliefs about masculinity and femininity conformed to previously articulated theoretical understandings. The common threads tying together masculine items on the one hand, and feminine items on the other were similar to or at least congruent with theoretical explanations of masculine and feminine behavior available in the literature of social science. The authors of both the PAQ and the BSRI have appealed to the theoretical formulations of Parsons and Bales (1955) and Bakan (1966), who respectively characterize the masculine dimension as "instrumental" or "agentic." These terms emphasize the concern for goal-attainment and accomplishment. The feminine

dimension, in contrast, is seen to be "expressive" or "communal," that is, concerned with emotion, interpersonal relationships, and so forth. The basic thesis of the dualistic model then, is that these two separate qualities are not mutually exclusive but can be and perhaps should be complementary. The authors of the PAQ and BSRI have argued that the item content of the M scales represent prima facie an instrumental-agentic orientation; likewise, the F scales are said to reflect expressive-communal or nurturant concerns.

Reports of factor analyses of the BSRI lend some support to the validity of the author's contentions regarding the underlying theoretical distinctions between M and F. Both Gaudreau (1977) using a sample of American adults and Whetton and Swindells (1977) using a sample of British students found that masculinity and femininity items loaded on separate factors. Gaudreau simply called the first and second factors Masculinity and Femininity respectively. The analysis of the highly loaded items supports the description of Factor I as instrumental-agentic, and of Factor II as expressive-communal. Highly loaded M items, for example, were "Has leadership abilities," "strong personality," "forceful," and "assertive." Highly loaded items on the Feminine factor were "Compassionate," "Eager to soothe hurt feelings," and "Sympathetic." Whetton and Swindells found that five major factors accounted for 17% of the variance in the BSRI. Of these five, two were composed primarily of items on the M scale which they labeled "Power" and "Autonomy" and one which was composed primarily of F items, labeled "Empathy." The difference in the labeling of the factors produced by the authors of these two

articles provokes speculation that perhaps it is more valid to regard these dimensions as somewhat less global than the titles Masculinity and Femininity would suggest. Nevertheless these factor analyses tend to be supportive rather than disconfirming of the validity of the theoretical explanations given by Bem and echoed by Spence and Helmreich.

To recapitulate, the theory underpinning both the PAQ and the BSRI emphasizes that sex-role orientation is not a trait, but a self-concept. That self-concept is defined relative to the societal conceptions of the differences between men and women rather than on a haphazard collection of actual differences between the sexes. The domain of the items on these scales is limited to personality descriptors rather than behaviors, attitudes, interests, or appearance. Masculine characteristics are gathered under the rubric of instrumental or agentic qualities; feminine characteristics are gathered under the rubric of expressive-communal qualities. In stark contrast to the empirical approach of using observed differences between males and females or homosexuals and heterosexuals as the criteria for M-F, Bem and Spence and Helmreich have with considerable specificity provided a conceptual grounding for the constructs of masculinity and femininity.

The restriction of the item content to personality characteristics may also contribute to the subtlety of the scale and thereby limit the stereotyping which is likely to occur in M-F measurement. In the case of the BSRI, this subtlety is enhanced by the inclusion of the Social Desirability items which are not only negative and

positive but are sex-neutral as well. It may well be that the problem of "fakability" is one that cannot be solved in a satisfactory fashion with the item selection process alone. An alternative suggested by Lunneborg (1970) with regard to M-F measurement is the use of a correction factor for sex-stereotyping in self-description.

The BSRI and the PAQ are genuine attempts to confront at least five of the six issues outlined as major criticisms of previous M-F measurement techniques, which were previously summarized. These two approaches use orthogonal dimensions to describe M and F, appeal to a theoretical understanding of the constructs, and select items on the basis of that theoretical understanding. The operations defining the construct strictly discriminate sex-role orientation from other related, but conceptually distinct constructs, although the precise boundaries of the constructs may prove to be a matter of theoretical dispute. The problem of subtlety is taken into account. The sixth issue, that of behavioral and experimental validation will be considered separately.

Current Behavioral Research

The Androgyny Model

To this point, the PAQ and BSRI have been described chiefly as direct responses to the problems implicit in the earlier forms of M-F measurement. While this is an accurate portrayal of the logic of the historical progression in M-F measurement, it falls short of explaining the renewed enthusiasm of researchers for this area. This enthusiasm was sparked more by the advancement of a distinctive new

model of sex-roles--the androgyny concept--than on any advances, however important, in measurement technique.

Bem (in press, 1974) posed an important question which has gone without resolution for decades: what does it mean to be more or less equally masculine and feminine? The implicit normative assumption behind the bipolar-unidimensional model of M-F seems to have been that a score falling towards the sex-appropriate endpoint for a subject was preferable in terms of global psychological adjustment. A score falling in the middle range seemed to signify sex-role confusion, indecision, or even sexual deviation (Berzins, 1975 ; Constantinople 1973). Bem hypothesized that masculine and feminine self-concepts were not incompatible and that a third group of individuals needed to be considered: those who had balanced amounts of masculinity and femininity.

At that juncture, it was not necessary, strictly speaking, to establish a complete conceptual separation of the qualities of masculinity and femininity in order to posit an androgynous middle range. Although Bem did separate the two scales empirically, the conceptual framework she employed, as reflected in the scoring system used on the BSRI, was not completely divorced from bipolar assumptions. In Bem's original formulation the scores on the two subscales of the BSRI were used to differentiate sex-typed from androgynous individuals on the basis of a t-statistic. This process essentially used two orthogonal scales to place individuals on a bipolar continuum from strongly sex-typed to androgynous to sex-reversed. Empirical considerations, criticism of the use of the t-statistic (Strahan 1975)

and the popularization of an approach which separated out low-low scorers from the androgynous category prompted Bem to abandon the three-group approach in favor of the four-group approach (Bem 1977). Since the original tripartite approach was used in several of the early behavioral studies it is necessary when re-examining these studies to give consideration to the accuracy of their interpretation and their implications for theory based on the quadripartite approach.

The four-group approach used by Spence et al. (1975) as well as other investigators (Heilbrun 1976; Kelly & Worell 1976; Kelly, Caudill, Hathorn, & O'Brien 1977) was adopted by Bem in light of the differences between high-high scorers and low-low scorers on an independent measure of self-esteem (Bem 1977). The approach demands that median scores be computed for both sexes combined on M and F, and individuals can then be classified as Masculine, Feminine, Androgynous, or Indeterminate depending on whether their two scores are above or below the appropriate median. While this may be preferable to the original approach used by Bem, it has been cautioned that it creates only broad typologies by placing individuals in one of four quadrants (Kelly & Worell 1977). While this scheme may be on firmer ground conceptually than Bem's original approach, another alternative is afforded by the use of multiple regression which examines linear relationships without regard to artificial groups.

Spence and Helmreich (1978) discuss this issue at some length and defend the use of typologies while noting that there are disadvantages attached to the use of such a categorical system. In

contrast to Bem (1977), they do not advocate the use of multiple regression. Spence and Helmreich argue that in their research multiple regression has not brought about "marked changes in the amount of variance accounted for in a number of subsequent analyses" (p. 36). They further express concern that the prediction of self-esteem was not improved by using M and F as predictor variables in a multiple regression equation "despite the fact that both were linearly and significantly related to each other and to self-esteem" (p. 36). It should be of some concern that the variables involved are positively correlated and multiple regression should be used precisely in such an instance since separate Pearson r 's might distort the relative importance of one of the predictor variables if the correlation of that variable to other predictor variables is not taken into account.

The newer dualistic measures of sex-role orientation were, at least in part, spurred by a new approach to the problem which will be referred to as the "Androgyny model." This model repudiated the implied norms of the bipolar model. It should not be assumed however that the Androgyny model is without its normative aspects, especially since these are quite the reverse of the normative implications of the bipolar-unidimensional model which preceded it. Bem hypothesized that androgyny implied greater flexibility for the individual. In vivid contrast to the reluctance of Spence and Helmreich (1978) to suggest that psychological masculinity and femininity will accurately predict behavior, Bem contends that there is a direct connection. This contention is echoed by Kelly and Worell who described Bem's "response repertoire model" in the following terms:

Simply stated, the highly sex-typed person is seen to have available a limited number of effective behavioral options to deal with situations. This restricted response repertoire is considered to be a function of the sex-typed person's reluctance or inability to engage in cross-typed behavior even in situations in which it might be adaptive (1977, p. 1102).

Bem's Original Research Program

The research program which Bem initiated to demonstrate the validity of this response repertoire model originally consisted of three experiments. Two additional experiments were designed to further investigate the failure of certain predictions to hold in one of the original three. The first experiment was designed to show that androgynous individuals would perform as well on a masculine sex-typed task as masculine individuals and at significantly higher levels than feminine individuals. The second was to show that androgynous individuals would perform a feminine task as well as feminine individuals and significantly better than masculine individuals. Results which partially disconfirmed predictions on the feminine task led to the design and completion of two additional studies designed to examine the second contention more fully. Finally, the third planned study was done to demonstrate that sex-typed individuals actively avoid sex-inappropriate tasks or situations even when they stand to gain from sex-inappropriate behavior. This research program will be reviewed in some detail.

In the first experiment (Bem 1975) a standard conformity-deception paradigm was employed to demonstrate the flexibility of androgynous individuals. Assuming that resistance to conformity in a group charged with rating a set of cartoons represented an "independent" and therefore masculine behavior, Bem predicted that for both sexes, masculine subjects and androgynous subjects taken together would perform significantly better, i.e., remain more independent, than would feminine subjects; further, the masculine and androgynous groups would not significantly differ on the dimension of independence. These predictions held. The mean number of trials on which feminine subjects conformed was significantly higher than the mean number of trials on which the combined masculine and androgynous subjects conformed. In addition, the mean number of conforming trials for the masculine and androgynous groups were not significantly different, although inspection of the means indicates that the actual means for androgynous individuals did fall in the direction of greater conformity. Since the definition of androgyny used in this case combined high-high and low-low scorers, this would be expected.

In retrospect, the data analysis of this study was overly simple. Instead of using the actual cartoon ratings ranging from 1 to 7, Bem would count a conforming trial as one in which an unfunny cartoon for example was given a rating higher than 4, the middle score. This point would seem to be trivial, however, since in the report of the experiment it was claimed that the use of actual scores yielded similar results. If as hypothesized by Bem, the actual M and F scores are somehow directly predictive of the repertoire of masculine and

feminine behaviors, a multiple regression on the dependent variable, resistance to conformity, should have shown a significant correlation between M and the masculine behavior called "independence" with femininity contributing little to the prediction. In a later paper (Bem 1977) reported re-analysis of this data in just such terms. The results were non-significant, i.e., no significant relationship was found between masculinity or femininity and independence for either males or females. Small partial correlations, positive for masculinity and negative for femininity seemed to have combined to create the significant differences between groups. It has not been noted in the literature that this effectively invalidates the claim that this study supports the Androgyny or "response repertoire" model.

In the second study planned to validate this model the dependent variable was observation of the subject playing with a small kitten. This behavior was intended to be a measure of expressive-nurturant or feminine behavior. It was anticipated that androgynous individuals would be as "nurturant" as feminine individuals toward the kitten and that the combined means of the two groups would be significantly higher than the mean of the masculine group. A set of variables based on coder's observations was constructed. On the basis of these variables, the predictions held only for male subjects. Androgynous and feminine males showed significantly more involvement with the kitten than masculine males, and the androgynous and feminine group means were not significantly different from one another. With females, however, androgynous subjects showed the highest overall involvement with the kitten, the masculine females a moderate amount, and the

feminine females showed the least involvement of the three groups. Thus the androgynous and feminine groups together were not significantly more involved or nurturant than the masculine females, and the androgynous and feminine groups did differ from one another significantly on two of the three dependent variables, as well as the summary measure, contrary to the original prediction.

These results were also reinterpreted using a fourfold model of sex-role orientation and a regression analysis (Bem 1977). The results for males demonstrated a significant relationship between femininity and responsiveness to the kitten ($\beta = .37, p < .05$). There was also a negative partial correlation between masculinity and responsiveness to the kitten but this was non-significant. The results for females however showed a significant positive partial correlation between femininity and kitten-playing and also a near-significant positive partial correlation between masculinity and kitten playing. These results might suggest that kitten playing was not the optimal behavior to use in the study of expressive-nurturant behavior. Alternatively, it might have suggested that the relationship of the F scale to "feminine behavior" was not great.

In order to examine these possibilities two more experimental situations designed to measure nurturance or expressive-communal behavior were designed by Bem, Martyna, and Watson (1976). The major factor in these studies distinguishing them from the kitten-playing study was the use of human beings as the objects of potential nurturance. Again, planned ratings by observers were used as the dependent measure and there were three sex-role groups. The first

study involved observation of interaction of subjects with a human baby. There was no significant effect for sex-role until, post hoc, subjects were redivided into four sex-role groups--masculine, feminine, androgynous, and indeterminate. The expected results were then observed with androgynous and feminine subjects showing significantly more interaction with the baby than masculine subjects, and androgynous and feminine subjects showing comparable levels of nurturance.

In the second of these two studies designed to take a closer look at feminine behavior, observed behavior and self-ratings were used as dependent measures in an experimental situation in which a confederate portrayed a lonely transfer student in order to elicit supportive or nurturant behavior from the subject. The predicted pattern of results was observed using the tripartite division of groups. Feminine subjects and androgynous subjects were not significantly different from one another. In addition, it was observed on several measures that females were significantly more nurturant than males.

On the basis of these two studies, one of which required post hoc reanalysis to evidence the predicted results, Bem et al. argue that behavioral support has been established for the hypothesis that androgynous individuals are more flexible in the behaviors they use than are sex-typed subjects. However, the redivision of Bem's three groups into four is accepted as legitimate, then the first study on masculinity and resistance to conformity ("independence") ceases to be supportive of the original hypothesis concerning the flexibility of androgynous individuals. If that redivision is not accepted,

then the infant-nurturance study is not supportive of the hypothesis. In either case, at most only two out of four studies can be accepted as legitimate support for the hypothesis. The direct support for the idea of enhanced flexibility based on these studies is therefore tentative and limited. Such flexibility, it must be said, would seem to be rather limited in its domain, and as Pleck (1975) has observed, the domain of sex-role related behaviors in comparison to all other behaviors is already a rather limited one.

The third of the studies originally planned to validate the androgyny model demonstrated the tendency of sex-typed individuals to avoid behaviors that were inconsistent with their particular sex-role orientation (Bem & Lenney 1976). In this study, individuals were offered small incentives (2-6¢) for performing a set of sex stereotyped activities (e.g., nailing two boards together, preparing a baby bottle) while being photographed. Higher incentives were always offered for cross-sex activities. Subjects were cautioned that familiarity at the task was not at issue. Of the thirty pairs of activities from which the subject selected, fifteen were choices between different combinations of feminine, masculine and neutral tasks. The other fifteen were both either masculine or feminine and paid different amounts. This allowed investigators to compare sex-stereotyped choices to a baseline and to later examine the discomfort of the subjects in performing cross-sex activities. The results showed that sex-typed subjects were significantly more stereotyped in their choice of activities despite the fact that it cost them money. This was especially true of males. Sex-typed subjects also reported greater

discomfort in performing cross-sex activities. These results are interpreted as support for the hypothesis that sex-role typing leads to a general constriction of behavior that androgyny does not.

This study surveys a wide variety of sex-linked behaviors and the results are complex but supportive of the idea that cross-sex activity is "motivationally problematic" (Bem & Lenney 1976, p. 48) for sex-typed individuals. It would seem that a stronger case can be built for this hypothesis than for the hypothesis that androgynous individuals have a greater repertoire of behaviors in situations. It is interesting that this final study used performance on a wider variety of sex-typed tasks rather than attempting a one-to-one correspondence between sex-role orientation and a given behavior. If Bem's program has provided support for the hypothesis that Androgynous individuals are more flexible than sex-typed individuals it is chiefly through this final study showing that sex-typed individuals inhibit performance of cross-sex behaviors, even when rewards are available and penalties would seem to have been minimized.

Sex-role Orientation and Social Skills

Like Bem, Kelly and his co-workers have chosen to emphasize the behavioral aspects of masculinity and femininity (Kelly & Worell 1977; Kelly, O'Brien, Hosford, & Kinsinger 1976). In particular, observing that the adjectives which compose the BSRI have largely interpersonal referents, they have examined the realm of interpersonal behavior. "It is our contention that sex-role style on a more basic level reflects certain sex-correlated social skills" (Kelly et al.,

p. 11). In the report of an experiment designed to demonstrate the linkage between sex-role orientation and relative performance advantages in assertive v. warm-expressive situations, data generally supportive of that hypothesized linkage was found.

However, the findings were most valuable in providing insight into the limitations of the predictive capabilities of sex-role variables. Using the BSRI and a median split procedure to separate four groups, these investigators asked subjects to role-play a series of scenes initiated by a description of the scene and a one-line prompt given by a female experimenter who served as a partner for the scene. Two types of scenes were employed; the first type was positive and designed to elicit warm and commendatory behavior. The second type was negative and designed to elicit "refusal-assertiveness." Sessions were audiotaped and reliable ratings describing latency, duration, loudness, affect, assertiveness and speech dysfluencies were used as dependent variables. The results reflected mixed patterns rather than clear-cut group differences across all categories of dependent variables.

A "striking" finding, in the view of the authors, was that the observation of sex-role group differences depended on the sex of the individual subject in combination with the type of scene. For females, the greatest group differences occurred in the negative scenes requiring assertiveness. However, for males, more sex-role category differences occurred in those scenes requiring warm, expressive behavior. Where category differences did occur there was general

support for the contention that androgynous individuals are flexible and interpersonally effective relative to sex-typed individuals, that masculine individuals are most effective where forceful assertive behaviors were required, that feminine individuals are most effective where warmth was required, and that indeterminates are least interpersonally effective. The authors attempt to explain the differences between the patterns as they related to sex of participant, but for the purposes of this review it is sufficient to note that what support this study provides for the behavioral validity for the BSRI is somewhat mitigated by the emergence of sex differences in the data.

BSRI Scores and Naive Judgments of Sex Role

In contradistinction to the studies reviewed above, which attempt to externally validate the BSRI and to verify the predictions of Bem's Androgyny model, at least one study has been done which was less concerned with the prediction of specific behaviors than with the interpretations by others of the expressive behaviors of individuals as masculine or feminine. In a pair of experiments, Lippa (1978) made a study of the relationship between sex-role variables by the BSRI and the perceptions of the subject's sex-role as rated by naive judges. Judges were asked to make ratings of short videotaped segments of 18 individuals giving a similar demonstration. There were three androgynous, three masculine and three feminine subjects of each sex represented in the 18 videotaped segments. These subjects were selected on the basis of the BSRI. Mean ratings for 48 judges were computed on four criterion variables: "Masculine," "Feminine,"

"Dominant-assertive," and "Compassionate-sensitive to others." The results demonstrated that individuals make judgments of masculinity-femininity on the basis of expressive cues rather than trait attributions, such as dominance or compassion. When the mean ratings by judges for "Femininity" were subtracted from the mean ratings for "Masculinity" to create a bipolar masculinity-femininity index, those ratings discriminated the three sex-role groups based on the independent BSRI measures of the stimulus individuals. In fact, these masculinity-femininity ratings by subjects correlated .90 and .79 respectively for males and females with the so-called Androgyny scores from the BSRI, another subtractive measure, accounting for 62-81% of the common variance.

While it is encouraging that masculinity-femininity based on the BSRI correlates with naive ratings of M-F, it is conceptually troubling that this is done on the basis of a bipolar conception on the part of judges, for whom M and F are highly negatively correlated. By way of contrast their ratings for "Dominant-assertive" and "Compassionate-sensitive to others," phrases which represent trait descriptions representing the M and F scales of the BSRI, did not reliably discriminate between the actual sex-role groups of the stimulus individuals. It is intriguing that the stimulus individuals were assigned to sex-role categories on the basis of self-attributions but the ratings of the judges on the sex-role dimension were not mediated by trait attributions--who in fact, could not discriminate on the basis of those attributions.

In the second experiment, components of the videotape presentation were separated to analyze the relative reliance by judges on looks, voice, bodily expression, facial expression, and so forth. Different sets of judges made similar ratings of the same stimulus persons based on hearing only the voice, seeing the video picture only, seeing only the head, seeing only the body, and seeing only a still photo made from the television picture. The results seemed to indicate that judges used different cues for judging males and females. One of Lippa's findings was that body cues were more highly correlated with overall estimation of sex-role for females than for males. This result may reflect a similar phenomenon to that described by Ickes and Barnes (1977) who examined self-monitoring as a moderating variable in the expression of a given trait. Within same-sex dyads, high self-monitoring males gestured less than low self-monitoring males. However, females showed the reverse pattern; high self-monitoring females gestured more than low self-monitoring females.

Ickes and Barnes (1978) extended this line of research in a way which provides provocative insight into the behavioral differences between sex-typed and androgynous individuals and support for the Androgyny model advocated by Bem. Subjects were observed and videotaped for five minutes in unstructured mixed sex dyads where the sex-role orientation of the partners represented the four possible combinations of either sex-typed or androgynous subjects (Androgynous male, androgynous female; Sex-typed male, sex-typed female; Sex-typed male, androgynous female; Androgynous male, sex-typed female). They hypothesized that an initial, unstructured interaction between a

masculine sex-typed male and a feminine sex-typed female would be stress inducing because of incompatibility between the instrumental and expressive orientations of the respective participants. The flexibility inherent in androgyny was supposed to permit a more compatible interaction. No clear prediction was made with regard to the androgynous/androgynous pairs.

The methodology was simple. On a pretext, experimental subjects were asked to wait for five minutes in a room which contained a concealed videotaped camera. Most of the dependent measures consisted of ratings made by judges from the videotapes. Post-interchange measures were also collected, including a measure of liking. Two of the four sets of pre-selected dependent measures showed clear evidence of a poorer quality of interaction within the masculine male-feminine dyads compared to dyads which included at least one androgynous partner. The authors examined a number of alternative explanations for the effects but none were as satisfactory as the Androgyny theory. The generalizability of these findings is limited by a number of factors: for example, the fact that many or most interactions are not unstructured but occur within a social context. Nevertheless, this would seem to be a clear instance of greater flexibility on the part of androgynous individuals in a situation which approaches naturalistic.

Summary of the Behavioral Studies

Taken case by case, the behavioral evidence for the external validity of the BSRI and the Androgyny model itself is not clear-cut. However, taken as a whole the nine studies which are reviewed above

provide a fairly sound foundation for the theory. Individuals who score high on both scales are more flexible in performing cross-sex behaviors (Bem 1975; Bem, Martyna & Watson 1976; Bem and Lenney 1976; Ickes & Barnes 1978; Kelly et al. 1976). Sex-typed individuals seem to have generally shown in these studies specific deficits in the predicted direction. One difficulty in these studies has been the specification of masculine and feminine behavior. If masculinity is defined as independence and subsequently tested against "independent" behavior, or if femininity is defined as nurturance and then tested against a "nurturant" behavior, little is added to our knowledge of masculinity and femininity that is not tautological. Crucial in this regard was the study by Lippa (1978) which demonstrated that judges could discriminate masculinity-femininity rather accurately when compared to the independent measure of the BSRI. This result was however perplexing because the implicit model of the raters was at variance with the model employed by the BSRI--a bipolar as opposed to a dualistic model.

What is lacking in the behavioral research is a clear tie between levels of masculinity and the performance of "masculine" behaviors as well as a similar tie for femininity and "feminine" behaviors. It is precisely this lack which makes it difficult to assess the impact of M and F over and above variables with which they are correlated in the prediction of actual behavior.

Self-Esteem and Sex-Role Orientation

In marked contrast to the complexity of the results of attempts to demonstrate the relation of sex-role orientation to sex-role

behaviors, the examination of personality correlates to sex-role orientation has produced at least one very salient result. This result involves the correlation of sex-role orientation, especially masculinity, to self-esteem. It has been pointed out earlier in this discussion that masculinity and femininity have been conceptualized as two orthogonal dimensions of positive self-attributes which differ in relative value for the sexes. This being so, three different kinds of relationships between these dimensions and self-esteem might be intuitively anticipated: (1) Masculinity would promote self-esteem in males; femininity in females, (2) both masculinity and femininity would promote self-esteem in either sex, and (3) either masculinity or femininity would be related to self-esteem in both sexes, while the other attribute would be independent of self-esteem.

The first of these possibilities would seem to express the traditional view of sex-roles: high self-esteem would be primarily related to high scores on the sex-appropriate scale with some possible secondary contribution deriving from the second, sex-reversed scale. Spence and Helmreich note that "One of the cores of women's and men's self-concept is the degree to which they believe they measure up, or believe it important to measure up, to their abstract conception of what it is to be a proper woman or man" (1978, p. 116). The discrepancy between the ideal self and the perceived self has often been used as a measure of self-esteem (Wells & Marwell 1976). It might be expected that the extent to which one's self-concept matches a sex-appropriate ideal might contribute to overall feeling of self-worth. Therefore, a male would tend to have higher

self-esteem if he were more masculine and a female would have increased self-esteem if she were more feminine.

A second possibility would be more congruent with the androgyny hypothesis advanced by Bem (in press). Since both masculinity and femininity represent positive realms of self-attributions both might contribute to overall self-esteem either in an additive or interactive fashion. Spence and her colleagues expressed this in the following manner:

The view that both masculine and feminine characteristics have positive implications for effective functioning suggests that contrary to conventional opinion, the relationship between men's self-esteem and their scores on the female-valued scale of the PAQ might also be positive. The same reasoning leads to the expectation that similar results might also be obtained in women, with higher scores on all of the subscales and not merely the female-valued subscale being related to higher self-esteem (Spence et al. 1975, p. 32).

In a similar vein, Kelly and Worell emphasized the hypothesized behavioral flexibility implied by androgyny:

Recent theorizing about androgyny would propose that the androgynous individual has at his or her disposal a larger or more diverse number of socially approved behaviors that are available for problem solution or for obtaining reinforcement. Persons who describe themselves as capable of more situationally appropriate behaviors using masculine-typed behavior for assertive purposes and feminine-typed behavior for expressive purposes should achieve a higher self-esteem score (1977, p. 1108).

Thus one might expect that rather than seeing self-esteem as a direct function of sex-appropriate sex-role orientation it might be a combined effect of both masculinity and femininity, acting in a more or less symmetrical fashion.

The third possibility is that self-esteem might be directly related to one but not both of the dimensions which describe sex-role

orientation, for either sex. This outcome is less consistent with the theoretical approaches underlying either the unidimensional or the dualistic measurement techniques. It would therefore require a re-examination of the theories in use and a close examination of the data. Because such a finding could not be explained by these theories, some theoretical understanding of it would have to be provided.

The first published study examining self-esteem in this context seemed to support the point of view that both masculinity and femininity were important factors in self-esteem for both sexes (Spence, et al. 1975). The measures used in this study were the PAQ as the measure of sex-role orientation and the Texas Social Behavior Inventory (Helmreich, Stapp, & Ervin 1974) as the measure of self-esteem.

The Texas Social Behavior Inventory (TSBI) was "designed to determine individuals' self-confidence and competence in interpersonal situations and is generally described as a measure of social self-esteem" (Spence, et al. 1975, p. 31). This instrument does not significantly discriminate between the sexes, is not correlated to intelligence, is only mildly related to social desirability, and then only for females ($r = .32$, $p < .01$) (Helmreich, et al. 1974). The support offered by its authors for its validity, other than face validity, was a correlation with the self-esteem scale of the California Personality Inventory ($r = .50$ for males, $p < .001$; $r = .52$ for females, $p < .001$) (Helmreich, et al. 1974). Fay and Brown (1979) provided further convergent validity in the form of correlations with four other self-esteem measures (including again the CPI

self-acceptance subscale) using female respondents only. These correlations ranged from .35 to .60 and were all significant. Although these correlations provide some support for the view that the TSBI measures self-esteem, it is perhaps germane to note the criticism of Wylie (cited in Wells & Marwell 1976, p. 79) that measures of self-esteem are often devised only for a particular study or set of studies without much effort to assess the adequacy of the measurement--therefore they are of unknown quality and tend to be short-lived. Interpretations based on the TSBI need not be discounted but some caution must be used in view of its as yet limited history as a self-esteem measure.

Spence et al. (1975) reported very high significant correlations between the TSBI and the M (male-valued) scale of the PAQ in a sample of college students (males, $r = .77$; females, $r = .83$) as well as mild correlations between self-esteem and the female valued scale of the PAQ (males, $r = .42$; females, $r = .30$). The impact of these correlations seemed to be additive because a significant main effect for sex-role group was found on self-esteem with the four groups significantly differing from one another in the following order: Androgynous, highest; Masculine; Feminine; Indeterminate, lowest. These results were essentially replicated in separate samples of college and high school students (Spence & Helmreich 1978, p. 55).

A problem exists in the interpretation of these data. Succinctly, the problem is the fact that the M and F scales of the PAQ are positively correlated and not orthogonal. For example, for males, the correlations reported for four samples between the M and F scales are all positive and range from .11 to .47; in females, they are also

all positive and range from .09 to .14 (Spence et al. 1975; Spence & Helmreich 1978). While, in practical terms, the scales may be said to be orthogonal for females, it would be disingenuous to claim that a correlation of .47 for college males represents orthogonality.

In fact, Spence and her colleagues go so far as to admit that the F and M-F scales of the PAQ do not significantly contribute to the prediction of self-esteem over and above the prediction made by the M score alone--precisely the reason why they chose to present their data in terms of the analysis of variance across four groups rather than using a multiple regression analysis. One might speculate that the variance which is shared between M and F contributes to the significantly higher scores shown by the Androgynous group but that the actual contribution of femininity as a distinct psychological construct is minimal. Stated another way, it might be expected that if that portion of the variance which is shared between M and F scores were partialled out, F would be a very ineffective linear predictor of self-esteem. Thus, although the authors would prefer to interpret their data as demonstrating that both M and F are important in relation to self-esteem, it is clear that the major impact is due to the M scale.

The careful observer will also note that although the scores of the male valued scale of the PAQ are significantly different for males and females, where the scores on the TSBI do not differ for the sexes, the size of the correlations between the TSBI and the M scale are generally higher than the correlations between TSBI scores and other measures of self-esteem. The range of correlations between the TSBI and other measures of self-esteem, it may be recalled,

ranged from .35 to .60. The correlations which have been reported by Spence and her colleagues between the PAQ M subscale and the TSBI range from .54 to .77 in males and from .59 to .83 in females. Where convergent validity and scale reliability are reported in terms of Pearson's r statistic, discrepancies of this magnitude raise some rather interesting questions about the titles of two highly correlated scales. The correlations between the TSBI and the PAQ M scale are generally of comparable magnitude to those between the PAQ M scale and the BSRI Masculinity subscale ($r = .75$ for males; $.73$ for females) (Spence & Helmreich 1978, p. 24).

Having noted these criticisms of the work employing the PAQ, we may note that other researchers have replicated the positive correlation found between masculinity and self-esteem. Bem (1977) found that in males self-esteem, as measured by the TSBI, was significantly related to the M scale of the BSRI ($\beta = .48$) but not to femininity; and that for females self-esteem was related to both M ($\beta = .54$) and F ($\beta = .28$). These partial correlations indicate that masculinity is more important in the prediction of self-esteem than femininity is, even for females. Wetter (Note 2) used the PRF ANDRO scales to measure sex-role orientation and developed an ad hoc self-esteem questionnaire, the SEQ. These measures were administered to large samples of high school and college students. There were sex differences on the self-esteem questionnaire with males having higher scores. The patterning of the results indicated a strong main effect for masculinity in the prediction of self-esteem scores. The authors point out the importance of the finding that feminine females, who comprise the largest single group of females in sex-role terms,

are significantly lower in self-esteem than masculine or androgynous females. Some caution must be exercised in interpretation of these results since the measures used are somewhat different from those used in other studies.

Fay & Brown (1979) described the results of a study, using only female subjects, which relied on a variety of measures for sex-role orientation and for self-esteem. Their conclusion was again that self-esteem in females is strongly related to their self-reported psychological masculinity. Though they used two other measures of sex-role and five measures of self-esteem, of most interest in this context are the correlations of the BSRI to the TSBI which were .08 and .77 for F and M respectively. The latter is the most extremely high correlation between M and self-esteem cited but the general pattern observed was one of virtual orthogonality for femininity and self-esteem in contrast to moderate to strong positive correlations between masculinity and self-esteem.

Deutsch and Gilbert (1976) used adjustment as a dependent variable rather than self-esteem but their findings merit mention here since they are congruent with the patterns of results reported by those examining self-esteem per se. These investigators used the Bem scales and two scales of intrapersonal and interpersonal adjustment. Unfortunately, they used a balance measure of androgyny rather than the fourfold approach, making it more difficult to draw direct inferences about the effects of masculinity or femininity separately. Nevertheless, their evidence suggested that for both males and females the greatest adjustment occurred with the following ranking (highest to lowest): Masculine, Androgynous, and Feminine. Since

the androgynous group would include those with relatively lower, but balanced, scores on both scales of the BSRI, this type of ranking would be anticipated where masculinity was strongly related to adjustment and femininity was not.

Thus far, the results of all of the correlational literature reported have strongly confirmed the role of masculinity. The contribution of femininity in relation to enhancement of self-esteem has varied but is relatively less important. There is, however, a study by Doherty and Schmidt (1978) which employed the Tennessee Self-Concept Scale, a more widely established measure of self-esteem than the TSBI, and found that in overall self-esteem feminine females showed higher scores than masculine females, even though the androgynous group was still the highest. The sample consisted of 140 female respondents from a random pool of 200 freshman women to whom the BSRI and Tennessee scales were mailed. Consequently, cross-sex comparisons are not available. On overall self-esteem the Androgynous group was significantly higher than the Undifferentiated and Masculine groups, and was higher although not significantly higher than the Feminine group. This study would tend to support the point of view that femininity as a self-concept does enhance the self-esteem of women, contrary to the evidence that even for females the major correlate of self-esteem is masculinity. In the face of the general trend in the studies presented above, the use of a mailed questionnaire, even with a high return rate, and the omission of the presentation of data for male subjects substantially limits the impact of this study in the context of the present discussion. This is

especially unfortunate in light of the use of the Tennessee Self-Concept Scale which has had a broader usage in the published literature as a measure of self-esteem and therefore might provide more basis for comparison across empirical studies than the TSBI or other measures.

At the outset of this section three hypothetical relationships between self-esteem and sex-role orientation were suggested as intuitive possibilities. At this point, it appears that the third possibility, i.e., that self-esteem would be chiefly related to only one of the two sex-role dimensions, has been confirmed by the evidence available. It seems safe to say that for self-esteem the principal source of shared variance is with the Masculinity scales, with the potency of the Femininity contribution to higher self-esteem in serious question. The influence of Femininity was maximized (1) where the PAQ (which has positively covarying M and F scales) was used, and (2) in the instance where the Tennessee Self-Concept Scale was employed as the measure of self-esteem. Overall, however, an incontrovertible pattern has emerged from the available data, emphasizing the communality between masculine self-attributions and measured self-esteem.

The central problem posed by this result is that of theoretical explanation. Many of the writers in this area have chosen to emphasize the relative superiority of the androgynous subjects to other sex-typed or indeterminate subjects, even in cases where masculine and androgynous subjects were not significantly different (Spence, et al. 1975; Spence & Helmreich 1978; Bem 1977; Wetter, 1975). As it has become increasingly clear that it is high masculinity which

is primarily accounting for high self-esteem in both males and females it has become necessary for some theoretical explanation to be advanced. Fay and Brown take the following approach in summing up their findings across a variety of sex-role and self-esteem measures:

Perhaps measures of masculinity, including Bem's conceptualization, have operationalized masculinity as competence. A simple examination of various masculinity test items leads one to a view of a masculine person as self-confident and competent. This competence may well be the 'stuff' of which self-esteem is made, particularly when contrasted to the emotional, sensitive, retiring quality of the femininity conceptualizations (Note 1, p. 2).

Kelly and Worell (1977) described the Masculine and Androgynous groups on the one hand, and the Feminine and Indeterminate groups on the other as clustering separately on self-esteem, life history data, and certain behavioral tasks. They suggested the possibility that feminine and masculine behaviors are differentially effective socially, masculine behaviors being more likely to lead to positive outcomes. "In this respect, further efforts to examine the relative effective contributions of masculinity and femininity to androgynous role orientations would appear useful" (p. 1113).

The problem of a theoretically adequate account for the relationship between higher M scores and higher self-esteem is far from being resolved. An alternative that has not been given consideration is the possibility that the measurement of masculinity is somehow confounded with the measurement of self-esteem. Fay and Brown (1979) suggest, for instance, that for women, Bem's Masculinity scale may in fact be more an independent measure of self-esteem than a measure of sex-role preference. The type of item which is used to measure M-F has been demonstrated to exhibit widely varying effects on the

empirical definition of sex-typing for any particular individual, as noted by Constantinople (1973). One of the effects of using self-attribution of personality descriptors as the measure of sex-role orientation may be that self-esteem variance is subtly incorporated into that measurement. The reason why that variance would be concentrated in the Masculinity score and not in the Femininity score is open to conjecture.

The fact that masculine tasks and typically masculine characteristics are more valued in society (Broverman, Broverman, Clarkson, Rosenkrantz, & Vogel 1970; Broverman, Vogel, Broverman, Clarkson, & Rosenkrantz 1972) reflects the pervasive "male supremacy" of contemporary society which may account for that relationship. A pertinent distinction can be made between the influence such societal attitudes have on the measurement of a hypothetical construct and the influence such attitudes have on the construct itself. Kelly and Worell's argument emphasizes that masculine attributes actually lead to enhanced social or behavioral outcomes. Another interpretation might be that the personality characteristics of the person with high self-esteem are felt to be more appropriate for men in our society because of the preferential status bestowed upon them. The fact that the same language--the language used in the masculinity scales of Bem and Spence et al.--is descriptive both of men and people with high self-esteem is not coincidental. But to use the same language to describe two phenomena does not mean that the psychological constructs themselves are identical. To gain a truer picture of the effects of masculinity on personality and behavior it may be necessary to

analyze that impact holding self-esteem constant or using self-esteem as a blocking variable. Wells and Marwell (1976) have stated that self-esteem has often been used in precisely this way, as a blocking variable to separate out the specific effects of self-esteem from the effects of other variables of interest.

Statement of the Problem

In establishing the importance of a psychological construct it is essential that the construct tell us more about the world than we already know. If a construct is highly correlated with a more frequently investigated concept, it must be shown that the newer construct provides unique information about the world that the earlier concept was incapable of producing. The survey of the sex-role orientation literature has produced two essential demands for further research. The first need is for further behavioral support for the dimensions of masculinity and femininity. The second is for demonstration that the sex-role questionnaires can produce predictions of experimental results above and beyond what could have been predicted given the known relationship between sex-role orientation and self-esteem measures. That is, if a given result which is attributed to sex-role orientation could have been predicted just as well by the use of a self-esteem measure alone, the use of a sex-role orientation measure is not justified. The observation by Kelly and Worell (1977) that the two high self-esteem groups (Masculine and Androgynous scorers) seem to cluster over and against the low self-esteem groups (Feminine and Indeterminate scorers) indicates an overview of the literature which almost insists that the relationship of self-esteem

to variables of interest be separated from the relationship of sex-role orientation to those variables.

The project of combining these two needs was the object of the present study. We have seen that the original study designed to demonstrate the relationship between masculinity and masculine behavior (Bem 1975) produced somewhat disappointing results when the data were re-analyzed. Multiple regression using "Independence" (resistance to conformity) as the dependent variable and Masculinity and Femininity scores as the two predictor variables failed to produce a significant result even though planned t-tests had shown that the Androgynous and Masculine groups did not differ from one another and were significantly different, taken together, from the Feminine groups in both sexes. In addition, the close identification of self-esteem and Masculinity is a clear result of the correlational literature. McGuire (1969, p. 251) has made two suggestions with regard to persuasibility and conformity in a review of the literature on attitude change which are pertinent here. First, although the evidence is conflicting, there is a good possibility that a negative relationship exists between chronic self-esteem and influenceability. Second, females are generally more easily influenced than males. Although sex differences and sex-role patterns are not identical, sex differences are suggestive of hypotheses with regard to masculinity and femininity.

The major objective of the present study was the examination of the relative importance of sex-role orientation and self-esteem in the prediction of influenceability or its reverse, independence.

Although the evidence was contradictory, a working hypothesis was formulated: that in fact, sex-role orientation would prove unimportant in predicting "Independence," i.e., that self-esteem would prove to be the only significant predictor of independence with masculinity, Femininity, and the interaction of M and F failing to add significantly to the multiple R.

A subsidiary task of the study was to analyze the relationship of Femininity scores to self-esteem by performing a more sophisticated analysis of the data than had been reported in the available literature. This analysis incorporated an interaction term, as well as M and F scores, in the prediction equation for which self-esteem was the dependent measure.

The division between those sex-role orientation researchers who advocated a "strong hypothesis" that sex-role orientation was not only a psychological entity but also has direct behavioral implications (e.g., Bem or Kelly) and those who advocated a "weak" hypothesis that sex-role orientation has cross-situational validity but cannot be expected to demonstrate clear and strong correlations with particular behaviors (e.g., Spence & Helmreich) was not directly at issue in this study. However, demonstration of significant predictive ability of M and F, over and above any effects observed for self-esteem, would provide support for the "strong" hypothesis.

CHAPTER III

METHODS

Subjects

Overall. Subjects were 165 male and 212 female undergraduate students enrolled in introductory or Developmental Psychology classes at the University of North Dakota during the 1978-1979 academic year. All of these students completed the Bem-Sex-Role Inventory and the Texas Social Behavior Inventory. Of the males, 84 completed these questionnaires in the Fall semester, 81 in the Spring semester. Of the females, 131 completed the instruments during the Fall, 82 during the Spring. Only those subjects who completed the questionnaires during the Spring semester were eligible for inclusion in the conformity study itself which was run during the Spring.

Conformity Induction Participants. The procedures by which the subjects for the conformity study were selected differed somewhat from those used by Bem in her original (1975) experiment and will therefore be described fully. The approach now advocated by Bem (1977) and others uses a combined-sex median-split procedure to divide subjects into four groups (above both medians, or androgynous; below both, or indeterminate; above the M median only, or masculine; above the F median only, or feminine). Since a normative sample of at least 165 individuals of each sex was available, groups were divided by mean rather than median. The practical impact of doing

so was, as expected, negligible. All scores were converted to T-scores ($M = 50$, $SD = 10$) initially; those subjects having either an M or F score falling within the 48-52 range (expressed as a T) were then dropped in an attempt to exclude those individuals whose assignment to sex-role orientation group was borderline. On this basis 19 males and 17 females were excluded from the total Spring sample. The remaining 62 males and 65 females from the Spring semester pool were then divided into sex-role categories and within each category they were rank-ordered according to their scores on the self-esteem measure, the Texas Social Behavior Inventory.

Subjects were then invited to participate in order of the degree to which their self-esteem scores varied from the mean of their sex-role group; i.e., for each sex, the high and low extreme scorers on self-esteem were invited first from each sex-role group followed by the next highest and the next lowest, respectively, etc. In the actual conformity study 26 females and 32 males participated.

Measures of Sex-Role Orientation and Self-Esteem

Sex-role orientation. The Bem Sex-Role Inventory (BSRI; Bem 1974) is a paper-and-pencil measure of sex-role orientation containing 20 masculine, 20 feminine, and 20 sex-neutral social desirability items. The subject is asked to rate him- or herself on a seven-point Likert scale for each item. The scale ranges from "Never or almost never true" to "Always or almost always true." Examples of items are "Self-reliant," (m); "Yielding," (F); and "Helpful," (S-D).

The BSRI therefore yields three separate scores, Masculinity, Femininity, and Social Desirability.

Self-esteem. The Texas Social Behavior Inventory (TSBI; Helmreich, Stapp & Ervin 1974) is a 32-item measure of self-esteem and social competence. The version used for this study was a short form, TSBIA, a 16-item version which correlated above $r = .97$ for both males and females according to Helmreich and Stapp (1974). This measure was chosen because of its previous use in the literature pertaining to sex-role orientation (Spence, et al. 1975; Bem 1977; Spence & Helmreich 1978; Fay & Brown 1979).

Stimuli

As in Bem's (1975) original study, cartoons were used as stimuli in the conformity induction. A set of 200 cartoons from a variety of periodicals were assembled. A group of judges, 12 males and 12 females, all undergraduates enrolled in Introductory Psychology classes, rated each of these cartoons on a scale from 1 ("Very unfunny") to 9 ("Very funny"). Those cartoons (34%) with the highest variance were eliminated. Of the 132 remaining cartoons, the 46 having the highest mean ratings across judges and the 46 having the lowest mean ratings were selected for use as stimuli. The overall mean rating for the "unfunny" cartoons was 3.86 and the range of the mean ratings for these cartoons was from 2.25 to 4.458. The overall mean rating for the "funny" cartoons was 5.77 and the range of the mean ratings extended from 5.125 to 7.25.

Procedure

Conformity induction. Experimental participants representing both sexes, four levels of sex-role orientation, and high and low self-esteem relative to their sex-role group, were then invited to participate in an experiment "on humor." On arriving for the experiment each subject was placed in a small room with a table and the following items: A standard consent form, a set of written instructions, a set of cartoons in a notebook, a microphone, a pair of earphones, and an answer sheet. The subject was asked to read and sign the consent form and the instructions before the experiment began. The instruction sheet read as follows:

The study in which you are participating is a duplication of an experiment done at Stanford University. We will be trying to follow their procedures as closely as possible. For this experiment, you are requested to rate a set of 92 cartoons on the following rating scale:

/-----/-----/-----/-----/-----/-----/
 1 2 3 4 5 6 7

Very
Unfunny

Very
Funny

Using this rating scale, we want you to give your rating out loud into the microphone and, as a double check for accuracy, we want you to write your rating down. You will be able to hear the experimenter through your earphones. Four people will be rating the same set of cartoons at the same time. Each of you is assigned a number. You are rater number 3.

For purposes of experimental control you are all sitting in different rooms. You will be instructed when to turn to each cartoon and then the experimenter will ask for your rating for that cartoon. For each cartoon the experimenter will call the different raters in a different order which has been randomly determined. Therefore it is essential that you listen carefully for your number and answer promptly. This experiment should take about an hour to finish.

After the instructions had been read and the consent form signed, the actual conformity induction began. The earphones were actually connected to a tape-player and the microphones were inoperative. Through the earphones the subject was able to hear the tape-recorded voice of the experimenter and three other "raters" who were actually confederates of the same sex as the subjects. Although the subjects were led to believe that the three other voices belonged to their co-participants, in actuality subjects were run in groups of only three at a time due to space limitations. This procedure differed from the original study in which subjects were run four at a time. Nevertheless, the illusion that four students were participating at a time was maintained, and none of the students whose data were used disclosed any doubt during debriefing that there had been fewer than four students participating. This was true even in one instance where a female subject failed to appear for the experiment and the procedure was run with only two real subjects.

Conformity was tested by having false evaluations given on 36 critical trials. On these trials, all of the recorded voices would give high ratings (6 or 7) to those cartoons which had been pre-rated as unfunny, or low ratings (1 or 2) to the cartoons which were pre-rated as funny. On critical trials the subject, as "rater number 3" was always called on last. In addition, on ten other trials the confederates gave uniformly false evaluations of the same magnitude even though the real subject was not called on last. On the remaining 46 trials, the non-critical trials, the subject could be called in any randomly determined sequence relative to the other raters.

Following the conformity induction, subjects were completely debriefed, informed about the purposes and methods of the study, and sworn to secrecy.

Taped Responses: It was necessary to run subjects separately by sex and to create two tapes of confederates, one for males and one for females. Females were run earlier in the semester than males. Two of the 26 females indicated suspicion of the deception in the experiment and their data were not used in the analyses.

Each tape had the voice of the experimenter and the voices of three confederates. The confederates were graduate and undergraduate students assigned the task of giving predetermined false ratings on the randomly determined critical and extra false feedback trials as well as giving their own personal ratings on all 46 non-critical trials. A comment made by the experimenter to the male confederates led to important differences between the original male and female tapes. Male confederates tried to use the mid-range ratings (3, 4, 5) more often on non-critical trials due to the offhand observation by the experimenter than because of the scripting those middle values might seem underrepresented on the tape. The impact of this was not foreseen but became obvious as the experimental subjects' answers were scored. The almost exclusive use of mid-range values on non-critical trials meant that a uniformity of response among the three taped voices was observable on almost every trial of the 92-trial series on that first male tape. As a result, the male subjects were conforming even more on the critical trials.

In order to remedy this situation, a new tape was made by three more confederates who were given only the same instructions as the

female confederates. In addition, correlations were computed between the sums of the raters' responses on the 46 non-critical trials in order to analyze stimulus differences. The correlations were between the two male tapes, the original mean ratings, and the female tape.

Unfortunately, the making of a new stimulus tape for males meant more delay and a further attrition of the subject pool as a result of that delay. It was also impossible to fill all cells of subjects since the number of indeterminate and feminine males in the pool was already limited and some of those subjects had already been through the induction and debriefing process. In the end, there were 15 male subjects run with the first tape and 14 with the corrected tape. Of this latter group, two did not meet the borderline exclusion criterion, that is, they had one sex-role score that fell near the mean. The data from three additional male subjects were not used because they indicated suspicion of the experimental manipulation during debriefing.

CHAPTER IV

RESULTS

Treatment of the Data

The two questionnaires, the Bem Sex-Role Inventory and the Texas Social Behavior Inventory, produced four scores: Masculinity (M), Femininity (F), Social Desirability (Soc), and Self-esteem (SE). The BSRI uses a seven-point Likert scale for each item, and each of its three scales contains twenty items. The TSBI (Short Form A) uses a five-point Likert scale for each item and contains 16 items.

In using multiple regression to analyze the relative predictive power of orthogonal sex-role variables it is crucial to isolate any variance that might be uniquely attributable to the interaction of M and F. It is possible to isolate the unique variance attributable to the interaction of two variables in regression analyses. This is done by using the product of the two as a separate predictor variable in combination with the two predictor variables of primary interest within a single regression equation. It can be argued that by introducing this extra variable into an equation the use of the multiple regression technique demonstrates far greater precision in the specification of effects than an analysis of variance. This process, it should be noted, is most accurate when scores of the two constituent variables have been standardized.

Consequently, in each of the analyses made where sex-role variables were to be included as predictor or independent variables, the raw scores on the M and F scales were first converted to T-scores ($\underline{M} = 50$, $\underline{SD} = 10$). This procedure not only prepares scores for examination of possible interactions but also circumvents any problems of comparability of raw scores between the M and F scales. For simple consistency, the other questionnaire data, Social Desirability scores, and Self-esteem scores, were also converted to T-scores.

Independence scores were computed in the following manner. Subjects gave ratings ranging from 1 (Very unfunny) to 7 (Very funny) for each cartoon. On one-half of the critical trials a funny cartoon was uniformly evaluated by the confederates as unfunny (1 or 2). On these trials, the higher the rating given by the subject, the more independent he or she was from the other "subjects." However, on the other half of the critical trials--where confederates gave unfunny cartoons very high ratings--it was necessary to reverse the subject's ratings so that a higher rating would signify greater independence. This was done, and all the ratings given by the subject were then summed to create an Independence scores. Higher independence scores reflected a general pattern of resistance to conformity. Low Independence scores reflected a general tendency to conform. The potential range of the Independence scores was from 36 to 252. Actual scores ranged from 81 to 185. Since these scores were to be used only as dependent variables, and would not therefore be subjected to multiplication for purposes of studying interactions, these data were left in raw form.

Finally, it was necessary in the analyses of the data obtained from male subjects (who heard two different stimulus tapes) to add an extra variable into the prediction equations to account for that part of the predicted variance that was uniquely attributable to differences in the stimulus conditions. For purposes of brevity, the first male stimulus tape may be regarded as the "error tape" and the second male tape as the "corrected tape." It may be recalled that the differences between those tapes were in the ratings given to the cartoons used on the 46 non-critical trials, since the other trials had a fixed script. To analyze the extent of the differences between the stimulus tapes (error, corrected, and the female tape) the sums of the ratings given by the confederates for each of the 46 cartoons were computed. These three sets of 46 summed ratings were correlated together and with the mean ratings on the same 46 cartoons given by the original 24 cartoon raters. The correlation matrix is reproduced in Table 1.

As anticipated, the summed ratings of the confederates on the first male tape (the error tape) correlated only mildly with the summed ratings of their counterparts on the female tape. Similarly, the error tape ratings correlated only mildly with the mean ratings of the original cartoon raters. The summed ratings of the three confederates who made the second or corrected male stimulus tape not only correlated moderately well with the judgments of the female confederates but were strongly correlated ($r = .75, p < .001$) with the mean ratings of the original raters. It seemed clear that some account must be taken of the impact of the first male tape in the multiple regression

Table 1
 Correlation Coefficients among Groups Rating 46
 Cartoons used in Non-critical Trials of the
 Conformity Induction

| | 1 | 2 | 3 | 4 |
|-------------------------------------|-----|------------------|------------------|------------------|
| 1. Female Confederates | 1.0 | .29 ^a | .50 ^b | .54 ^b |
| 2. Male tape--error | | 1.0 | .37 ^a | .30 ^a |
| 3. Male tape--corrected | | | 1.0 | .75 ^b |
| 4. Original raters of both sexes | | | | 1.0 |

^a
p < .05

^b
p < .001

analyses involving data from male subjects, if the data from the first fifteen male subjects were to be used.

Overall Characteristics of the Sample

The overall distribution of subjects into sex-role categories and high and low levels of self-esteem is presented in Table 2. Since the regression analyses to be employed are concerned with continuous dimensions rather than with group differences, the presentation of this data is strictly for descriptive purposes. The distribution is fairly typical with the largest proportion of subjects falling into the sex-appropriate category and the smallest into the opposite-sex-appropriate category.

Table 3 presents the distribution of the actual participants in the conformity induction based on the same categories. The problems associated with the known correlation between Masculinity and Self-esteem are evident in the unavailability of subjects in certain categories. For example, there are no masculine women who actually fell below the mean level of self-esteem.

The simple correlations between questionnaire variables are presented separately for males and females in Table 4. It is clear that the correlation between Self-esteem and Masculinity has been replicated in this sample. Similarly, the Masculinity and Femininity scales of the BSRI are again demonstrated to be empirically orthogonal. The most important result described in this table, however, is the clear independence of Femininity and Self-esteem for females as well as for males. Very mild correlations are also observed between Bem's so-called Social Desirability scale and the separate M

Table 2
 Distribution of 165 Male and 212 Female Undergraduates
 By Sex-Role Category and Level of Self-Esteem

| | Sex-Role Category ^a | | | | S-E % |
|---------------------------|--------------------------------|-----|-----|-----|----------|
| | A | M | F | I | |
| Males (<u>N</u> = 165) | | | | | |
| Self-Esteem | | | | | |
| High | 22 | 59 | 7 | 12 | 61% |
| Low | 10 | 19 | 9 | 27 | 39% |
| Sex-Role % | 19% | 47% | 10% | 24% | |
| Females (<u>N</u> = 212) | | | | | |
| Self-Esteem | | | | | |
| High | 37 | 13 | 36 | 12 | 46% |
| Low | 12 | 10 | 46 | 20 | 54% |
| Sex-Role % | 23% | 11% | 46% | 20% | |

^aSex-Role Categories

A = Androgynous

M = Masculine

F = Feminine

I = Indeterminate

Table 3
 Distribution of Conformity Induction Participants
 by Sex-Role Category and Level of Self-esteem

| | Sex-Role Category* | | | |
|----------------------|--------------------|---|---|---|
| | A | M | F | I |
| Females ($N = 24$) | | | | |
| Self-esteem | | | | |
| High | 3 | 6 | 4 | 3 |
| Low | 3 | 0 | 2 | 3 |
| Males ($N = 29$) | | | | |
| Self-esteem | | | | |
| High | 4 | 5 | 2 | 1 |
| Low | 3 | 5 | 4 | 5 |

*Sex-Role Categories

A = Androgynous

M = Masculine

F = Feminine

I = Indeterminate

Table 4
 Correlation Matrices of the Questionnaire
 Variables for Males and Females

| Males | | | |
|------------------------------|------|------------------|-----------------------|
| | F | Soc | TSBI (Self-esteem) |
| BSRI | | | |
| 1. Masculinity (M) | .02 | .18 ^a | .63 ^b |
| 2. Femininity (F) | | .27 ^b | .10 |
| 3. Social Desirability (Soc) | | | .38 ^b |
| Females | | | |
| BSRI | | | |
| 1. Masculinity (M) | -.11 | .22 ^b | .65 ^b |
| 2. Femininity (F) | | .28 ^b | .02 |
| 3. Social Desirability (Soc) | | | .44 ^b |

^ap < .05

^bp < .001

and F scales. This would suggest a mild tendency for an individual who describes himself or herself as more masculine or more feminine to also endorse sex-neutral adjectives. Of some interest is the fact that moderate correlations between the Social Desirability scale and the TSBI occur, indicating the possibility that the inclusion of the Soc subscale, composed of positive and negative sex-neutral adjectives, enhances the prediction of the self-esteem scores when using BSRI variables as predictors. These correlations then provide a framework for the consideration of the first of the two questions this study was designed to answer.

The Relationship of Sex-Role Variables to Self-esteem

In order to directly examine the impact of Femininity scores on self-esteem, which has been an area of some contention in the literature, two separate regression equations having self-esteem scores from the TSBI as the dependent variable were constructed for each sex. The first equation for both sexes employed four predictor variables--Masculinity, Femininity, the MF product or interaction term, and Social Desirability (see Tables 5 and 7). The second eliminated Femininity and MF from the equation, depending solely on Masculinity and Social Desirability to predict Self-esteem (see Tables 6 and 8). A comparison of the squared multiple R 's reveals that the two variable solution explains 51.019% of the common variance for females, and that the inclusion of Femininity and the MF product term in the four-variable equation improves this prediction only to 51.585% of the variance. Similarly, for males, the two variable solution, using M and Soc as predictor variables, explains 47.712% of the self-esteem

Table 5
 Multiple Regression: Females
 Predicting Self-Esteem from Masculinity, Femininity, Social
 Desirability, and the Masculinity-Femininity Product

| Source | <u>df</u> | <u>SS</u> | <u>MS</u> | <u>F</u> | <u>p</u> |
|---------------------|-----------|-----------|-----------|---------------------------|----------|
| Masculinity | 1 | 5.532 | 5.532 | .107 | n.s. |
| Femininity | 1 | 121.05 | 121.05 | 2.339 | n.s. |
| Social Desirability | 1 | 1893.426 | 1893.426 | 36.578 | .0001 |
| M x F | 1 | 125.033 | 125.033 | 2.415 | n.s. |
| Residual | 207 | 10714.237 | 51.7596 | | |
| Total SS = 22129.81 | | | | $R^2 = .51585, p < .0005$ | |

Table 6
Multiple Regression: Females
Predicting Self-Esteem from Masculinity and Social Desirability

| Source | <u>df</u> | <u>SS</u> | <u>MS</u> | <u>F</u> | <u>p</u> |
|----------------------|-----------|-----------|-----------|---------------------------|----------|
| Masculinity | 1 | 7043.191 | 7043.191 | 135.8 | .0001 |
| Social Desirability | 1 | 2030.97 | 2030.97 | 39.16 | .0001 |
| Residual | 209 | 10839.3 | 51.86 | | |
| Total SS = 22129.806 | | | | $R^2 = .51019, p < .0005$ | |

Table 7
 Multiple Regression: Males
 Predicting Self-Esteem from Masculinity, Femininity, Social
 Desirability, and the Masculinity-Femininity Product

| Source | <u>df</u> | <u>SS</u> | <u>MS</u> | <u>F</u> | <u>p</u> |
|---------------------|-----------|-----------|-----------|---------------------------|----------|
| Masculinity | 1 | 79.28 | 79.28 | 1.79 | n.s. |
| Femininity | 1 | 35.62 | 35.62 | .80 | n.s. |
| Social Desirability | 1 | 987.424 | 987.424 | 22.27 | .0001 |
| M x F | 1 | 39.846 | 39.846 | .897 | n.s. |
| Residual | 160 | 7093.76 | 44.34 | | |
| Total SS = 13645.99 | | | | $R^2 = .48016, p < .0005$ | |

Table 8

Multiple Regression: Males

Predicting Self-Esteem from Masculinity and Social Desirability

| Source | <u>df</u> | <u>SS</u> | <u>MS</u> | <u>F</u> | <u>p</u> |
|----------------------|-----------|-----------|-----------|---------------------------|----------|
| Masculinity | 1 | 4498.8 | 4498.8 | 102.14 | .0001 |
| Social Desirability | 1 | 1045.74 | 1045.74 | 23.744 | .0001 |
| Residual | 162 | 73135.21 | 44.04 | | |
| Total SS = 13645.993 | | | | $R^2 = .47712, p < .0005$ | |

variance and the four-variable equation explains 48.016%. In both cases the gain in precision would seem to be negligible and the appropriate F-tests confirm this view (for females, $F(2,207) = 1.21$, n.s.; for males, $F(2,160) = .468$, n.s.).

In this sample there would seem to be substantial variance held in common between the M score (based on a set of male-appropriate positive adjectives), taken together with the Soc score (based on a set of sex-neutral positive and negative adjectives), on the one hand and Self-esteem as measured by the TSBI on the other. No support is given to the contention that higher levels of femininity are associated with enhanced self-esteem, even for female subjects, for whom femininity is presumed to be sex-appropriate. Neither is there support for the idea that Masculinity and Femininity interact to enhance self-esteem.

For females in this analysis, Masculinity alone accounted for 42% of the TSBI Self-esteem variance ($r = .65$, $p < .0001$) and Social Desirability contributed uniquely to that linear prediction above the prediction based on Masculinity ($R = .71$, $F(2,209) = 108.85$, $p < .0001$). For males, Masculinity alone accounted for 40% of the Self-esteem variance ($r = .63$, $p < .0001$) and Social Desirability again made a unique contribution above that ($R = .69$, $F(2,160) = 73.91$, $p < .0001$). The question of the contribution of femininity to enhanced self-esteem would be answered negatively by these data, a fact which presents a distinct theoretical problem for those who advocate the superiority of the androgynous individual chiefly on the basis of higher levels of self-esteem.

Self-esteem, Sex-Role Orientation and Independence

The most important issue under consideration in this study was the replicability of the Conformity Induction findings reported by Bem (1975), and the respective contributions of Masculinity and Self-esteem to the prediction of independence, as defined by the conformity task. It has just been seen that there exists a strong empirical tie between Masculinity and Self-esteem. Earlier it was pointed out that although no definitive prediction was possible on the basis of a review of previous research, a "working hypothesis" was formed. This hypothesis stated that Self-esteem would predict the dimension of resistance to conformity, or "Independence," and that Masculinity, insofar as it is related to Self-esteem would reflect this correlation between Self-esteem and Independence. However, it was predicted that Masculinity would not contribute uniquely to the explained variance over and above the prediction based solely on TSBI self-esteem scores. The results did not support this hypothesis. Since it has been noted that methodological irregularities affected the data collected from male subjects, the results from the female subjects will be considered first.

In order to examine the effects of the questionnaire variables on Independence a regression equation was formed using Independence as the dependent variable and Masculinity, Femininity, MF, and Self-esteem as the independent variables. The results of this analysis appear in Table 9. For the female subjects, the four variables taken together produced a multiple R of .35 ($F(4, 19) = .677, n.s.$). When the impact of the three sex-role predictors over and above the

Table 9
Multiple Regression: Females
Predicting Independence from Masculinity, Femininity,
M x F, and Self-Esteem

| Source | <u>df</u> | <u>SS</u> | <u>MS</u> | <u>F</u> | <u>p</u> |
|--------------------|-----------|-----------|-----------|---------------------|----------|
| Masculinity | 1 | 5.0 | 5.0 | .0127 | n.s. |
| Femininity | 1 | 0.0 | 0.0 | .0 | n.s. |
| Self-Esteem | 1 | 167.35 | 167.35 | .407 | n.s. |
| M x F | 1 | 20.91 | 20.91 | .051 | n.s. |
| Residual | 19 | 7819.41 | 411.55 | | |
| Total SS = 8934.96 | | | | $R^2 = .125$, n.s. | |

prediction of Self-esteem was analyzed it was also shown to be non-significant ($F(3, 19) = .895, \underline{n.s.}$). The simple correlation between Independence and Masculinity was both small and in the opposite direction to that predicted by Bem ($r = -.20, p = .17$). Similarly, the simple correlation between Self-esteem and Independence revealed the two to be virtually orthogonal ($r = .03, p = .44$). Thus for females, performance in the conformity induction seemed unrelated to any of the variables under consideration.

The results were similar for males, although the analyses were more complicated. A fifth predictor variable (TAPNUM for Tape number) was added to account for variance uniquely explained by the differences in stimulus tapes used with male subjects. There was a significant, positive, correlation between male independence scores and tape number ($r = .42, p < .025$) indicating that scores on the error tape were significantly lower than scores on the second, corrected tape. An equation predicting Independence from M, F, MF, SE and TAPNUM was constructed. The results of this analysis are shown in Table 10, and again were non-significant. When the variance exclusively attributable to the tape-group was removed statistically, the resulting F-test for the other four predictor variables was not significant ($F(4, 23) = 1.276, \underline{n.s.}$). These data indicate that none of the predictor variables significantly explain the variance of the observed Independence scores. Thus for males as well as females, neither Self-esteem nor Masculinity were effective predictors of conformity performance.

Table 10
 Multiple Regression: Males
 Predicting Independence from Masculinity, Femininity,
 M x F, Self-Esteem, and Tape Number

| Source | <u>df</u> | <u>SS</u> | <u>MS</u> | <u>F</u> | <u>p</u> |
|----------------------|-----------|-----------|-----------|---------------------|----------|
| Masculinity | 1 | 2184.296 | 2184.296 | 3.298 | .10 |
| Femininity | 1 | 1731.79 | 1731.79 | 2.615 | n.s. |
| M x F | 1 | 2012.59 | 2012.59 | 3.039 | .10 |
| Self-Esteem | 1 | 357.245 | 357.245 | .539 | n.s. |
| Tape Number | 1 | 2579.65 | 2579.63 | 3.895 | .10 |
| Residual | 23 | 15232.743 | 662.29 | | |
| Total SS = 22682.207 | | | | $R^2 = .328$, n.s. | |

Sex Differences

One additional question remains, regarding sex differences which have been observed in other conformity induction experiments. For this analysis, female subjects were coded as belonging to the corrected male tape group. This is consistent with the observation made earlier of the closer covariance observed between the ratings of "non-critical" cartoons by female confederates to the ratings on the corrected male stimulus tape than to the ratings on the error tape.

To answer the question of sex differences, an F was computed to test the significance of the contribution of Sex to the prediction of Independence scores over and above the prediction made based on TAPNUM or tape group. The results of the analysis are summarized in Table 11. The F-test computed for the variance uniquely predicted by sex of subject proved to be non-significant as well ($F(1,50) = 1.34, n.s.$). This would indicate that for these samples, there were no differences observed in Independence scores which were directly related to sex of subject. It might be germane to note that these subjects were especially selected to evenly represent sex-role categories that are not distributed in the population evenly, but for the fact that no effect on Independence was observed for sex-role for either sex.

Summary

To summarize, the evidence presented indicates:

1. Although Self-esteem is positively and significantly related to Masculinity, and although the linear prediction based on Masculinity may be improved by consideration of the Social Desirability

Table 11
 Multiple Regression: Males and Females
 Predicting Independence from Sex and Tape Number

| Source | <u>df</u> | <u>SS</u> | <u>MS</u> | <u>F</u> | <u>p</u> |
|--------------------|-----------|-----------|-----------|-------------------------|----------|
| Sex | 1 | 738.299 | 738.299 | 1.34 | n.s. |
| Tape Number | 1 | 4068.846 | 4068.846 | 7.38 | .01 |
| Residual | 50 | 27548.12 | 550.96 | | |
| Total SS = 37629.9 | | | | $R^2 = .268, p < .0005$ | |

score as well, there was no evidence that Femininity contributed to the prediction of self-esteem in either an additive or and interactive fashion.

2. Subjects in this study were no more or less likely to conform in their ratings of stimulus cartoons because of their sex, their sex-role orientation, or their measured self-esteem.

3. The relationship demonstrated by Bem between masculinity and independence was not replicated, and the results obtained failed to support the contention that the behavioral dimension of resistance to conformity is predictable by knowledge of the subject's self-attributed sex-role orientation.

CHAPTER V

DISCUSSION AND CONCLUSIONS

This study undertook the investigation of two aspects of the relationship between sex-role orientation and self-esteem. The first of these aspects was the nature of the relationship between the orthogonal dimensions of Masculinity and Femininity on the one hand, and Self-esteem on the other. The results presented indicated that, as the previous literature implied, there exists a fairly strong correlation between Masculinity and Self-esteem but that Femininity is virtually orthogonal to level of Self-esteem for both sexes. The second aspect under consideration reflected a concern with the relative importance of Masculinity and Self-esteem, which had been shown to be empirically related variables, to the prediction of a particular behavior. The hypothesis under consideration stated that given the shared variance between Masculinity and Self-esteem, Masculinity would not significantly predict performance by subjects in a standard conformity paradigm situation over and above the prediction based on Self-esteem alone. Insofar as neither Self-esteem or Masculinity proved predictive of performance in the conformity situation, this hypothesis was disconfirmed.

The replication of the finding that Femininity is not related to Self-esteem even for females is theoretically problematic for those who have argued on the basis of mild correlations that both

Masculinity and Femininity are important to high levels of Self-esteem. It may be recalled that the normative assumption of the "Androgyny model" is that the androgynous person is at a psychological advantage by virtue of the integration of masculine and feminine characteristics or behaviors. The advantage that has been attributed to androgynous individuals based on higher levels of self-esteem would seem to be spurious.

A further implication of this finding is that there would appear to be no direct relationship between a well-established sex-appropriate sex-role and enhanced self-esteem, at least for females. If developing a feminine self-concept is unimportant to establishing a healthy and positive sense of self the socialization of girls in our society would seem to be systematically relegating them to lower levels of self-esteem. The implications are particularly far-reaching when consideration is given to the fact that nearly half the women participating in the study were classified as feminine in sex-role orientation.

The results of this study with regard to the connection between Masculinity and Self-esteem would seem to prompt more questions than are answered. The meaning of this correlation remains unspecified. If we assume that the observed correlation represents an accurate view of the relationship between the psychological constructs of masculine self-concept and self-esteem, the implication is that feminine individuals must have low self-esteem, an extremely counter-intuitive interpretation. Alternatively, it could be said that self-esteem measures accurately reflect the general superiority or higher status of males in our culture.

Another type of assumption might be that for some as yet not understood reason, the measurement of Masculinity is artifactually incorporating self-esteem variance. Until such time as other means of measuring or effectively operationalizing masculinity are developed it will be difficult to establish this possibility with certainty or to rule it out. This is clearly one area on which future research could focus. It may be recalled that several writers have pointed to the fact that M-F measurement varies depending on the type of item employed in the questionnaire (Constantinople 1973; Worrell 1979). The restriction of the universe of masculinity-femininity items to personality characteristics can be seen in some ways to be effective as a means of improving the psychometric accuracy and integrity of an instrument. However, it can also in some ways be seen as arbitrary. The speculative hypothesis may be advanced that the relationship of the M scale of the BSRI to the global construct of "masculinity" parallels the relationship of the Block Design subtest of the Wechsler Adult Intelligence Scale to overall intelligence. It may represent a significant part of the whole, but merely one aspect or dimension. If this were the case it might explain the apparent anomaly of a connection between sex-role orientation and self-esteem in which masculinity is the only important correlate to self-esteem for females as well as males.

With regard to the relationships between self-esteem, sex-role orientation and independence, two questions emerge from the results: first, what were the flaws in the present study that could have obscured or confounded the predicted relationships between these

variables; and, second, is the conformity paradigm, selected by Bem for her original study and replicated here, a valid and adequate criterion behavior for masculinity as a construct. These will be considered separately.

Replication: Empirical Considerations

In general terms, this study replicated Bem's methodological description fairly closely. The deception employed appeared to work fairly convincingly in all but a few cases. Unlike the original study subjects were divided into four categories rather than three. The basis of this division was not a median split but a split based on statistical means with the removal of borderline cases from inclusion in the conformity induction. As advocated by Bem (1977), a multiple regression analysis was made with the addition of an interaction term to help explore the combined effects of M and F.

Aside from these differences there were three departures from Bem's methodology which may have had an impact on the final results: (1) cartoons to be rated were placed in notebooks rather than flashed on a screen, (2) automatic data-recording devices were not available, and subjects were therefore requested to record their own responses on an answer sheet, and (3) the stimulus tape for one-half of the male sample tended to display a marked uniformity between the confederates even on non-critical trials.

With regard to the fact that the cartoons were bound in notebooks, it is to be conceded that this may have invited speculation on the part of subjects that the cartoons they were rating were in fact somehow different than those being rated by the other subjects. It

can only be said that subjects were explicitly told at the outset that they would be rating the same cartoons. Since Bem's subjects were sitting in separate booths also, it is not inconceivable that her subjects could have experienced the same doubts.

The recording of the data on answer sheets may have been a major source of problems with the data. Very often subjects would in fact jot down their response before hearing all of the voices of the confederates giving artificial ratings. The differences due to the error tape used with the males were statistically controlled. It should be noted that the pattern of results was the same for males as for females, indicating that the impact of the error tape was probably not crucial.

Replication: Theoretical Considerations

In light of the fact that the multiple regression analyses performed by Bem to re-examine the data from the original (1975) conformity study were also non-significant and the multiple R 's predicting independence from masculinity and femininity rather small, the question of the use of the conformity paradigm as an appropriate criterion for masculine or independent behavior assumes some importance. It was briefly noted in the review of the literature that in general females seemed to be more conforming than males (McGuire 1969). Since the focus of this study has been toward emphasizing the differences between sex-role groupings rather than differences based on sex, this fact was not elaborately discussed.

There is evidence however that this general trend toward more conformity in females is modified by the type of stimulus used in

the conformity situation. Sistrunk and McDavid (1971) performed a series of conformity type experiments altering the judgmental tasks. On the basis of the observed results they argue that the sex-determined variance observed on independence could be reduced in one of three ways: (1) by controlling the personality or motivational variables, (2) by controlling the content of the task, and (3) by controlling the sex of the influencer relative to the subject. We would expect females to conform more, for example, when the content of the task is masculine. It may well be that rating cartoons is a sex-neutral task which would result in no sex differences and very possibly no sex-role differences.

The only substantiation offered by Bem for the validity of this task as a criterion for masculine behavior was the ratings given by judges to the item, "Saying what you believe, even when you know those around you disagree." On a scale ranging from -3 (Very masculine) to +3 (Very feminine) both males' and females' average ratings for this item were approximately -1 (Bem 1975). The combined evidence of the re-analysis of her data using multiple regression and the evidence of this study suggests that either this type of conformity paradigm is not an adequate measure of masculine behavior or that the M scale of the BSRI is not predictive of masculine behavior. The evidence of Sistrunk and McDavid might suggest the former. The arguments of Spence and Helmreich (1978) against a "strong hypothesis" would admit to the latter.

Summary

The overarching concern of this investigation was the relative importance of Masculinity and Self-esteem in predicting a third measure which was intended to represent an operational definition of masculinity. It is clear from our sample that a strong correlation exists between Self-esteem as measured by the Texas Social Behavior Inventory (Form A) and Masculinity as measured by the Bem Sex-Role Inventory. In the absence of significant prediction of independence based on either variable it is impossible to specify the exact nature of the relationship between Masculinity and Self-esteem. The question of validity for the claim that the BSRI is an accurate technique for appraisal of the global qualities of Masculinity and Femininity is still an open and unanswered one. It has been speculated that resistance to conformity is an inadequate criterion for masculine behavior. Furthermore, the possibility has been raised, as a direction for research, that the extreme correlation of Masculinity and Self-esteem may result from the limited scope of the items employed to measure masculinity. These issues deserve further theoretical and empirical consideration.

APPENDIX

APPENDIX

RAW SCORES ON THE BEM SEX-ROLE INVENTORY AND THE TEXAS
SOCIAL BEHAVIOR INVENTORY FOR 212 FEMALE AND 165 MALE SUBJECTS

MALES

FALL SEMESTER

| <u>Subject Number</u> | <u>Masculinity</u> | <u>Femininity</u> | <u>Social Desirability</u> | <u>Self- Esteem</u> |
|---------------------------|--------------------|-------------------|--------------------------------|-------------------------|
| 1 | 106 | 84 | 99 | 45 |
| 2 | 106 | 90 | 105 | 54 |
| 3 | 96 | 91 | 113 | 44 |
| 5 | 69 | 76 | 91 | 26 |
| 10 | 101 | 91 | 100 | 43 |
| 11 | 107 | 95 | 89 | 42 |
| 12 | 97 | 92 | 118 | 48 |
| 13 | 90 | 93 | 97 | 28 |
| 14 | 106 | 73 | 96 | 37 |
| 28 | 105 | 99 | 99 | 51 |
| 33 | 113 | 90 | 82 | 44 |
| 34 | 114 | 85 | 90 | 44 |
| 37 | 112 | 94 | 110 | 53 |
| 42 | 100 | 107 | 98 | 41 |
| 43 | 109 | 74 | 103 | 36 |
| 46 | 102 | 96 | 106 | 56 |
| 47 | 104 | 90 | 97 | 53 |
| 50 | 67 | 80 | 98 | 25 |
| 54 | 112 | 94 | 106 | 47 |
| 58 | 99 | 112 | 105 | 52 |
| 62 | 85 | 83 | 89 | 24 |
| 65 | 103 | 107 | 110 | 55 |
| 86 | 92 | 106 | 99 | 32 |
| 87 | 123 | 59 | 99 | 56 |
| 88 | 90 | 100 | 95 | 40 |
| 89 | 110 | 86 | 92 | 46 |
| 90 | 90 | 97 | 108 | 45 |
| 91 | 124 | 113 | 107 | 52 |
| 92 | 106 | 82 | 98 | 48 |
| 93 | 112 | 83 | 92 | 36 |
| 94 | 93 | 88 | 102 | 45 |
| 95 | 99 | 93 | 102 | 44 |
| 96 | 127 | 85 | 108 | 61 |
| 97 | 123 | 82 | 103 | 49 |
| 98 | 95 | 88 | 93 | 42 |
| 103 | 115 | 93 | 99 | 51 |
| 104 | 89 | 97 | 113 | 43 |
| 105 | 114 | 85 | 91 | 52 |
| 112 | 110 | 88 | 108 | 46 |
| 117 | 81 | 104 | 102 | 27 |

| <u>Subject Number</u> | <u>Masculinity</u> | <u>Femininity</u> | <u>Social Desirability</u> | <u>Self- Esteem</u> |
|---------------------------|--------------------|-------------------|--------------------------------|-------------------------|
| 121 | 123 | 80 | 90 | 48 |
| 122 | 83 | 98 | 107 | 42 |
| 123 | 104 | 99 | 83 | 41 |
| 125 | 112 | 81 | 98 | 32 |
| 126 | 99 | 107 | 101 | 37 |
| 127 | 79 | 97 | 86 | 23 |
| 128 | 111 | 88 | 91 | 41 |
| 129 | 97 | 99 | 100 | 37 |
| 130 | 99 | 86 | 98 | 43 |
| 132 | 114 | 92 | 66 | 37 |
| 133 | 99 | 98 | 97 | 35 |
| 135 | 90 | 99 | 101 | 41 |
| 139 | 95 | 82 | 110 | 46 |
| 147 | 107 | 97 | 96 | 38 |
| 153 | 99 | 83 | 97 | 31 |
| 159 | 76 | 89 | 108 | 33 |
| 166 | 99 | 113 | 93 | 35 |
| 168 | 107 | 84 | 95 | 43 |
| 169 | 96 | 89 | 84 | 30 |
| 186 | 93 | 101 | 103 | 48 |
| 192 | 112 | 104 | 103 | 42 |
| 193 | 107 | 77 | 102 | 45 |
| 195 | 95 | 91 | 87 | 38 |
| 196 | 101 | 59 | 90 | 28 |
| 201 | 95 | 82 | 97 | 35 |
| 204 | 116 | 88 | 121 | 57 |
| 206 | 106 | 98 | 99 | 33 |
| 207 | 103 | 110 | 110 | 42 |
| 208 | 115 | 98 | 83 | 49 |
| 209 | 99 | 115 | 100 | 27 |
| 210 | 118 | 101 | 110 | 59 |
| 213 | 90 | 91 | 91 | 41 |
| 217 | 101 | 90 | 108 | 42 |
| 218 | 95 | 76 | 94 | 40 |
| 219 | 74 | 93 | 82 | 29 |
| 224 | 96 | 93 | 105 | 45 |
| 231 | 110 | 87 | 82 | 32 |
| 232 | 103 | 97 | 96 | 43 |
| 235 | 88 | 85 | 104 | 40 |
| 236 | 96 | 87 | 102 | 49 |
| 237 | 111 | 88 | 98 | 47 |
| 239 | 111 | 84 | 99 | 51 |
| 241 | 92 | 119 | 120 | 46 |
| 247 | 80 | 114 | 98 | 36 |

SPRING SEMESTER

| <u>Subject Number</u> | <u>Masculinity</u> | <u>Femininity</u> | <u>Social Desirability</u> | <u>Self- Esteem</u> |
|---------------------------|--------------------|-------------------|--------------------------------|-------------------------|
| 601 | 105 | 109 | 116 | 48 |
| 602 | 97 | 83 | 99 | 47 |
| 603 | 104 | 103 | 101 | 42 |
| 604 | 110 | 89 | 96 | 45 |
| 609 | 93 | 85 | 93 | 30 |
| 611 | 115 | 94 | 106 | 42 |
| 612 | 93 | 93 | 118 | 39 |
| 613 | 92 | 84 | 95 | 37 |
| 614 | 109 | 87 | 104 | 46 |
| 615 | 106 | 99 | 97 | 42 |
| 616 | 84 | 90 | 104 | 35 |
| 617 | 66 | 69 | 62 | 38 |
| 618 | 98 | 96 | 102 | 36 |
| 620 | 106 | 84 | 100 | 46 |
| 621 | 102 | 71 | 100 | 31 |
| 624 | 99 | 77 | 105 | 41 |
| 626 | 108 | 93 | 94 | 34 |
| 627 | 111 | 91 | 115 | 39 |
| 628 | 85 | 86 | 96 | 33 |
| 633 | 111 | 99 | 103 | 44 |
| 637 | 85 | 101 | 104 | 32 |
| 641 | 91 | 92 | 116 | 45 |
| 651 | 99 | 90 | 109 | 43 |
| 652 | 95 | 102 | 108 | 40 |
| 653 | 100 | 103 | 104 | 38 |
| 654 | 118 | 110 | 97 | 47 |
| 657 | 114 | 71 | 101 | 48 |
| 658 | 82 | 90 | 106 | 31 |
| 660 | 93 | 103 | 108 | 43 |
| 661 | 120 | 93 | 101 | 52 |
| 662 | 124 | 110 | 102 | 53 |
| 666 | 70 | 87 | 86 | 23 |
| 668 | 96 | 77 | 75 | 29 |
| 669 | 120 | 102 | 94 | 39 |
| 670 | 107 | 94 | 111 | 43 |
| 677 | 110 | 93 | 102 | 38 |
| 680 | 118 | 92 | 118 | 55 |
| 688 | 112 | 88 | 98 | 48 |
| 691 | 107 | 97 | 109 | 39 |
| 692 | 100 | 77 | 104 | 37 |
| 693 | 114 | 102 | 92 | 42 |
| 694 | 104 | 82 | 108 | 44 |
| 695 | 113 | 91 | 115 | 44 |
| 696 | 121 | 86 | 95 | 56 |

| <u>Subject Number</u> | <u>Masculinity</u> | <u>Femininity</u> | <u>Social Desirability</u> | <u>Self- Esteem</u> |
|---------------------------|--------------------|-------------------|--------------------------------|-------------------------|
| 697 | 111 | 97 | 107 | 41 |
| 700 | 98 | 83 | 104 | 44 |
| 702 | 116 | 89 | 103 | 50 |
| 703 | 94 | 106 | 100 | 37 |
| 705 | 97 | 100 | 87 | 33 |
| 710 | 74 | 69 | 82 | 33 |
| 711 | 69 | 94 | 102 | 38 |
| 712 | 129 | 81 | 93 | 50 |
| 714 | 103 | 83 | 105 | 35 |
| 718 | 84 | 89 | 84 | 30 |
| 723 | 97 | 98 | 104 | 46 |
| 727 | 107 | 104 | 110 | 42 |
| 728 | 116 | 97 | 99 | 54 |
| 729 | 107 | 99 | 96 | 43 |
| 731 | 84 | 101 | 101 | 38 |
| 732 | 87 | 75 | 86 | 42 |
| 733 | 100 | 81 | 93 | 44 |
| 737 | 102 | 82 | 96 | 48 |
| 738 | 107 | 99 | 107 | 47 |
| 739 | 108 | 90 | 107 | 47 |
| 741 | 106 | 96 | 103 | 53 |
| 743 | 97 | 96 | 97 | 44 |
| 744 | 118 | 100 | 100 | 48 |
| 747 | 120 | 94 | 105 | 55 |
| 748 | 92 | 83 | 98 | 30 |
| 749 | 96 | 111 | 86 | 47 |
| 750 | 102 | 97 | 103 | 40 |
| 752 | 101 | 111 | 96 | 40 |
| 754 | 97 | 101 | 89 | 34 |
| 755 | 81 | 94 | 91 | 33 |
| 756 | 106 | 87 | 103 | 41 |
| 760 | 115 | 85 | 101 | 43 |
| 763 | 95 | 83 | 90 | 47 |
| 764 | 113 | 89 | 102 | 54 |
| 767 | 103 | 73 | 97 | 35 |
| 768 | 93 | 76 | 93 | 36 |

FEMALESFALL SEMESTER

| | | | | |
|----|-----|-----|-----|----|
| 2 | 79 | 117 | 107 | 28 |
| 4 | 106 | 120 | 104 | 39 |
| 6 | 100 | 89 | 112 | 40 |
| 7 | 82 | 111 | 101 | 48 |
| 8 | 99 | 101 | 106 | 44 |
| 9 | 78 | 104 | 112 | 46 |
| 15 | 111 | 110 | 120 | 58 |

| <u>Subject Number</u> | <u>Masculinity</u> | <u>Femininity</u> | <u>Social Desirability</u> | <u>Self- Esteem</u> |
|---------------------------|--------------------|-------------------|--------------------------------|-------------------------|
| 16 | 106 | 96 | 101 | 46 |
| 17 | 91 | 107 | 105 | 41 |
| 18 | 93 | 110 | 107 | 36 |
| 19 | 110 | 100 | 103 | 48 |
| 20 | 99 | 105 | 96 | 31 |
| 20 | 87 | 111 | 109 | 45 |
| 23 | 87 | 96 | 93 | 35 |
| 24 | 66 | 105 | 102 | 31 |
| 26 | 82 | 107 | 87 | 28 |
| 27 | 110 | 99 | 113 | 45 |
| 29 | 105 | 100 | 115 | 44 |
| 30 | 101 | 103 | 112 | 45 |
| 31 | 57 | 93 | 96 | 16 |
| 32 | 99 | 117 | 105 | 45 |
| 35 | 92 | 114 | 125 | 44 |
| 36 | 83 | 99 | 104 | 36 |
| 38 | 93 | 107 | 96 | 40 |
| 39 | 78 | 83 | 93 | 35 |
| 40 | 68 | 85 | 95 | 28 |
| 41 | 97 | 113 | 124 | 43 |
| 44 | 72 | 108 | 103 | 33 |
| 45 | 63 | 103 | 95 | 29 |
| 48 | 88 | 99 | 105 | 40 |
| 49 | 110 | 102 | 92 | 46 |
| 51 | 85 | 112 | 103 | 32 |
| 52 | 105 | 107 | 112 | 46 |
| 53 | 92 | 118 | 116 | 46 |
| 55 | 88 | 106 | 96 | 36 |
| 56 | 84 | 120 | 106 | 33 |
| 57 | 83 | 112 | 99 | 17 |
| 59 | 66 | 95 | 111 | 45 |
| 60 | 101 | 70 | 90 | 40 |
| 61 | 92 | 85 | 99 | 51 |
| 63 | 95 | 91 | 104 | 40 |
| 64 | 82 | 103 | 106 | 26 |
| 99 | 100 | 80 | 84 | 29 |
| 100 | 76 | 110 | 105 | 34 |
| 101 | 80 | 111 | 111 | 44 |
| 102 | 112 | 102 | 112 | 39 |
| 106 | 98 | 97 | 96 | 38 |
| 107 | 77 | 114 | 106 | 28 |
| 108 | 103 | 77 | 86 | 35 |
| 109 | 88 | 90 | 106 | 36 |
| 110 | 82 | 95 | 110 | 34 |
| 113 | 86 | 110 | 102 | 39 |
| 115 | 74 | 109 | 103 | 42 |
| 116 | 74 | 105 | 97 | 31 |

| <u>Subject Number</u> | <u>Masculinity</u> | <u>Femininity</u> | <u>Social Desirability</u> | <u>Self- Esteem</u> |
|---------------------------|--------------------|-------------------|--------------------------------|-------------------------|
| 118 | 74 | 105 | 97 | 31 |
| 119 | 91 | 99 | 113 | 39 |
| 120 | 99 | 118 | 121 | 53 |
| 124 | 68 | 99 | 99 | 21 |
| 127 | 116 | 100 | 97 | 56 |
| 136 | 92 | 112 | 110 | 42 |
| 138 | 122 | 102 | 96 | 38 |
| 141 | 118 | 110 | 105 | 43 |
| 143 | 97 | 114 | 97 | 46 |
| 144 | 69 | 116 | 92 | 21 |
| 146 | 89 | 105 | 96 | 31 |
| 148 | 78 | 94 | 101 | 24 |
| 149 | 102 | 112 | 108 | 53 |
| 150 | 70 | 103 | 98 | 30 |
| 151 | 87 | 103 | 98 | 39 |
| 154 | 89 | 112 | 117 | 49 |
| 155 | 89 | 98 | 96 | 30 |
| 157 | 90 | 115 | 94 | 39 |
| 158 | 96 | 84 | 90 | 23 |
| 160 | 78 | 116 | 110 | 24 |
| 160 | 97 | 110 | 94 | 50 |
| 161 | 98 | 107 | 106 | 52 |
| 162 | 107 | 104 | 94 | 49 |
| 163 | 93 | 93 | 107 | 36 |
| 164 | 92 | 95 | 95 | 41 |
| 165 | 97 | 96 | 107 | 34 |
| 167 | 71 | 108 | 98 | 31 |
| 170 | 106 | 105 | 114 | 46 |
| 171 | 114 | 116 | 107 | 60 |
| 172 | 81 | 117 | 96 | 33 |
| 173 | 92 | 113 | 120 | 34 |
| 174 | 100 | 108 | 108 | 37 |
| 175 | 75 | 113 | 102 | 40 |
| 176 | 106 | 93 | 101 | 34 |
| 177 | 69 | 115 | 97 | 26 |
| 178 | 81 | 112 | 97 | 32 |
| 181 | 93 | 102 | 108 | 45 |
| 182 | 90 | 101 | 103 | 38 |
| 183 | 66 | 91 | 94 | 33 |
| 184 | 95 | 103 | 87 | 31 |
| 185 | 95 | 112 | 107 | 35 |
| 187 | 75 | 104 | 98 | 35 |
| 188 | 87 | 101 | 103 | 29 |
| 189 | 91 | 109 | 100 | 41 |
| 190 | 95 | 89 | 106 | 45 |
| 191 | 83 | 99 | 94 | 41 |
| 194 | 89 | 106 | 90 | 43 |

| <u>Subject Number</u> | <u>Masculinity</u> | <u>Femininity</u> | <u>Social Desirability</u> | <u>Self- Esteem</u> |
|---------------------------|--------------------|-------------------|--------------------------------|-------------------------|
| 198 | 75 | 106 | 97 | 37 |
| 199 | 100 | 100 | 103 | 44 |
| 202 | 83 | 121 | 100 | 39 |
| 203 | 70 | 92 | 98 | 25 |
| 211 | 85 | 85 | 75 | 21 |
| 212 | 104 | 90 | 95 | 47 |
| 214 | 103 | 93 | 97 | 31 |
| 215 | 90 | 90 | 96 | 36 |
| 216 | 98 | 120 | 102 | 44 |
| 220 | 78 | 93 | 85 | 18 |
| 221 | 65 | 101 | 95 | 27 |
| 222 | 97 | 99 | 108 | 40 |
| 223 | 81 | 102 | 95 | 49 |
| 225 | 75 | 102 | 104 | 41 |
| 226 | 86 | 106 | 92 | 32 |
| 227 | 90 | 115 | 106 | 46 |
| 229 | 73 | 107 | 99 | 21 |
| 230 | 85 | 92 | 98 | 40 |
| 233 | 88 | 95 | 85 | 34 |
| 234 | 83 | 112 | 91 | 41 |
| 238 | 87 | 119 | 111 | 42 |
| 240 | 107 | 92 | 102 | 41 |
| 242 | 100 | 102 | 102 | 47 |
| 243 | 92 | 107 | 97 | 44 |
| 244 | 93 | 110 | 102 | 40 |
| 245 | 87 | 105 | 100 | 33 |
| 246 | 98 | 99 | 104 | 40 |
| 248 | 92 | 95 | 118 | 47 |
| 250 | 96 | 94 | 103 | 35 |
| 502 | 59 | 112 | 108 | 14 |

SPRING SEMESTER

| | | | | |
|-----|-----|-----|-----|----|
| 600 | 84 | 105 | 107 | 32 |
| 605 | 108 | 89 | 93 | 41 |
| 606 | 89 | 112 | 116 | 47 |
| 607 | 71 | 100 | 101 | 32 |
| 608 | 96 | 124 | 110 | 42 |
| 610 | 94 | 108 | 98 | 45 |
| 622 | 71 | 99 | 95 | 39 |
| 623 | 96 | 95 | 105 | 42 |
| 625 | 104 | 90 | 98 | 53 |
| 629 | 115 | 92 | 120 | 54 |
| 630 | 88 | 102 | 105 | 42 |
| 631 | 94 | 106 | 111 | 42 |
| 632 | 81 | 81 | 108 | 49 |
| 634 | 88 | 98 | 108 | 46 |
| 635 | 83 | 117 | 118 | 49 |

| <u>Subject Number</u> | <u>Masculinity</u> | <u>Femininity</u> | <u>Social Desirability</u> | <u>Self- Esteem</u> |
|---------------------------|--------------------|-------------------|--------------------------------|-------------------------|
| 636 | 94 | 92 | 114 | 42 |
| 638 | 100 | 94 | 116 | 46 |
| 639 | 81 | 117 | 104 | 26 |
| 640 | 114 | 99 | 106 | 41 |
| 642 | 105 | 101 | 96 | 36 |
| 643 | 30 | 92 | 82 | 5 |
| 644 | 85 | 81 | 80 | 28 |
| 645 | 101 | 109 | 101 | 48 |
| 646 | 84 | 103 | 89 | 36 |
| 647 | 115 | 89 | 103 | 46 |
| 648 | 78 | 108 | 106 | 37 |
| 649 | 110 | 101 | 106 | 47 |
| 650 | 79 | 94 | 103 | 35 |
| 655 | 123 | 115 | 99 | 48 |
| 656 | 70 | 106 | 107 | 30 |
| 659 | 75 | 102 | 106 | 37 |
| 663 | 99 | 105 | 96 | 38 |
| 664 | 93 | 97 | 99 | 39 |
| 665 | 73 | 93 | 97 | 38 |
| 667 | 79 | 110 | 91 | 23 |
| 671 | 114 | 95 | 102 | 51 |
| 672 | 115 | 114 | 110 | 48 |
| 673 | 110 | 88 | 114 | 46 |
| 674 | 82 | 105 | 111 | 39 |
| 676 | 100 | 110 | 111 | 41 |
| 679 | 85 | 96 | 101 | 31 |
| 682 | 109 | 102 | 104 | 36 |
| 683 | 91 | 98 | 109 | 51 |
| 684 | 90 | 111 | 104 | 46 |
| 685 | 105 | 101 | 111 | 56 |
| 686 | 91 | 104 | 91 | 48 |
| 687 | 91 | 93 | 95 | 44 |
| 689 | 80 | 114 | 106 | 42 |
| 690 | 102 | 91 | 100 | 47 |
| 698 | 103 | 114 | 110 | 51 |
| 699 | 87 | 81 | 104 | 42 |
| 701 | 102 | 114 | 110 | 44 |
| 704 | 106 | 100 | 91 | 46 |
| 706 | 90 | 110 | 99 | 36 |
| 707 | 86 | 121 | 107 | 35 |
| 708 | 74 | 118 | 108 | 31 |
| 709 | 97 | 103 | 104 | 39 |
| 713 | 101 | 98 | 101 | 47 |
| 715 | 92 | 104 | 100 | 40 |
| 716 | 96 | 93 | 113 | 37 |
| 717 | 96 | 122 | 106 | 35 |
| 720 | 84 | 114 | 94 | 22 |
| 721 | 82 | 111 | 97 | 31 |
| 722 | 114 | 96 | 102 | 50 |

| <u>Subject Number</u> | <u>Masculinity</u> | <u>Femininity</u> | <u>Social Desirability</u> | <u>Self- Esteem</u> |
|---------------------------|--------------------|-------------------|--------------------------------|-------------------------|
| 724 | 84 | 102 | 121 | 47 |
| 725 | 89 | 88 | 99 | 44 |
| 726 | 74 | 102 | 101 | 31 |
| 730 | 88 | 115 | 106 | 41 |
| 734 | 96 | 81 | 95 | 35 |
| 735 | 106 | 101 | 98 | 47 |
| 736 | 100 | 107 | 101 | 47 |
| 740 | 119 | 78 | 105 | 51 |
| 742 | 92 | 85 | 114 | 31 |
| 745 | 101 | 113 | 97 | 30 |
| 746 | 83 | 101 | 90 | 27 |
| 751 | 110 | 76 | 113 | 50 |
| 753 | 77 | 94 | 92 | 28 |
| 758 | 90 | 91 | 103 | 47 |
| 759 | 87 | 103 | 111 | 45 |
| 761 | 92 | 110 | 97 | 41 |
| 762 | 73 | 112 | 104 | 30 |

REFERENCES

- Bakan, D. The duality of human existence. Chicago: Rand McNally, 1966.
- Bem, S. L. Psychology looks at sex roles: Where have all the androgynous people gone? Paper presented at the UCLA Symposium on Women, May 1972. (Available from Sandra Bem, Stanford University Department of Psychology.)
- 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. On the utility of alternative procedures for assessing psychological androgyny. Journal of Consulting and Clinical Psychology, 1977, 45, 196-205.
- Bem, S. L. Beyond androgyny: Some presumptuous prescriptions for a liberated sexual identity. In J. Sherman & F. Denmark (Eds.). Psychology of women: Future directions and research. New York: Psychological Dimensions, in press.
- Bem, S. L., & Lenney, E. Sex typing and the avoidance of cross-sex behavior. Journal of Personality and Social Psychology, 1976, 33, 48-54.
- Bem, S. L., Martyna, 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. New perspectives on sex roles and personality dimensions. In R. Bednar (Chair), Sex roles: Masculine, Feminine, Androgynous, or none of the above? Symposium presented at the meeting of the American Psychological Association, Chicago, August 1975.
- 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.
- Broverman, I. K., Broverman, D. M., Clarkson, F. E., Rosenkrantz, P. S., & Vogel, S. R. Sex role stereotypes and clinical judgments of mental health. Journal of Consulting and Clinical Psychology, 1970, 34, 1-7.

- Broverman, I. K., Vogel, S. R., Broverman, D. M., Clarkson, F. E., & Rosenkrantz, P. S. Sex role stereotypes: A current appraisal. Journal of Social Issues, 1972, 28, 59-78.
- Constantinople, A. Masculinity-femininity: An exception to the famous dictum. Psychological Bulletin, 1973, 80, 389-407.
- Deutsch, C. J., & Gilbert, L. A. Sex role stereotypes: Effects on perceptions of self and others on personal adjustment. Journal of Counseling Psychology, 1976, 4, 373-379.
- Doherty, P. A., & Schmidt, M. R. Sex-typing and self-concept in college women. Journal of College Student Personnel, 1978, 19, 493-497.
- Edwards, A., & Abbott, R. Measurement of personality traits: Theory and technique. Annual Review of Psychology, 1973, 24, 241-278.
- Fay, T. L. & Brown, I. S. Sex-role preference and self-esteem in women. Abstract of a paper presented at the meeting of the Mid-western Psychological Association, Chicago, May 1979.
- Gaudreau, P. Factor analysis of the Bem Sex-Role Inventory. Journal of Consulting and Clinical Psychology, 1977, 45, 299-302.
- Heilbrun, A. B. Measurement of masculine and feminine sex role identities as independent dimensions. Journal of Consulting and Clinical Psychology, 1976, 44, 183-190.
- Helmreich, R., & Stapp, J. Short forms of the Texas Social Behavior Inventory (TSBI), an objective measure of self-esteem. Bulletin of the Psychonomic Society, 1974, 4, 473-475.
- Helmreich, R., Stapp, J., & Ervin, C. The Texas Social Behavior Inventory: An objective measure of self-esteem or social competence. JSAS Catalog of Selected Documents in Psychology, 1974, 4, 79. (Ms. No. 681)
- Ickes, W., & Barnes, R. D. The role of sex and self-monitoring in unstructured dyadic interactions. Journal of Personality and Social Psychology, 1977, 35, 315-330.
- Ickes, W., & Barnes, R. D. Boys and girls together--and alienated: On enacting stereotyped sex roles in mixed-sex dyads. Journal of Personality and Social Psychology, 1978, 36, 669-683.
- Jackson, D. N. Personality Research Form Manual. Goshen, N.Y.: Research Psychologists Press, 1967.
- Kelly, J. A., Caudill, S., Hathorn, S., & O'Brien, C. G. Socially undesirable sex-correlated characteristics: Implications for androgyny and adjustment. Journal of Consulting and Clinical Psychology, 1977, 45, 1186-1187.

- Kelly, J. A., O'Brien, C. G., Hosford, R., & Kinsinger, E. Sex roles as social skills: A behavioral analysis of "masculinity," "femininity," and "psychological androgyny." Paper presented at the meeting of the Association for the Advancement of Behavior Therapy, New York, December 1976.
- 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.
- Kelly, J. A., & Worell, L. Parent behaviors related to masculine, feminine, and androgynous sex role orientations. Journal of Consulting and Clinical Psychology, 1976, 44, 843-851.
- Lippa, R. The naive perception of masculinity-femininity on the basis of expressive cues. Journal of Research in Personality, 1978, 12, 1-14.
- Lunneborg, P. W. Stereotypic aspects in masculinity-femininity measurement. Journal of Consulting and Clinical Psychology, 1970, 34, 113-118.
- Lunneborg, P. W. Dimensionality of MF. Journal of Clinical Psychology, 1972, 28, 313-317.
- Lunneborg, P. W., & Lunneborg, C. E. Factor structure of MF scales and items. Journal of Clinical Psychology, 1970, 26, 360-366.
- McGuire, W. J. The nature of attitudes and attitude change. In G. Lindzey and E. Aronson (Eds.), The Handbook of Social Psychology. Reading, Massachusetts: Addison-Wesley Publishing Company, 1969.
- Nichols, R. C. Subtle, obvious, and stereotype measures of masculinity-femininity. Educational and Psychological Measurement, 1962, 22, 449-461.
- Parsons, T., & Bales, R. F. Family, socialization and interaction process. New York: Free Press of Glencoe, 1955.
- Pleck, J. H. Masculinity-femininity: Current and alternative paradigms. Sex Roles, 1975, 1, 161-178.
- Sistrunk, F. & McDavid, J. W. Sex variable in conforming behavior. Journal of Personality and Social Psychology, 1971, 17, 200-207.
- Spence, J. T., & Helmreich, R. L. Masculinity & Femininity, Their Psychological Dimensions, Correlates, & Antecedents. Austin, Texas: University of Texas Press, 1978.

- Spence, J. T., Helmreich, R., & 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. (Ms. No. 617)
- Spence, J. T., Helmreich, R., & Stapp, J. Ratings of self and peers on sex role attributes and their relation to self-esteem and conceptions of masculinity and femininity. Journal of Personality and Social Psychology, 1975, 32, 29-39.
- Strahan, R. F. Remarks on Bem's measurement of psychological androgyny: Alternative methods and a supplementary analysis. Journal of Consulting and Clinical Psychology, 1975, 43, 568-571.
- Tyler, L. Individual differences: Sex differences. In D. Sills (Ed.), International Encyclopedia of the Social Sciences. New York: Macmillan, 1968.
- Wells, L. E., & Marwell, G. Self-Esteem, Its Conceptualization and Measurement. Beverly Hills: Sage Publications, 1976.
- Wetter, R. E. Levels of self-esteem associated with four sex role categories. In R. Bednar (Chair) Sex roles: Masculine, Feminine, Androgynous, or none of the above? Symposium presented at the meeting of the American Psychological Association, Chicago, August 1975.
- Whetton, C., & Swindells, T. A factor analysis of the Bem Sex-Role Inventory. Journal of Clinical Psychology, 1977, 33, 150-153.
- Worell, J. Issues in the measurement and development of sex-role styles. In R. Bednar (Chair) Sex roles: Masculine, Feminine, Androgynous or none of the above? Symposium presented at the meeting of the American Psychological Association, Chicago, August 1975.