

3-31-2003

Perception of Leadership Qualities in Higher Education: Impact of Professor Gender, Professor Leader Style, Situation, and Participant Gender

Michela A. LaRocca
University of South Florida

Follow this and additional works at: <https://scholarcommons.usf.edu/etd>

 Part of the [American Studies Commons](#)

Scholar Commons Citation

LaRocca, Michela A., "Perception of Leadership Qualities in Higher Education: Impact of Professor Gender, Professor Leader Style, Situation, and Participant Gender" (2003). *Graduate Theses and Dissertations*.
<https://scholarcommons.usf.edu/etd/1415>

This Dissertation is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.

PERCEPTION OF LEADERSHIP QUALITIES IN HIGHER EDUCATION: IMPACT
OF PROFESSOR GENDER, PROFESSOR LEADER STYLE, SITUATION, AND
PARTICIPANT GENDER

by

MICHELA A. LAROCCA

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Educational Measurement & Research
College of Education
University of South Florida

Co-Major Professor: Robert F. Dedrick, Ph.D.

Co-Major Professor: John M. Ferron, Ph.D.

Bruce W. Hall, Ed.D.

Kathy Mcnelis, Ph.D.

Daphne Thomas, Ph.D.

Date of Approval:

March 31, 2003

Keywords: leadership dimension, task/personal, democratic/autocratic,
gender stereotypes, theoretical implications

© Copyright 2003 , Michela A. Larocca

Table of Contents

List of Tables	v
List of Figures	vii
Abstract	x
Chapter I - Introduction	1
Problem Statement and Purpose of the Study	3
Organization of the Study	5
Chapter II – Review of the Literature	7
Section I – Gender Role Models	7
Physical Characteristics as they Relate to Task/Person Dimension	9
Gender-role spillover	10
Gender-role congruence	11
Social role theory	12
Section II – Observed Gender Differences in Leadership Styles	13
Section III – Perception of Gender Differences in Leadership Positions and Leader Evaluation	17
Gender and Leader Competence	19
Gender and Leader Likeability	23
Interaction Between Rater Gender and Leader Gender	29
Gender and Leadership Issues in Academia	30
Section IV – Conceptual Issues	34
Leadership Dimensions	34
Summary and Purpose of Current Study	38

Chapter III – Method	41
Section I – Overview	41
Hypotheses	41
Main Effects	42
Dimension of leader competence	42
Dimension of leader masculinity/femininity	42
Dimension of leader likeability	42
Interaction Effects	42
Interaction between participant gender, leader style, and situation on competence and likeability ratings	42
Section II – Research Participants	45
Section III – Instrument Development	52
Experimental Conditions	52
Task Situation	53
Personal Situation	53
Professor Nonverbal Democratic Leadership Style for Task Situations	53
Professor Nonverbal Autocratic Leadership Style for Task Situations	54
Professor Nonverbal Democratic Leadership Style for Personal Situations	54
Professor Nonverbal Autocratic Leadership Style for Personal Situations	54
Professor – Study Dyads	54
Development of Dependent Measures	55
Procedure	57
Chapter IV – Results	59
Section I – Psychometric Analysis of the Four Leadership Dimensions	59
Competence	59

Masculinity/Femininity	59
Likeability	60
Professional Traits	60
Principal axis factoring specifying four factors	61
Principal axis procedure with two factors	64
Comparison of two factors results with originally proposed four factors	71
Overview of Confirmatory Factor Analysis Procedure (CFA)	72
Four Factors Revisited	74
Section II – Hypotheses Testing	78
Hypothesis 1	87
Hypothesis 2	87
Hypothesis 3	91
Hypothesis 4	93
Hypothesis 5	94
Additional Findings for the Competence Factor	94
Interaction effects	94
Additional Findings for the Masculinity Factor	98
Additional Findings for the Likeability Factor	100
Chapter V – Discussion	107
Section I – Summary	108
Section II – Psychometric Properties of Leadership Construct	111
Section III – Hypotheses Testing	111
Hypothesis 1: Democratic professors will be rated as significantly more competent than autocratic professors	111
Hypothesis 2: Autocratic professor will be rated as significantly more masculine than democratic professors	114
Hypothesis 3: Democratic professors will be rated as significantly more likeable than autocratic professors	113

Hypothesis 4: There will be an interaction effect between participant gender, professor gender, professor leader style and type of situation on the ratings of leader competence	114
Hypothesis 5: There will be an interaction effect between participant gender, professor gender, professor leader style and type of situation on the ratings of leader likeability	117
Additional Findings	118
College of Arts and Science	118
College of Education	119
Section IV – Future Research	120
Design	120
Implications	123
References	126
Appendices	134
Appendix A. Pilot I: Validation of Scenario Manipulation	135
Appendix B. Pilot II: Manipulation Check on Dependent Measures	151
Appendix C. Task Vignette Booklets	157
Appendix D. Personal Vignette Booklets	164
Appendix E. Survey Introduction Letter	171
About the Author	End Page

List of Tables

Table 1. Colleges Within the Participating University	45
Table 2. Student Race/Ethnicity Profile of Sample and University Population	46
Table 3. Source Information of Study Sample	47
Table 4. Reported Majors for Male and Female Arts and Science Students	48
Table 5. Reported Majors for Male and Female Education Students	49
Table 6. Summary Statistics for Age by College and Participant Gender	50
Table 7. Frequency Distribution of School Standing by College and Gender	51
Table 8. Racial/Ethnicity Frequency Distribution by College and Gender	52
Table 9. Initial Items for the Four Leadership Dimensions	57
Table 10. Summary Statistics for Proposed Competence Factor	60
Table 11. Summary Statistics for Proposed Masculinity/Femininity Factor	60
Table 12. Summary Statistics for Proposed Likeability Factor	61
Table 13. Summary Statistics for Proposed Professional Traits Factor	61
Table 14. Eigenvalues for Principal Factor Extraction Method compared to Null Model	63
Table 15. Factor Pattern and Structure Matrices for Four-Factor Solution	65
Table 16. Factor Pattern and Structure Matrices for the Two-Factor Solution	67
Table 17. Factor Names and Indicator Items	70

Table 18. Summary Statistics for Competence and Masculinity/Felinity Factors	70
Table 19. Summary Statistics of the Seven Largest Chisquare Change Statistics	74
Table 20. Likeability Indicator Item Intercorrelations	75
Table 21. Summary Statistics for Leadership Factors	77
Table 22. Descriptive Statistics for Male Research Participants in the College of Arts and Science for Leader Competence	80
Table 23. Descriptive Statistics for Female Research Participants in the College of Arts and Science for Leader Competence	80
Table 24. Descriptive Statistics for Male Research Participants in the College of Education for Leader Competence	81
Table 25. Descriptive Statistics for Female Research Participants in the College of Education for Leader Competence	81
Table 26. Descriptive Statistics for Male Research Participants in the College of Arts and Science for Leader Masculinity	82
Table 27. Descriptive Statistics for Female Research Participants in the College of Arts and Science for Leader Masculinity	82
Table 28. Descriptive Statistics for Male Research Participants in the College of Education for Leader Masculinity	83
Table 29. Descriptive Statistics for Female Research Participants in the College of Education for Leader Masculinity	83
Table 30. Descriptive Statistics for Male Research Participants in the College of Arts and Sciences for Leader Likeability	84
Table 31. Descriptive Statistics for Female Research Participants in the College of Arts and Sciences for Leader Likeability	84
Table 32. Descriptive Statistics for Male Research Participants in the College of Education for Leader Likeability	85
vii	
Table 33. Descriptive Statistics for Female Research Participants in the College of Education for Leader Likeability	85
Table 34. Modified Levene's Test Summary Statistics for the College of Arts and Sciences and Education by Booklet Form and Participant Gender	86

Table 35. ANOVA Source Table for the College of Arts and Science on Competence	88
Table 36. ANOVA Source Table for the College of Arts and Science on Masculinity/Femininity	89
Table 36. ANOVA Source Table for the College of Education on Masculinity/Femininity	90
Table 39. ANOVA Source Table for the College of Arts and Science on Likeability	92
Table 40. ANOVA Source Table for the College of Education on Likeability	93

List of Figures

Figure 1. Competence Ratings By Male and Female Participants for Professors Acting Democratically in a Task Oriented Situation	43
Figure 2. Competence Ratings By Male and Female Participants for Professors Acting Democratically in a Personal-Oriented Democratic Personal	44
Figure 3. Democratic Personal	44
Figure 4. Autocratic Personal	44
Figure 5. Two-factor confirmatory factor analysis model for leadership measures	77
Figure 6. Competence Ratings for the College of Arts and Science: 3-Way Interaction Effects for Prof Leaders Style, Situation, and Rater Gender (Male Participants)	95
Figure 7. Competence Ratings for the College of Arts & Science: 3-Way Interaction Between Prof Leader Style, Situation, and Rater Gender (Females)	96
Figure 8. Competence Ratings for the College of Education: 3-Way Interaction between Prof Gender, Prof Leader Style, and Situation (Ratings for Male Professors)	97
Figure 9. Competence Ratings for the College of Education: 3-Way Interaction between Prof Gender, Prof Leader Style, and Situation (Ratings for Female Professors)	98
Figure 10. Masculinity Ratings for the College of Arts & Science: 2-way Interaction Effect between Prof Leader Style and Participant Gender	100
Figure 11. Likeability Ratings for the College of Arts & Science: 2-way Interaction Effect between Type of Situation and Rater Gender	101
Figure 12. Likeability Ratings for the College of Arts & Science: 2-way Interaction between Prof Leader Style and Rater Gender	103

Figure 13. Likeability Ratings for the College of Arts & Science: 3-way Interaction Effect between Prof Leader Style, Situation, and Rater Gender (Male Raters)	104
Figure 14. Likeability Ratings for the College of Arts & Science: 3-way Interaction Effect between Prof Leader Style, Situation, and Rater Gender (Female Raters)	105
Figure 15. Likeability Ratings for the College of Education: 2-way Interaction Effect between Professor Leader Style and Type of Situation	106

ABSTRACT

This experimental study used eight written vignettes to analyze the effects of professor gender, professor leadership style (democratic/autocratic), and type of situation (task/personal) and participant gender on evaluations of professors' competence, likeability and masculinity characteristics. Undergraduates from the College of Arts and Science (N=932; Males=464, Females=467), and the College of Education (N=722; Males=140, Females=582) were used. Results indicated that research participants rated democratic professors significantly more competent, likeable, and more feminine than autocratic professors. Contrary to expectations derived from gender spill-over and gender congruency theories, male participants did not rate female professors more negatively than their male counterparts when they acted autocratically in a personal situation (i.e., gender incongruent manner.) Exploratory results revealed trends that are discussed along with theoretical and practical implications.

Chapter I

Introduction

An extensive amount of empirical research has examined how men and women in leadership positions are evaluated. The majority of these studies have been conducted within an organizational setting and have evaluated leadership styles along two dimensions: (a) task accomplishment/interpersonal relationships, and (b) autocratic/democratic style. Task accomplishment refers to a leadership style that concentrates on following rules, performing assigned tasks, and making leader and subordinate roles explicit; the interpersonal style is more concerned with employees' morale and their relationships. Even though aspects of the task leadership style are considered to be independent of the interpersonal leadership style, the democratic (participative)/autocratic (directive) leadership style dimension (e.g., Lewin, 1937) is not considered to be independent of the task-person dimension. For example, Eagly and Johnson (1990) suggested that the democratic style describes a more narrow aspect of the interpersonal style, while the autocratic style describes precise behaviors of the more general task oriented style. Western culture generally associates the task and autocratic style with male gender stereotypes (e.g., masculine, dominant, aggressive, ambitious, controlled), while the personal and democratic styles are closely associated with female gender stereotypes (e.g., warm, friendly, submissive, concerned for others). General beliefs about a competent leader are most closely associated with masculine gender stereotypes, while female gender stereotypes are not seen as being essential attributes for a good

leader (e.g., Broverman, Vogel, Broverman, Clarkson & Rosendrantz, 1972). While most researchers agree that gender stereotypes, in part, play a role in female underrepresentation in leadership roles, there are several major competing theories explaining the reasons for this disparity. For example, gender-role spillover theory (e.g., Heilman, Block, Martell, & Simon, 1989) states that gender roles carry over from (Nieva & Gutek, 1980) into organizational settings. According to role congruence theory leaders who behave inconsistently with their gender role expectations are more negatively evaluated than leaders who behave in a manner that is consistent with their gender roles. For this reason, Deaux and Major (1987) proposed that it is important to examine the degree to which a job emphasizes gender typed skills and behaviors.

Empirical evidence of gender bias in leadership evaluation is not consistent. While meta-analytic research found that women consistently were more democratic in their leadership style (Eagly et al., 1990), research evaluating the perception of leadership outcomes most consistently found that women were negatively evaluated when (a) assessed by males, and (b) when supervising in a male dominated environment (e.g., military) (e.g., Eagly., 1987). In addition, studies consistently showed evidence that female leaders who used stereotypically autocratic styles were evaluated most negatively by male raters, while female raters did not show this rating bias (e.g., Eagly et al., 1995).

Eagly, Makhijani and Klonsky (1992), in their meta-analytic review, found a multitude of methods used to measure leadership outcomes. They commented that these methods increased the complexity of outcome comparisons. For example, leader competence was frequently measured using leader productivity, expertise, effort, and effectiveness. Similarly, leader satisfaction was also assessed in terms of leader

likeability, desire to work with the leader, and group cohesiveness. Leadership ability was measured by some studies in terms of ability to organize, plan, and execute managerial duties. In addition to leadership outcome measures, studies frequently included scales to assess leader masculinity/femininity personality traits (e.g., Rojahn & Willemsen, 1994). However, little factor analytic research has examined the measures of leadership. And even though an extensive amount of research has explored the link between leader gender, leadership styles, and leader ratings, only one study attempted to integrate both task/personal and autocratic/democratic leadership dimensions in their examination of leader outcome ratings (Remland, Jacobson, & Jones 1983). This particular study was of interest because the authors integrated both leadership dimensions into their design (along with participant gender, participants' psychological gender, and gender of manager described in the scenario).

Problem Statement and Purpose of the Study

The literature review revealed that many studies have used single items to measure leader attributes, or composites of two or more items. The literature review of leadership studies has further revealed that even though the two leadership dimensions (task/personal, democratic/autocratic) are generally believed to be equally important as factors influencing leader evaluations, the majority of studies have not examined possible interactions between these dimensions (except for Remland et al.'s study). Finally, while most leadership research has been conducted within an organizational setting, leadership studies within academia have been much less frequent, and more importantly dated from the 70's and 80's and in need of updating.

The purpose of this study was two-pronged: (a) develop a psychometrically sound leadership evaluation scale for an academic setting, and (b) create an experimental design that incorporated both leadership style dimensions (task/personal and democratic/autocratic) in order to examine the predictive accuracy of two well established gender role theories (gender role spillover and gender congruency theory) to explain leader evaluations.

An experimental design was created using written vignettes describing a professor/student interaction. The researcher manipulated the following variables: professor gender, professor leadership style (democratic or autocratic), and type of situation (task or personal). This study also included participant gender as a variable, which created a 2 (professor gender) x 2 (professor leader style) x 2 (type of situation) x 2 (participant gender) between factors design. The dependent measures consisted of 24 leadership items representing four leadership factors: competence, likeability, masculinity/femininity, and professional traits. This study tested five hypotheses that were derived from a combination of two gender role theories (gender role spill over and gender role congruency theory) and empirical evidence. Research participants were undergraduate students from the College of Arts and Science and the College of Education from one university.

Hypothesis 1: Democratic professors will be rated as significantly more competent than autocratic professors.

Hypothesis 2: Autocratic professors will be rated as significantly more masculine than democratic professors.

Hypothesis 3: Democratic professors will be rated as significantly more likeable than autocratic professors.

Hypothesis 4: There will be an interaction effect between participant gender, professor gender, professor leadership style, and the type of situation on the ratings of leader competence.

Hypothesis 5: There will be an interaction effect between participant gender, professor gender, professor leadership style, and the type of situation on the ratings of leader likeability.

Hypotheses 4 and 5 anticipated differences between male and female research participants with regard to competence and likeability ratings. First, it was predicted that female research participants would not be influenced by professor gender in their ratings of leader competence and likeability, only by leadership style (see hypotheses 1 and 3). On the other hand, competence and likeability ratings for male research participants were anticipated to be influenced not only by leadership style (hypotheses 1 and 3), but also by professor gender and type of situation. First, it was anticipated that male participants would judge autocratic female professors more harshly than their autocratic male counterparts. In addition, these anticipated differences were predicted to be largest in the personal condition, that is, male participants would rate autocratic female professors in personal situations the least competent and the least likeably compared to all other conditions.

Organization of the Study

The remainder of this study is organized in the following manner. Chapter II reviews the literature on gender differences in leadership evaluations within

organizational and academic settings. Specifically, this review summarizes studies that examined the relationships between gender role theories, leadership dimensions (task/personal, democratic/autocratic), and rater gender with regard to leadership evaluations of male and female leaders.

Chapter III presents the procedures and methods used in this experiment. Sample selection, instrument and leadership factors, and experimental procedures are discussed.

Chapter IV presents the results of factor analytic procedures used to analyze the psychometric properties of the leadership factors (using the whole sample), followed by an in depth presentation of the results that test the five hypotheses. This chapter also presents detailed results of exploratory findings. Hypothesis testing and exploratory findings are presented separately for the College of Arts and Science and the College of Education.

Chapter V discusses the results of the five hypotheses and exploratory findings as they relate to gender role theories and findings of previously published studies. Theoretical and practical implications are evaluated along with limitations concerning population and ecological validity. Finally, suggestions for future research are presented.

Chapter II

Review of the Literature

The review of the literature is divided into four sections. Using an historic perspective, the first section discusses the major theories (e.g., Kanter, 1976) that attempt to explain differential distribution of men and women in positions of authority and leadership. The second section examines empirical evidence of gender differences with regard to leadership behaviors. Specifically this section will establish the link between leadership style and gender stereotypes with regard to verbal and non-verbal behaviors believed to express masculine and feminine leadership styles.

The third section summarizes studies examining the relationship between gender and the perception of leadership effectiveness and leadership ability within non-academic as well as academic environments.

The fourth section discusses methodological issues. Specifically, the development and use of a written vignette/scenario format and other alternative formats are explored along with measurement issues regarding operational definitions that capture the measures of leadership ability and effectiveness.

Section I – Gender Role Models

Based on their research in 1994 on managers and secretaries, Vinnicombe and Colwill found that over 95% of secretaries in the North America and Western Europe were women, whereas over 95% of senior managers were men. In general, women are over-represented in low status jobs and under-represented in high status occupations.

These subspecialties within male oriented professions are usually rated lower in terms of status and prestige by members of that profession. With the exception of nursing, occupations traditionally associated with females tend to have male supervisors in the highest positions.

Most researchers agree that one of the most important reasons that women are not equally represented in organizational leadership positions is that women are socialized into specific gender roles (e.g., Eagly, 1987; Payne, Fuqua, & Canegami, 1998). Payne et al., for example, stated that the socialization process promotes differential gender expectations for men and women in many aspects of social interactions. With regard to the workforce, it is believed that men are expected to exhibit greater qualifications for leadership than do women. Kruse and Wintermantel (1986), in their review of gender stereotyping in leadership positions, summarized what other researchers have stated repeatedly (e.g., Rosenkrantz et al., 1968), which is that stereotypes of how men differ from women match the general perception of how leaders differ from followers. In their review of the gender stereotype research literature, Kruse et al. noted that, in general, men are believed to be aggressive, independent, objective, active, dominant, competitive, and decisive, while women are believed to be gentle, emotional, sensitive, dependent and submissive. These gender beliefs are generally accepted to be normal and healthy by mental health professionals and have been established over many generations (Kruse et al., 1986). As a result of gender typing of leadership positions, Terborg (1977) suggested that women needed to adopt masculine qualities if they wanted to succeed in management.

Extensive research on gender stereotypes in social behaviors (e.g., Broverman et al., 1972) has established that the majority of the beliefs that people hold about the differences between men and women can be described along the personal/task dimension (Eagly, 1987). The personal dimension involves behaviors that are of a caring and nurturing nature. Such behaviors include concern for the welfare of others, loving children, being affectionate, able to devote self completely to others, eager to soothe hurt feelings, helpful, kind, and sympathetic. Additional traits included in the personal dimension describe a person's emotional expressiveness (e.g., expresses tender feelings), interpersonal sensitivity (e.g., aware of feelings of others), and personal style (e.g., social orientation, femininity). It is believed that women exhibit personal-oriented behaviors more strongly than men (e.g., Eagly, 1987).

In contrast, men are more strongly associated with the task dimension. Behaviors consistent with the task style include tendencies to be assertive, controlling, aggressive, ambitious, dominant, forceful, and acting as a leader. Additional attributes include independent, self-reliant, self-sufficient as well as self-confident and feeling superior. A task-oriented person acts in a direct and adventurous manner and does not give up easily. The task/personal dimension has also been labeled communal/agentive, and socially/instrumentality oriented. The current study will refer to this dimension as task/personal.

Physical Characteristics as They Relate to Task/Personal Dimension

Ecological theory has suggested that male and female physical attributes may be directly linked to task and personal qualities. Deaux and Lewis (1984) found empirical evidence that masculine descriptions such as strong and sturdy were associated with the

task style, while feminine descriptors such as graceful and soft were linked to the personal style. Deaux et al. stated that traditionally, women in society have been socialized to be nurturing, likeable, affectionate, soft-spoken, warm, yielding, selfless, gentle, compassionate and dependent, while men have been traditionally socialized to be ambitious, aggressive, dominant, self-reliant, strong, individualistic, and independent. According to Klenke (1996), attributes and behaviors associated with women typically are inconsistent with leadership behaviors.

Schein (1972) was one of the pioneers who explored gender stereotypes in organizational settings by presenting a list of 92 male and female characteristics to a sample of managers. Results of her study suggested that both men and women managers believed that a successful manager possessed a higher degree of masculine traits as compared to feminine traits.

Gutek (1993), in their review of stereotypes of women in management, found that Schein's findings in the 1970s were still valid in the late 1980s. In the late 1980s, women in management were perceived as less aggressive and independent than men in management. However, studies also showed that women were seen as having higher degrees of interpersonal skills (e.g., Brenner, Tomkiewicz, & Schein, 1989; Frank, 1989). Gutek concluded that these results were compatible with the fact that women are socialized to be cooperative while men are socialized to be competitive.

Gender-role spillover. Phillips and Lord (1982) proposed that in organizational settings people develop specific expectations about appropriate leadership behaviors which take precedence over more gender based expectations. Gender–role spillover theory (Gutek & Morasch, 1982) rejects that notion, and instead stipulates that gender

roles continue to carry over into organizational settings. Specifically, Gutek et al. maintain that women and men are not regarded as generic managers, but instead, are evaluated by both their gender and their position within the organization. Gender-role spillover affects women in leadership positions differently than men because of people's expectations about leader appropriate behaviors more closely match prototypical masculine qualities. According to numerous researchers (e.g., Kruse & Wintermantel, 1986; Ragins & Sundstrom, 1989) women who want to be successful in their leadership positions have to adopt male like qualities in their leadership style. This managerial model (Terborg, 1977) however forces women to violate conventions concerning appropriate female behavior, and as a consequence, women are evaluated less favorably than their male counterparts. In addition to being more negatively evaluated, women in leadership roles, who adopt a masculine style, may also be perceived to be more extreme in their behavior than male leaders.

Gender-role congruence. Nieva and Gutek (1981) proposed a more detailed set of predictions in terms of leadership evaluation. Their theory maintains that leaders who behave consistent with their gender role expectations are evaluated more favorably than leaders who behave in a manner inconsistent with their gender role expectation. The gender-role congruency model predicts that women who choose a feminine leadership style (gender congruent) are predicted to be more positively evaluated than women who choose a masculine or counter-stereotypic leadership style. Empirical studies show mixed results. For example, Eagly, Makhijani and Klonski (1992) reported that women who displayed a masculine leadership style were more negatively evaluated than women who displayed a more feminine leadership style. However, Petty and Buning (1980) failed to

find support for the gender congruency hypotheses and found that male and female leaders who showed consideration were evaluated as equally effective. Nieva and Gutek propose that women who lead in a style congruent with their gender role may reduce the role conflict that is experienced when occupying a leadership position, as well as escape the negative evaluations that are experienced when behaving counter to gender-role expectations. Similarly, Eagly and Karau (1991) suggested that gender congruent leadership behaviors may ease the role conflict women experience in managerial positions and lead to a less stressful situation.

Social role theory. Social Role theory as it pertains to leadership positions (e.g., Baron & Kenny, 1986), stated that in order to consistently account for gender differences in leadership, moderator and mediator variables have to be considered. For example, Deaux and Major (1987) suggested that the degree to which a job emphasized gender-differentiated skills and abilities could act as a moderator variable. Specifically, the authors predicted that women may be attracted to a job that requires primarily personal-oriented activities, whereas men may be more attracted to jobs requiring predominantly task-oriented activities. Some support for this notion was reported by Wood (1987) who devised a group of tasks that were either categorized to require complex personal-oriented activity, or that called for predominantly task-oriented activities. These tasks were then assigned to all-male and all-female groups. Results indicated that female groups were superior in their performance of tasks requiring personal-oriented skills while male groups excelled at task-oriented projects. Social role theory predicts that women will experience role conflict when taking on leadership roles because conventions

regarding appropriate female behaviors conflict with the general expectation that leaders behave in a masculine, or task-oriented style (Eagly, 1987).

Section II - Observed Gender Differences in Leadership Styles

Eagly and Blair (1990) noted two approaches to the study of gender differences in leadership positions. The first approach is represented in books and articles written by management experts who base their information on their own experiences in organizations, as well as on personal interviews with men and women employed in leadership roles. The second approach is based on formal inquiries by social scientists. Most experts who base their information on experience maintain that men and women in leadership positions behave in accordance to their gender role. For example, books on the practice of management (Loden, 1985) have argued that women prefer to manage in a style that allows for cooperation and collaboration between managers and subordinates. This style gives less control to the leader, and encourages problem solving based on intuition, empathy, as well as rationality. Sargent (1981) agrees that even though men and women in management behave according to gender stereotypes, he argued that a movement towards androgyny would improve both men and women's performance as leaders.

On the other hand, empirical research data do not show conclusive evidence of that claim. Reviews from empirical research (e.g., Bartol & Martin, 1986) have led to the general consensus that there are few differences between males and females with regard to leadership style. Eagly and Johnson (1990) point out, however, that Bartol and Martin (1986), for example, based their generalizations on a sample of only eight studies. Furthermore, relatively informal methods were used to draw these conclusions

(summarizing research findings). To remedy this approach, Eagly et al. used meta-analytic techniques in their review of the leadership literature in order to come up with a systematic, quantitative integration of available research. The authors found that the majority of the studies examined four aspects of leadership styles. These four aspects were usually organized into the leader style dimensions of task/personal and democratic/autocratic. The task/personal leadership dimension was first coined by Bales in 1950 and further developed by the Ohio State studies on leadership (e.g., Halpin & Winer, 1957). These studies maintained that the task dimension included behaviors such as maintaining high standards for performance, and making leader and subordinate roles explicit, while the personal leadership dimension included behaviors such as being friendly and helpful to subordinates. The democratic/autocratic leadership dimension has its roots in experimental leadership studies (e.g., Lewin & Lippitt, 1938), and is generally considered to define a narrower aspect of the task/personal dimension (Bass, 1981). Eagly et al. along with other researchers state that the democratic/autocratic dimension relates to gender stereotypes in the same manner as the task/personal dimension, that is, men are believed to be more dominant and controlling (i.e., more autocratic) than women. Other variables coded for each study were publication format, possible confounding of gender with age, education, and management seniority. Finally, three types of study settings were examined in the meta-analysis: (a) organizational studies within educational, business and government settings; (b) assessment studies using college undergraduate and graduate students as research participants; and (c) experimental studies using college undergraduate and graduate students.

Eagly et al. examined all instances in the literature that pertained to the assessment of the task/personal dimension, as well as the democratic/autocratic dimension. Measuring these dimensions included the use of direct observation (coding leadership behaviors as they occurred), as well as assessing leadership styles indirectly using measures of attitude or personality questionnaires. Other important criteria called for the samples to be drawn from a typical population of the United States or Canada (i.e., adults). Finally, meta-analytic research called for sufficient information in order to calculate effect size estimates for gender with regard to the leadership dimensions. Using these criteria, Eagly et al.'s meta-analytic review included 162 studies. All studies, published and unpublished, were conducted between 1961 and 1987 with a median date of 1981.

Eagly et al. predicted that gender differences in leadership styles would be less pronounced in studies using organizational settings compared to experimental environments. Beyond this prediction the authors maintained that the purpose of their inquiry was primarily descriptive and exploratory.

Eagly et al. found that in general, leadership styles as computed by weighted means across all types of style were statistically significantly stereotypic, however the effect sizes of those results were very small. When computing the means for type of leadership style separately, no statistically significant gender difference was found for the task dimension as well as for the personal dimension. Calculated means revealed a significant gender difference when comparing men and women on the personal leadership dimension such that women tended to use the personal leadership style more often than their male counterpart. The calculated effect size for this difference was very

small ($d=-0.03$) compared to the average effect size range of 0.00 to 1.2 within the area of social and personality psychology research (Eagly, 1987). Gender differences were also found along the democratic/autocratic dimension revealing that women tended to act in a more democratic style. The calculated effect size for this estimate after outlier removal was small ($d= 0.27$).

Eagly's hypothesis that gender differences with regard to leadership style are less pronounced in organizational settings than in experimental and assessment studies was supported in the meta-analytic findings. Specifically, with regard to the task style, gender comparisons in organizational studies were statistically significantly less stereotypic than results obtained in assessment and laboratory studies. However, when comparing the personal and task leadership styles to the democratic and autocratic leadership styles, study settings had no impact on the outcomes.

Based on the results of their meta-analytic findings, Eagly et al. concluded that gender differences in leadership behaviors follow a more complex pattern than previously suggested. Results suggested that, consistent with findings from research on gender differences in social behaviors (e.g., Eagly, 1987; Hall, 1984), women's leadership styles were more democratic than men's regardless of research setting which may reflect underlying gender differences in either social skills, or may indicate subtle status differences men and women occupy within the same organizational role (Eagly et al., 1990). On the other hand, within organizational settings men and women did not differ with regard to task leadership styles, which may indicate that within organizational hierarchies, leadership roles take precedence over gender roles.

*Section III - Perception of Gender Differences in Leadership Positions
and Leader Evaluation*

Many researchers speculate that the lack of representation of women in higher levels of leadership in organizations may be because women's credentials and performance are not fairly evaluated (e.g., Nieva & Gutek, 1981). Fiske, Bersoff, Borgida, Deaux, and Heilman (1991) argued that gender discrimination of women in leadership roles is due to the influence of gender stereotypes on leadership performance. In an effort to examine gender discrimination, Eagly, Makhijani and Klonsky used meta-analytic procedures to evaluate the issue of whether women are more negatively evaluated than their male counterparts. Their meta-analytic research was confined to experimental research that controlled leadership characteristics and behaviors while manipulating leader gender. Eagly et al. reported that experimental studies reviewed, either presented their information in a written vignette format, or trained confederates to act out roles. Eagly et al.'s predictions for the meta-analysis were based on gender-role theory (Eagly, 1987), that is, people develop sets of expectations of what is considered appropriate behavior for roles within organizational settings (i.e., leader or subordinate). However, this theory maintains that gender based expectations have "carryover" effects similar to the gender-role spillover theory (Gutek & Morasch, 1982) in that people's expectations of leaders are influenced by their expectations of gender appropriate behaviors. Based on this theoretical framework, Eagly et al. predicted that woman leaders would be evaluated more negatively than their male leaders. Based on this framework the authors also predicted that the negative evaluation of female leaders would be most pronounced when female leaders use an autocratic and directive style. Finally, the authors

predicted that male research participants would be more prone to evaluate female more negatively than their male counterparts because this pattern was evident in numerous studies.

This meta-analysis included 61 experimental studies with a median publication date of 1980. The majority of research participants in the included studies were college undergraduates, and business management and other graduate students. Most of the studies portrayed the organizational context using examples of business, manufacturing, educational, or small groups vignettes. Leadership styles portrayed in the reviewed studies included examples of the task/personal dimension and democratic/autocratic leadership dimension. The majority of outcome measures of those studies consisted of rating scales for leader competence, leadership style, as well as satisfaction with leader.

Results indicated a significant gender effect when comparing leader competence and leader satisfaction with perceptions of leadership style. Specifically, Eagly et al. found that when measuring leader satisfaction and leader competence, research participants showed a greater tendency to evaluate female leaders more negatively compared to male leaders. As predicted, male research participants showed a greater preference for male leaders than female research participants. Inconsistencies emerged because results indicated that female leaders were perceived to be significantly more task style oriented than male leaders. A priori contrasts indicated that men who displayed an autocratic style were perceived more favorably than task oriented women, providing support for the gender-role congruency theory that women are negatively evaluated when they display a male-oriented leadership style. The fact that women tended to be perceived more negatively when displaying an autocratic or task type leadership style fits with the

notion that a directive and autocratic style most contradicts the traditional female patterns of accepted behaviors. Comparing men and women along the masculine/feminine dimension produced a larger difference between men and women with regard to the masculine dimension compared to the feminine dimension. Men leading in a feminine manner were not perceived negatively relative to women, however masculine women were more negatively evaluated than masculine men. Finally, studies showed a trend that male leaders were preferred over female leaders most often when men were a majority within that environment. This trend did not emerge when there was equal male and female representation by both genders. This result is consistent with gender-role spillover theory that states that women who violate gender-role expectations are more negatively evaluated. Inconsistent findings however indicate some inadequacies of gender theories. For example, results of this meta-analysis indicated that the negative evaluation of female leaders occurred not when directing male subordinates but when directing female subordinates. Male and female research participants favored female leaders with male subordinates and preferred male leaders with female subordinates.

Eagly et al. concluded that in general men were evaluated more favorably than women. However, the calculated effect size estimate for this trend was weak (0.07). More specifically, the effect size estimate for gender differences in leader satisfaction was 0.10, while gender differences with regard to leadership competence produced an estimated effect size of 0.09.

Gender and Leader Competence.

Even though a number of social and organizational psychologists (e.g., Denmark, 1993) have argued that women are not less effective than men in leadership roles, some

authors (e.g., Hollander, 1993) have stated that there may be some situations when female leaders appear to be more effective, while male leaders appear to be more effective in other situations. Bass (1990) argued that male leaders are more favorably evaluated with regard to competence because of existing gender stereotypes. Hunt (1991) defined leader competence in terms of the ability of a leader to facilitate group or organizational goals. There is some degree of consensus among organizational psychologists (e.g., Hunt, 1991) that competence should be regarded as an outcome of a leader's ability. In order to assess the level of leader competence numerous criteria have been suggested. For example, empirical studies have assessed the level of competence by measuring: (a) group or organizational productivity, (b) subordinate leader satisfaction, and (c) leadership performance as well as reputational ratings using the perspectives of superiors, subordinates, peers, and leaders themselves (Eagly et al., 1995).

In their review of the literature on leadership competence, Eagly, Karau and Makhijani (1995) found that the majority of studies were conducted within an organizational setting examining male and female managers in comparable leadership positions. A smaller body of literature investigated leadership competence using experimental designs with college students as research participants. The purpose of Eagly et al.'s meta-analysis was to provide a systematic and quantitative synthesis of existing research in order to address the general question of whether male and female leaders differed in their degree of leadership competence, as well as to evaluate possible moderating conditions that could influence competence ratings for men and women.

In order to predict gender differences in leadership competence the authors used social role theory (Eagly, 1987). That theory states that women experience increased

degrees of conflict in their leadership roles because existing social pressure to behave consistent with the female gender role is inconsistent with the characteristics of the masculine behaviors the leadership role traditionally demands. The more leadership roles are male dominated, the higher the degree of role conflict in women. According to this theory, female managers are predicted to be viewed more competent if they adopt a feminine leadership style, than women who adopt a masculine leadership style. Support for social role theory is evident in the results of Eagly and Johnson's meta-analysis (1990), which showed evidence that female leaders led in a more democratic and participatory leadership style than their male counterpart. Meta-analytic findings provided evidence that women were most negatively evaluated when they occupied male dominated positions, or when women were evaluated by men (Eagly, Makhijani, & Klonsky, 1992). The authors included 96 studies with a median publication date of 1980 (range: 1962-1992) in their meta-analysis of which 70% were organizational studies. Careful consideration was given to the classification of measures of competence. Competence measures were defined as objective or subjective. Objective measures included information about the manager's expert knowledge in his/her field, as well as measures of production goals met by subordinates. Subjective competence measures included rating scales assessing leader performance, ability, effectiveness, satisfaction with leader, and effort and motivation. In order to examine the interaction between the perceived degree of gender-stereotypic nature of leadership roles with perceived leadership effectiveness of men and women, the authors conducted a separate study to measure the degree of perceived gender stereotypic content of leadership roles. The resulting match between gender-roles and leadership roles was then defined along the

dimension of gender congeniality. Specifically, leadership roles that were judged to require female stereotypic qualities (e.g., interpersonally oriented qualities of the role) were defined as female-congenial while the male-congenial aspects of leadership roles had been judged to require masculine qualities (e.g., task-oriented qualities).

Results of this meta-analysis indicated that in general women and men did not differ in their leadership effectiveness. No gender differences were found when comparing objective performance measures with subjective measures, or when comparing results from organizational studies with experimental studies. However, studies conducted within the military organization produced statistically significant outliers, that is, military studies deviated strongly from all of the studies in the meta-analysis. Within the military organization men fared statistically significantly better with regard to the competence measure. The mean weighted effect size estimate was 0.42.

A statistically significant model was produced when leadership type was examined according to the categories of first-level (line-level) and second-level (middle-level) management. Specifically, men were perceived to be more competent when occupying line-level leadership roles (effect size estimate: 0.19), while women were perceived to be more competent when managing at the middle-level (effect size estimate: 0.18). When considering the gender-congeniality dimension, results of this meta-analysis showed a relationship between the perceived degree of male-stereotypic content of a leadership role and the perceived leadership competence rating. Thus, women who occupied increasingly male-congenial leadership roles were perceived to be less competent than their male counterparts. Analogously, women were judged to be

more competent than their male counterpart when the leadership role was perceived to require female stereotypic qualities.

In their conclusion, Eagly et al. underscored that overall women did not differ from men in terms of competence evaluations. However, the authors also point out that even though there were no differences found overall, men were judged to be more competent leaders in leadership roles considered to possess a high degree of masculine qualities, while female leaders tended to be judged as more competent in leadership roles considered to possess a high degree of feminine qualities. Results of the military setting provided evidence for the gender-role spillover model that states that gender-based expectations carry over into the work place (Nieva & Gutek, 1981). Eagly et al. argued that men were rated to be more competent in line management while women were preferred in middle management. These results fit the gender-congeniality dimension when examining the skills involved for line and middle management. Thus, line-level management can be described to be male congenial because it requires the involvement of a relatively high degree of technical skills, while middle management fits the dimension of female congeniality because of the emphasis on human relations skills.

Gender and Leader Likeability

Rojahn and Willemsen (1994) examined the sex-role congruency theory in the context of leadership. In their pilot study they found that male participants were more likely to stereotype behaviors than female participants. Thus, the authors hypothesized that gender-role appropriate behaviors would be evaluated favorably by men. The experimental scenario consisted of two male and two female students assigned to do a team class project. Via a written description, the authors created the positions of team

leader for each female and male member of the group. They also manipulated the scenario to include descriptors of task and personal leadership style behaviors. Four versions were created: two conditions representing a male task leader (gender appropriate) and two conditions representing a female task leader (gender inappropriate). Similarly, two personal male conditions (gender inappropriate), and two personal female conditions (gender appropriate) were created. Examples of the task condition described the leader to be pragmatic, suggesting that differences be settled by drawing lots, and a suggestion by the leader to encourage the weakest member (always opposite gender to leader) to drop out. The personal condition described the leader to be sensitive to the conflicts among group members, making considerable efforts to maintain good relations with members and offer to help the weak member of the group in order to avoid having the person drop out.

Participants included 342 female and 154 male Dutch undergraduate psychology students who were randomly assigned to one of four experimental conditions. After reading the scenario, participants filled out questionnaires concerning leader competence, leader likeability, and leader personality traits. Three questions were asked for leader competence (1. Will the task be finished on time? 2. Will the manager maintain group morale? 3. How does the manager contribute to the goal?), and leader likeability (1. How much would you like to work for the manager? 2. How comfortable would you feel working for him/her? 3. How likeable do you find the manager?). Leader personality traits consisted of 29 items describing stereotypical male and female traits. Male adjectives included active, ambitious, cool, dominant, independent, influential, persistent, taking charge, determined, forceful and self-confident, balanced, convincing, demanding,

efficient, and full of initiative. Female adjectives included appreciating, considerate, cooperative, helpful, social, sympathetic, tolerant, understanding, warm, reliable and talkative. All dependent variables were arranged on a scale ranging from 1 (not at all) to 100 (extremely). The between subjects factors included leader gender, leadership style, and participant gender. Results revealed that men judged leader competence differently depending upon leader gender. Specifically, female task leaders were judged to be less competent than their male counterparts. Similarly, female personal leaders were judged to be more competent by male participants than her male counterpart. The fact that male research participants rated both male and female gender role inappropriate leaders more negatively with regard to leader competence was unexpected. Overall, results for leader competence supported the gender-role congruency theory for male participants only. Female research participants were not influenced by leader gender in their evaluation of leadership measures. Other findings showed that overall, personal style leaders were judged to be as competent as the task style leaders. Personal leaders were rated to be more likable and more competent with regard to maintaining group morale. Estimated effect sizes for leader competence and likeability were comparable to the calculated effect size estimates of meta-analytic research (e.g., Eagly, Makhijani, & Klonsky, 1992; Swim, Borgida, Maruyama, & Myers, 1989).

Many organizations provide women with leadership training in an effort to overcome the status disadvantage women face when taking a leadership position (e.g., Brown, Dovidio & Ellyson, 1990). However, some researchers claim that these strategies can potentially backfire because they may cause men to like women less as leaders, and thus may be less likely influenced by women who underwent leadership training. For

example, Carli (1990) found that when women were perceived to be likeable they were also perceived to exert more influence over men than when they were perceived to be less likeable. Carli, LaFleur, and Loeber (1995) examined different strategies that optimized the relationship between perceived likeability and influence given a male or female audience. Specifically, the authors examined whether the dominant nonverbal style was a more effective strategy for men than women when addressing a male audience. Further, the authors examined whether the combination of projecting warmth and competence produced better results especially for a male audience. Carli et al. also evaluated possible interaction effects between gender, influence, perceived competence, and likeability. Based on the premise that women in the Western culture are more often associated with a lower social status than men (e.g., Berger, Rosenholtz, & Zelditch, 1980), Carli et al. hypothesized that male research participants would perceive a dominant female leader to be more threatening and less likeable compared to her male counterpart. Based on previous research the authors predicted that the degree of influence for women would depend on whether she was perceived to be likeable, while they predicted that the perception of likeability of male leaders would not impact the perception of influence. A preliminary study (95 male and 114 female undergraduates) was conducted to determine a gender-neutral topic that was interesting to both men and women. As a result, the topic of the university's current student meal plan was selected. The authors composed an argument in favor of the current meal plan and trained four (two male and two female) student confederates as a part of a dyad member. Confederates were trained to deliver a positive argument in favor of the student meal plan using either a dominant, submissive, task, or social style. The dominant style was characterized by speaking in a loud voice,

pointing at the other person in an intrusive manner, staring at the other person while talking, and maintaining stern facial expression. The submissive style was characterized as speaking in a soft pleading tone, using many hesitation and stumbles in the speech, slumped body posture, nervous hand gestures and avoiding eye contact. The social style was characterized as speaking in a moderately loud voice, having relaxed body posture with a body leaning towards the listener, friendly facial expression, and moderately high amount of eye contact. The task style was defined as speaking rapidly in a firm tone of voice, using few hesitations in the speech, having an upright body posture, using calm hand gestures, and displaying a moderately high amount of eye contact. The other student of the confederate dyad posed as a student listening to the speaker. For the study, 80 male and 80 female students from introductory psychology classes were recruited. Students were informed that the purpose of the study was to examine first impressions and subsequent group interactions. Thus, before meeting their selected partner for the subsequent campus issue debate, students were asked to watch a video of that student member and to provide information about their impression of the student they watched in the video. Using a nine point Likert-type scale, research participants rated the following statements: (a) how well they could work with the speaker and (b) the extent to which the speaker was likeable, trustworthy, competent, persuasive, powerful, knowledgeable, confident, condescending, influential, anxious, intelligent, intimidating, threatening, group-oriented, friendly, and believable. Research participants were also asked to indicate whether the person listening to the speaker on the tape was male or female. This question was asked to determine whether participants held stereotypes about nonverbal styles in conjunction with listeners. A factor analysis on the adjectives extracted five

factors: likeability ($\alpha=.81$), competence ($\alpha=.78$), power ($\alpha=.88$), threat ($\alpha=.85$) and anxiety ($\alpha=.72$). A factorial ANOVA using gender of participant, gender of speaker and nonverbal style was used to analyze the data. The influence of the speaker with regard to participants' opinion change was measured by how well participants agreed with the speaker's message. Results revealed that speakers using the task or social nonverbal style were more influential than those using the submissive or dominant styles in changing participant's attitudes. Male participants rated female task speakers to be less influential than their male counterparts. Regression analysis used opinion change as the dependent variable and the estimated factor scores as predictor variables. Results showed that higher levels of competence and likeability were associated with greater influence, while higher levels of threat were associated with reduced influence. Male participants did not weigh competence differentially with regard to male and female speakers. However, likeability seemed to be more important for female speakers in terms of influencing a male audience. Female participants were not influenced by the likeability variable. Other results indicated that social speakers were liked more than task speakers, who were liked more than submissive speakers. Dominant speakers were liked the least. Dominant speakers were rated to be more threatening and higher in power than task and social speakers. Male participants liked male task speakers better than female task speakers. However, male participants did not differ in their likeability ratings of dominant male and female speakers. Finally, contrary to predictions, men did not perceive dominant female speakers to be more threatening than dominant male speakers. Interestingly, female participants rated male task speakers to be more threatening than female task speakers, but did not differentiate between dominant male and female speakers. Also, a statistically

significant association was found between the assigned gender of listener and speaker style. Research participants were more likely to believe that submissive and dominant speakers were addressing men, and that social and task speakers were addressing women. The authors suggested that persons are generally threatened by competent opposite-sex speakers more so than by competent same-sex speakers. And even though likeability and competence were important predictors of influence, likeability was a more important predictor of influence with a male audience when dealing with female speakers. Results demonstrated that both male and female speakers benefited from the use of a social oriented style.

Interaction Between Rater Gender and Leader Gender

Evidence in meta-analytic reviews of leadership evaluation indicate that male research participants have a greater tendency to evaluate female leaders more negatively than female research participants (e.g., Eagly et al., 1992). However, a recent study by Luthar (1996) found evidence contradicting these previous research findings. The author's experiment examined the relationship between research participant gender, leader gender, and leader style (democratic or autocratic) on the evaluation of leadership ability and leadership performance. Based on published research, Luthar hypothesized that male research participants would tend to evaluate male leaders more favorably than female leaders, while female research participants were predicted not to show any gender bias in their evaluation of leadership ability and performance. In his study, Luthar randomly assigned 130 female and 160 male undergraduate business students to the following vignette conditions: a democratic or autocratic male or female general manager described as facing the problem to increase the number of membership holders in a health

club. After reading one version of the vignette, research participants were asked to rate, using nine point Likert type rating scales, management style and leadership ability. Consistent with previously published studies, Luthar found that democratic managers were rated higher on leadership ability and performance than their autocratic counterparts. However, contrary to previous empirical evidence, this study revealed a rater gender bias for both men and women. Specifically, the interaction effect revealed an opposite rater gender trend, that is, male research participants rated male managers more positively, while female research participants rated female managers more favorably. Male participants did not perceive a difference between male and female autocratic managers on leadership ability, while women rated autocratic male managers significantly lower compared to autocratic female managers on the leadership ability scale.

Gender and Leadership Issues in Academia

This group of studies examined perceptions of performance of teaching faculty within the academic environment. Studies examined the effects of instructor gender and teaching style on performance ratings. Other studies used experimental designs to manipulate teacher behaviors such as friendliness and smiling to examine their impact on teacher performance.

After reviewing results of studies examining student evaluation of teacher performance, Unger (1979) concluded that there was no correlation between student achievement and evaluation of instructors. Instead, ratings appeared to be a reflection of student affective reactions to personal characteristics of instructors. For example, Wittrock and Lumsdaine (1977) suggested the non-task related instructor characteristics

of warmth, popularity, reputation and enthusiasm had a significant impact on student ratings of teacher performance. Wittrock et al. suggested that a student's own set of values and frame of reference impacted the way instructors were evaluated. Research regarding gender-role stereotypes and evaluation bias (e.g., Broverman, Vogel, Broverman, Clarkson, & Rosenkrantz, 1972) provided evidence that instructor gender influenced student evaluations of teacher performance. Harris (1976) found that instructors with a feminine style were perceived to be less competent than instructors with a masculine style. Unger (1979) examined the effect of instructor gender on responses to course evaluation questionnaires. The author recruited 40 members of a psychology department of an urban northeastern liberal arts college. After elimination of minority faculty, the sample included 12 women and 26 men ranking from instructor to full professor. Evaluations were administered during the last few weeks of the semester for each class that instructors taught. For each instructor averages were calculated for teaching effectiveness and perceived grading difficulty. Ranks for grading difficulty were correlated with ranked items for teaching performance. Results revealed that while no correlation was found between teaching effectiveness and difficulty of grading among male instructors, a negative correlation between the two variables was established for female instructors. Women who were perceived as difficult graders received lower teacher effectiveness scores than women instructors who were perceived as easy graders. Unger concluded that more demanding female professors were perceived as acting in a gender inappropriate manner, and thus were more negatively evaluated than their gender appropriate female counterparts.

Similarly, a study by Martin (1984) revealed that students exhibited ambivalent expectations about female faculty. The highest rated female instructor in Martin's research had a combination of feminine qualities such as friendliness, supportiveness and warmth as well as the traditionally masculine quality of preparedness. Kierstead, D'Agostino, and Dill (1988) explored how traditionally female qualities such as warmth and friendliness influenced students in their rating of the instructor. In their first study, the authors constructed a course vignette that described behaviors of friendliness such as out-of class socializing (i.e., having lunch with students from the class). While the educational characteristics (such as type of course, office hours kept and classroom behavior) of the vignette remained constant, the gender of the instructor was manipulated along with a set of friendly behaviors. A sample of 20 male and 20 female college students were asked to read through one version of the scenario and to evaluate the instructor. Results revealed an instructor gender main effect, that is, male instructors were perceived more favorably than female instructors. In their second study, Kierstead et al. manipulated smiling behavior. They created a slide tape presentation of a lecture on the anatomy of the eye. One male and one female model posing as instructors were trained for the smiling or non-smiling condition. While the lecture was held constant, the smiling condition was manipulated during the presentation of slides. After 20 male and 20 female research participants watched the presentation, they were asked to rate the performance of the instructor, and were asked to indicate whether they would take a course with this instructor. Results revealed an interaction between instructor gender and smiling; whereas smiling did not affect ratings for the male instructor, a smiling female instructor received statistically significantly more favorable ratings than the non-smiling female

counterpart. Female instructors were perceived more negatively when they deviated from stereotypical expectation (gender role congruency theory), whereas smiling behaviors did not impact ratings of male instructors.

Section IV- Conceptual Issues

Leadership Dimensions

The review of the literature on the differential impact of leadership style on the evaluation of leadership performance of male and female leaders revealed that most studies examined leadership styles along two dimensions: (a) task/personal, and (b) democratic/autocratic (e.g., Eagly et al., 1992). Even though these dimensions are usually treated as independent of each other, Eagly et al. noted that the democratic/autocratic leadership style described a more precise aspect of the more general task/personal dimension. Luthar (1996) pointed out that despite different terminologies used in published leadership studies, the concept of the democratic/autocratic leadership style was virtually contained in all of them. Luthar pointed out that most leadership classifications used the democratic/autocratic leadership dimension either directly or indirectly. Many management consultants do not favor the autocratic leadership style (e.g., Naisbitt, 1982) because of its potentially demoralizing effect on employees. Even so, many experts in the field of leadership believe that the effectiveness of a leadership style may be contingent upon situational cues (e.g., Wood, 1987). For example, House (1971) suggested that the task style could be more effective when tasks are not well defined. A task style thus could overcome ambiguity by imposing structure. A personal style, on the other hand, could be more beneficial when work conditions are routine (Drenth & Koopman, 1984). However, Eagly et al. in their 1995 meta-analytic review on

leadership effectiveness, pointed out that the empirical literature rarely investigated possible interaction effects between the task/personal and the democratic/autocratic leadership styles on leadership evaluations. One study conducted in 1983 by Remland, Jacobson and Jones manipulated leader gender, leadership style, and type of situation the exchange occurred to measure possible interactions among those variables. Specifically, Remland et al. examined the effect of research participants' psychological gender (measured by Spence and Helmreich's Personal Attributes Questionnaire) on the performance evaluation of nontraditional gender-role behaviors of male and female managers. The authors hypothesized that psychological gender (masculine, feminine, and androgynous) of male and female research participants would interact with leader gender, leadership style, and type of situation on leadership measures. Research participants were 139 male and 150 female undergraduate students enrolled in a communication class. Prior to the experiment, participants filled out Spence and Helmreich's (1974) 24-item psychological gender orientation questionnaire. After participants were categorized according to their psychological gender (gender typed masculine or feminine, androgynous, and undifferentiated) they were randomly assigned to one of six experimental conditions. The authors described the six conditions as versions of a written vignette in which a manager (male or female) interacted with a male subordinate about a task problem (how to improve productivity output), followed by a personal problem (how the subordinate could deal with a family conflict). The authors created a total of six experimental conditions. In versions one and two, a task vignette was presented followed by a personal vignette, with a supportive acting male (female) manager in both scenarios (the authors called these versions SS). Versions three and four described a non-supportive

male (female) manager in the task-oriented vignette, and a supportive male (female) manager in the personal problem (NS for not supportive/supportive versions). Versions five and six described a non-supportive male (female) manager in both the task and the personal vignettes (NN for non-supportive versions in both scenarios). The order of vignette placement in all six experimental conditions remained constant, with the task vignette presented first, followed by the personal vignette. All experimental conditions held subordinate gender within the dyad constant (male). This experiment did not include an SN version to test a supportive manager (male or female) for the task problem, or a version portraying a non-supportive manager (male or female) for the personal vignette. The nonverbal behaviors depicting a supportive or non-supportive manager for the vignettes consisted of nonverbal descriptions. The supportive manager was described as leaning forward, touching the subordinate, speaking in a soft voice, smiling, nodding, and maintaining high eye contact, while the non-supportive manager was described as leaning back, keeping physical distance, speaking in a firm voice, having a serious facial expression, interrupting, no eye contact, and turning away from the subordinate. Remland et al.'s nonverbal descriptions of managers within dyads coincide with behaviors identified in the literature as democratic and autocratic (e.g., Eagly et al., 1990). The experimental design was a 2 (manager gender) x 3 (nonverbal behavior pattern: SS, NS, NN) x 3 (participant psychological gender: androgynous, gender typed masculine or feminine, and undifferentiated) x 2 (participant gender) between factors design. Research participants, after having been categorized according to their psychological gender, were randomly assigned to one of the six conditions. Manager performance in the vignettes was measured using 18 items that were rated on a seven-point bipolar scale. These items

captured two factors: task behavior and consideration of manager performance. It is not clear what type of rotation was used (orthogonal or oblique) with the factor analysis, or which items were included for each scale in the final factor score estimates. Examples for the consideration factor as best discerned included the following items: Lowering employee's confidence/raising employee's confidence, concerned with employee's approval/ not concerned with employee's approval, degrading/upgrading, boosting employee's ego/damaging employee's ego, treating like an inferior/treating like a superior, poor/excellent, and appropriate/inappropriate. Items for the task factor included: clear/unclear, inefficient/efficient, organized/unorganized, and concerned with detail/unconcerned with detail. The internal consistency estimate for the consideration factor as assessed by Cronbach alpha was .91 and .75 for the task factor. Remland et al. hypothesized that psychological gender of the research participants would interact with the ratings of non-traditional manager performance. Results, however, did not support this prediction. Instead, results of this study revealed that managers were rated most considerate in the SS condition followed by the NS version. Managers in the NN condition were rated as the least considerate. Surprisingly, no statistical difference in terms of consideration was found between manager dyads in the SS and NS conditions. Manager dyads in the SS conditions were rated to be more task-oriented than manager dyads in the NN conditions. Again, no statistical difference was found in terms of the perception of task style between the SS and the NS conditions. Finally, male research participants rated manager dyads as more considerate than female research participants.

Although this study had several strong points, such as an experimental design, as well as a combination of leadership behaviors with task and personal situations, this study

had several weaknesses. First, it is not clear why the authors combined the task-oriented vignettes with the person-oriented vignettes, thus reducing the conditions to SS, NS, and NN. Furthermore, after having combined the task vignette with the person vignette, the authors did not add a condition needed to test for a supportive manager in the task condition, and a non-supportive manager in the personal condition (i.e., the SN condition). The order effect problem may be very prominent in this study because the authors always presented the task vignette before the personal vignette. Another potential problem was found with the vignette manager-subordinate dyad. Specifically, the authors failed to vary subordinate gender within that dyad.

With regard to the 18-item scale measuring the dimensions of consideration and task behavior there were several weaknesses. First, most published leadership research studies have measured leadership performance in terms of leadership productivity, effectiveness, likeability, competence, and influence (e.g., Eagly et al., 1992). The measures of this study were inconsistent with current measures and theories of leadership performance. It is unclear why items such as effective/ineffective and inefficient/efficient were not used in the analysis of leadership effectiveness. The relatively weak internal consistency estimate for the task factor (.75) may be an indication that the items used do not form a strong cluster. Even though it is not of interest for the present study, there is a question regarding how research participants were distributed into the categories of psychological gender. And finally, this study was conducted in the early 80's and may not be relevant in today's society, because as Gergen (1986) has pointed out, research results in social psychology are in part a reflection of societal values at a given point in history.

Summary and Purpose of Current Study

A large number of studies on gender differences in leadership behaviors and leadership evaluation have been summarized in the literature using meta-analytic techniques (e.g., Eagly et al., 1990). Meta-analyses have evaluated three important questions: 1) Is there a statistical difference with regard to the evaluation of leadership behavior in men and women, 2) do these differences vary as a function of various moderator variables that include types of situations such as social settings (e.g., Baron & Kenny, 1987), and 3) how important are those differences estimated via the calculation of effect sizes (e.g., Hedges & Olkin, 1985). Results of meta-analyses (Eagly et al., 1990) have found that women in management positions tend to behave more democratically compared to men in those positions. Meta-analytic results also indicated that female leaders were more negatively evaluated than their male counterparts, especially in highly masculine typed leadership roles such as the military (Eagly et al., 1995). In general, however, calculated effect size estimates with regard to gender differences in leadership behavior, and leadership evaluation were consistently small. Eagly and Wood (1991) pointed out that these effect sizes were comparable to those usually obtained in social behavior research. They argued that dismissing these findings of gender differences as trivial could be misleading. Similarly, Eagly, Karau, and Makhijani (1995) have pointed out that even though there were no gender differences with regard to the evaluation of leadership competence, more in depth analysis revealed that gender did matter because gender appropriate leaders were rated to be more competent than leaders with gender inappropriate behaviors (these findings were not consistent, however). Small effect sizes may indicate a change in attitude towards women in leadership roles, due to the fact that

more and more women enter all levels of managerial positions. Nevertheless, Eagly et al. cautioned not to ignore gender as a relevant issue since women still fare much worse than men when they occupy increasingly masculine typed leadership positions.

Deaux and Major (1987) speculated that gender differences in social behavior are context dependent. Specifically, the authors argued that the context in which social interaction occurs influences the resultant behavior. In the review of the literature, gender differences in leadership behavior and leadership evaluation were examined most commonly along the task/personal dimension without integrating the context in which those behaviors were occurring. One experiment (Remland et al., 1983) manipulated the context as well as gender and leader behavior in their evaluation of gender differences in leadership evaluation. Even though the study's methodological weaknesses may have limited its results, Remland et al. addressed a number of important questions related to the evaluation of male and female managers leading democratically or autocratically in a task or personal environment. By improving upon Remland et al.'s study design, the present study aims to contribute to the empirical and theoretical literature on the perception of leader gender incorporating the task/personal as well as the democratic/autocratic leadership dimensions. Furthermore, since studies have consistently found that male participants perceived female leaders more negatively than female participants, it was important to include rater gender into the design. This study also attempted to create a set of psychometrically sound leadership outcome measures, because most leadership studies measured leadership outcomes using single items, or, if more than one item was used, they usually did not undergo any psychometric evaluation. This study psychometrically evaluated a set of leadership outcomes to come up with a set

of psychometrically sound factors. With regard to the study setting, most research on gender differences in leadership evaluations were conducted in non academic organizational settings. This study examined leadership evaluation within an academic setting, because most research of gender differences in teacher evaluation is dated (mid 70's to 80's). By operationally defining leadership roles along the leader dimensions of task/personal and democratic/autocratic, using a written vignette format, this study maximized the manipulation of important leadership variables in order to examine hypothesized complex interaction effects, that otherwise could not have been possible.

Chapter III

Method

This chapter is divided into three sections. The first section presents an overview of the study with the research hypotheses. Section II describes the research participants and the rationale for their selection. Section III describes the process used in the instrument development, along with the development of the outcome measures.

Section I - Overview

The purpose of this experiment was to examine the effect of professor (leader) gender, professor nonverbal behavior (democratic/autocratic), and type of situation (task/person), on the evaluation of leader performance by male and female research participants. Written scenarios of a professor-student interaction were created in this experimental design, to manipulate professor gender, professor nonverbal behavior (democratic/autocratic), and the situation in which the interaction occurred (task versus person-oriented). Leadership evaluations (i.e., leader competence, masculinity/femininity, leader likeability) were provided by male and female research participants. The design of this experiment was a 2 (professor gender) x 2 (professor leadership style) x 2 (type of situation) x 2 (participant gender) between factors design.

Hypotheses

The following hypotheses were based both on theory and empirical evidence from previously published research.

Main Effects

Dimension of leader competence. It was predicted that male and female undergraduate student research participants would rate democratic professors (male or female) higher on the competence scale (Table) than autocratic professors regardless of the situation (task or personal). See Figures 1- 4 .

Dimension of leader masculinity/femininity. It was predicted that male and female research participants would rate autocratic professors more masculine than democratic professors.

Dimension of leader likeability. This study predicted that all research participants (male and female) would rate the democratic professors higher on the likeability scale than the autocratic counterpart regardless of type of situation (task or personal).

Interaction Effects

Interaction between professor gender, professor leadership style, type of situation, and participant gender with regard to competence and likeability ratings.

Figures 1-4 represent a visual depiction of the predicted four-way interaction effects for the competence and likeability factors. Specifically, the visual display illustrates the differential predictions for male and female research participants. Comparing Figures 1 and 2 with Figures 3 and 4 illustrates the prediction that male research participants will be similar to female research participants in their competence and likeability ratings in the democratic condition. In the autocratic conditions, however, male and female research participants are predicted to differ, that is, male research participants (but not female research participants) would also be influenced by professor gender and type of situation when evaluating leader competence and leader likeability. Using gender role

spill over theory, it was predicted that male research participants would rate autocratic female professors in a personal situation the least competent, followed by autocratic female professors in a task situation (Figures 1- 4).

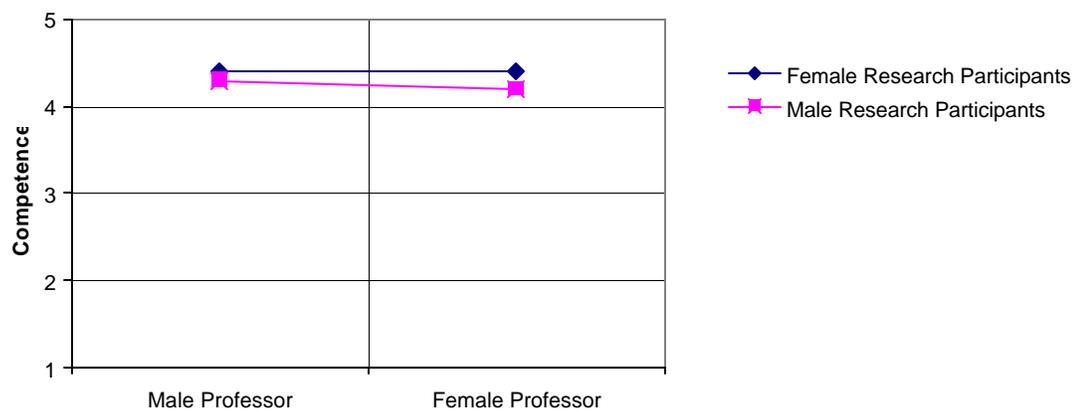


Figure 1 . Competence Ratings By Male and Female Participants for Professors Acting Democratically in a Task-Oriented Situation

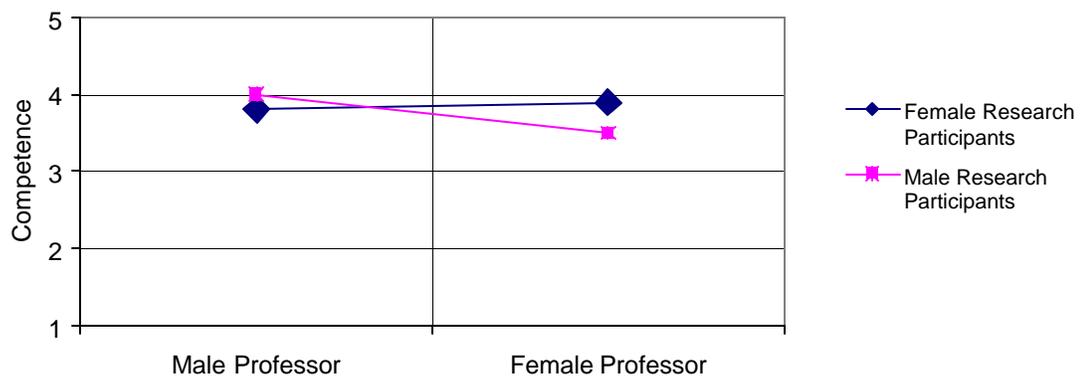


Figure 2 . Competence Ratings By Male and Female Participants for Professors Acting Democratically in a Personal-Oriented

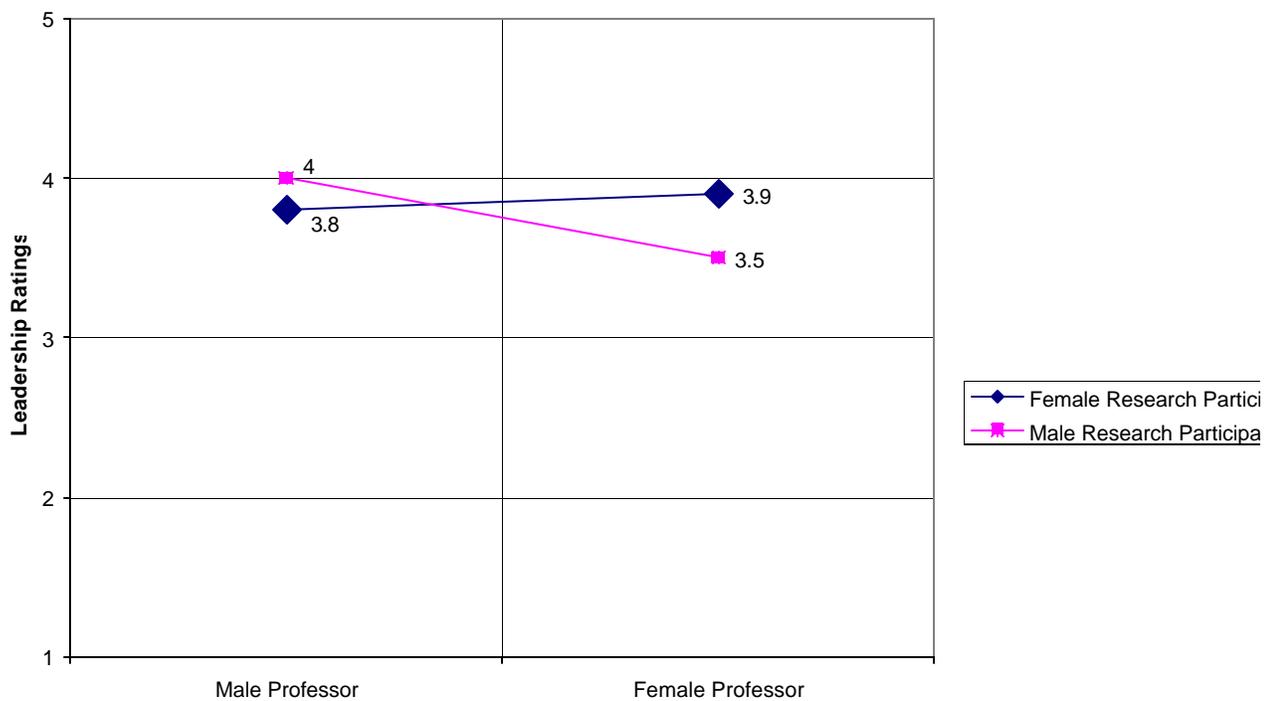


Figure 3. Democratic Personal

Section II - Research Participants

Because past leadership and gender research has generally revealed small to medium effect sizes it was determined that a large sample size was needed to achieve a level of power of .80 (Keppel, 1991). The sample size for this study, a 2 (professor gender) x 2 (professor leadership style) x 2 (type of situation) x 2 (participant gender) between factors design, was estimated using Stevens's power tables. Based on these power calculations it was determined that a minimum of 100 research participants per cell (16 cells total) was needed, resulting in a sample size of 1600 participants.

The research was conducted at the second largest metropolitan state university in the southeast. Out of the 37,500 enrolled students, approximately 27,000 students are enrolled in undergraduate programs in 10 Colleges (Table 1). This study selected the College of Arts and Science and the College of Education because of their size, range of departments and their willingness to participate.

Table 1

Colleges Within the Participating University

Colleges
Architecture and Community Design
Arts & Sciences *
Business Administration
Education *
Engineering
Marine Science
Medicine
Nursing
Public Health
Visual and Performing Arts

Note. Research participants were selected from Colleges marked with *

The sample consisted of 1653 undergraduate students from the College of Art and Science (n=931) and the College of Education (n=722). Table 2 compares the student race/ethnicity profile of the sample used in the study with the race/ethnicity profile of the total student population of the university. Racial distribution of the sample was similar to the University population.

Table 2

Student Race/Ethnicity Profile of Sample and University Population

	Study Sample		University Total Student Population	
	N=1653	%	N=37500	%
African/American	147	9	1305	11
American/Indian	14	1	146	1
Asian	60	4	1924	5
Caucasian	1148	70	26798	71
Hispanic	164	10	3528	9
Other	119	7	1146	3

Note. percentages were rounded to whole numbers

The study consisted of 1049 (63%) females and 604 (37%) males, as compared to 59% female and 41% male for the total student population. The average age for the sample was 21.92 years (SD=4.92 years) compared to 24 years (unknown SD) for the undergraduate student population as reported for the Fall of 2001. Of the 931 research participants from the College of Arts and Science, 464 (50%) were males and 467 (50%) were females. Of the 722 research participants from the College of Education 140 (19%) were males, and 582 (81%) were females.

Three departments within the College of Arts and Science, and three departments within the College of Education were used for data collection (Table 3). The goal was to access introductory classes in order to maximize research participant volunteers. Class sizes ranged from 30 to 200 students. Exact class sizes could not be established because student attendance varied.

Table 3

Source Information of Study Sample

College	Department	Number of Classes	Class Size Range
Arts & Science	Geography	3	100-200
	Language/Linguistic	10	10-40
	Communication	10	20-40
Education	Language/Acquisition	8	20-100
	Social Foundation	15	15-30
	Measurement	6	10-30
Total		52	

Note. Class size ranges are given because exact numbers were not available

Because most classes where data collection occurred were introductory classes for both colleges, the reported majors were more reflective of the college than the department where testing took place. Tables 4 and 5 list reported majors by gender separately for each College.

The most frequent major for male and female research participants for the College of Arts and Science was Communication. Both genders were comparable in their reported majors.

The majority of female research participants from the College of Education were Elementary education majors, whereas male research participants were not concentrated in any one area.

Table 4

Reported Majors for Male and Female Arts and Science Students (N=931)

	Males N=464		Females N=467	
	N	%	N	%
Communication	133	34	119	30
Criminology	47	12	19	5
ISS	44	11	32	8
Psychology	25	6	47	12
Languages	14	4	34	8
Biology	30	8	56	14
Geography	16	4	12	3
Geology	4	1	6	2
Mass Communication	18	5	37	9
Sociology	8	2	2	1
History	13	3	11	3
Classics	10	3	6	2
Chemistry	18	5	8	2
Political. Science	9	2	0	0
English	9	2	15	4

Note. Percentages were rounded to whole numbers.

Table 5

Reported Majors for Male and Female Education Students (N=722)

	Males N=140		Females N=582	
	N	%	N	%
Elementary Ed	16	12	293	52
Special Ed	20	14	71	13
Early Child	3	2	65	12
Physical Ed	26	19	20	4
Math Ed	11	8	16	3
Social Sciences	19	14	24	4
Secondary Ed	16	12	25	5
Ed Leadership	2	1	2	1
Language Ed	5	4	5	1
English Ed	10	8	24	4
Adult Ed	1	1	5	1
Music Art Ed	10	8	12	2

Note. Percentages were rounded to whole numbers.

There were differences between the samples from the College of Arts and Science and the College of Education in terms of age distribution, race/ethnicity, and distribution of school standing. With regard to age, results from an ANOVA revealed a main effect for college and gender. Specifically, research participants for the College of Education were statistically significantly older than their Arts and Science counterparts, $F(1, 1345)=93.53, p<.01$. The mean age as reported by the research participants for the College of Arts and Science was 20.86 ($SD=3.79$), while the mean age of research participants from the College of Education was 23.36 ($SD=5.84$). The effect size estimate for this difference was medium ($d=0.5$). Results further revealed a statistically significant gender effect for age, $F(1, 1345)=21.14, p<.01$. Male research participants had a mean age of 22.20 ($SD=5.68$), while female research participants had a mean age of 21.76 ($SD=4.41$). The effect size estimate for this difference was

very small ($d=0.09$). No interaction effect between college and gender was found, $F(1, 1345)=0.02, p>.05$. Curiously, female participants from the College of Education had a higher tendency of omitting the question of age. Table 6 lists age summary statistics for male and female research participants for each college.

Table 6

Summary Statistics for Age by College and Participant Gender

	Arts & Science		Education	
	Males (N=464)	Females (N=467)	Males (N=140)	Females (N=582)
Mean Age	21.55	20.20	24.53	23.10
Range	18-56	18-27	18-57	18-48
SD	5.1	1.67	7.10	5.50
Missing	79	69	31	120

Differences between the two samples of this study were also found in the distribution of school standing. Results of testing the differences in proportions indicated that the College of Arts and Science had a significantly higher proportion of freshmen, $\chi^2(1, N=1653)=108.63, p<.01$. The College of Arts and Science had a statistically significant higher proportion of sophomores than the College of Education, $\chi^2(1, N=1653)=24.91, p<.01$. On the other hand, the College of Education had a statistically significant higher proportion of juniors in this study than the College of Arts and Science, $\chi^2(1, N=1653)=47.36, p<.01$. The College of Education also had a statistically significant higher proportion of seniors in this study than the College of Arts and Science, $\chi^2(1, N=1653)=19.22, p<.01$. Table 7 lists the frequency distributions for class standing separately for each college.

Table 7

Frequency Distribution of School Standing by College and Gender

Class Standing	Arts & Science				Education			
	Males N=464		Females N=467		Males N=140		Females N=582	
	N	%	N	%	N	%	N	%
Freshman	114	25	82	17	5	4	20	4
Sophomore	116	25	118	25	20	14	89	16
Junior	105	23	149	32	55	40	259	45
Senior	124	27	117	25	58	42	201	35
Missing	5		1		2		13	
N=1653								

Differences were found in the race/ethnicity profile of the two colleges in this study. Results of testing the differences in proportions indicated no statistically significant difference between the colleges in the frequency distribution of the American/Indian category, $\chi^2(1, N=1653)=2.4.3, p >.05$. A statistically significant difference was found in the proportion of African/American students, $\chi^2(1, N=1653)=12.39, p <.01$. The proportion of African/Americans in Arts and Science was statistically significantly higher than the College of Education counterpart. The proportion of Caucasians was statistically significantly higher in the College of Education than in the College of Arts and Science, $\chi^2(1, N=1653)=62.74, p <.01$. The College of Arts and Science had a statistically significantly higher proportion of Asians than the College of Education, $\chi^2(1, N=1653)=20.81, p <.01$. Finally, the College of Arts and Science had a statistically significantly higher proportion of Hispanics than the College of Education, $\chi^2(1, N=1653)=19.51, p <.01$. Table 8 lists the frequency distribution of race/ethnicity by gender and separately for each college.

Table 8

Racial/Ethnicity Frequency Distribution by College and Gender

Race/Ethnicity	Arts & Science				Education			
	Males N=464		Females N=467		Males N=140		Females N=582	
	N	%	N	%	N	%	N	%
African/American	47	10	56	12	10	7	34	6
American/Indian	3	1	2	1	2	1	7	1
Asian	22	5	29	5	1	1	8	1
Caucasian	296	64	277	59	114	81	461	79
Hispanic	62	14	57	12	3	2	42	7
Other	33	7	46	10	10	7	30	5
N=1653								

*Section III – Instrument Development**Experimental Conditions*

Most research evaluating the effects of gender stereotypes on the evaluations of leaders have typically used written scenarios or vignettes (Rojahn & Willemsen, 1994). Consistent with this paradigm, the present study used the vignette format in order to manipulate type of situation, professor gender, and professor leadership style, while keeping all other characteristics of the scenario constant. Specifically, two scenarios were created, one describing a task-oriented situation, and the second describing a person-oriented situation. Embedded within these scenarios was a professor-student dyad, with the professor described as either male or female, and the professor's interaction style with the student described as either democratic or autocratic. All other descriptions of the scenario were held constant. Crossing all possible combinations of the independent variables (situation, professor gender, professor leadership style) resulted in eight

vignette versions (see Appendices C and D for complete set of booklet versions). Prior to the experiment, a pilot was conducted to obtain a measure of construct validity for the scenario manipulations. Appendix A details the procedures and results supporting the validity of the scenario manipulations.

Task Situation

A 175-word vignette was created of a professor-student dyad interaction (Appendix C) that involved a task-oriented problem. Specifically, the scenario described a student, enrolled in a course taught by this professor, entering the professor's office in order to discuss a solution to increase student participation during class. The student suggests access to the professor's notes before class as a solution to the problem.

Personal Situation

A 175-word vignette was created that described a professor- student dyad interaction that involved a personal situation (Appendix D). Specifically, the scenario described a student of the professor, entering the professor's office in a state of emotional distress revealing despair over a recent relationship breakup.

Professor Nonverbal Democratic Leadership Style for Task Situations

The nonverbal democratic leadership style between professor and student was manipulated using the following nonverbal descriptors: smiling encouragingly, nodding his/her head while the student talks, leaning forward towards the student while listening to the student, gently tugging at his/her collar while listening to the student, and accompanying the student to the door at the end of the conversation (Appendix C).

Professor Nonverbal Autocratic Leadership Style for Task Situations

The nonverbal autocratic leadership style between professor and student was manipulated using the following nonverbal descriptors: frowning, sitting straight in his/her chair and staring at the student, leaning back in his/her chair, lowering his/her eye brows, shaking his/her head, pointing his/her finger at the student, and abruptly standing up and motioning the student to the door at the end of the conversation (Appendix C).

Professor Nonverbal Democratic Leadership Style for Personal Situations

The nonverbal democratic leadership style between professor and student for this condition was manipulated using the following nonverbal descriptors: nodding encouragingly, sitting in a chair opposite the student, leaning forward towards the student, and folding his/her arms in his/her lap while the student tells the story. Finally, the professor remains seated and waits for the student to finish talking before offering to make an appointment (Appendix D).

Professor Nonverbal Autocratic Leadership Style for Personal Situations

The nonverbal autocratic leadership style between professor and student for this condition was manipulated using the following nonverbal descriptors: not paying attention to the student, displaying impatience, sighing, crossing his/her arms over his/her chest while the students speaks, and staring at the student (Appendix D).

Professor –Student Dyads

Professor gender within every professor–student dyad was manipulated while student gender remained unidentified. As a result, every type of scenario combination had a version for male and female professor-student dyads.

Development of Dependent Measures

Meta-analytic studies (e.g., Eagly, Makhijani, & Klonsky, 1992) have revealed that most studies measured leader evaluation using the following aspects of leadership performance: (a) competence, expertise, effort, productivity, and general evaluation, (b) leader likeability, desire to work with him or her, and group cohesiveness, and (c) leader power, authority, and influence. Most recent studies (e.g., Carli, LaFleur, & Loeber, 1995; Luthar, 1996) also included lists of adjectives measuring leader effectiveness, likeability, and personality traits. For example, Rojahn and Willemssen (1996) used a list of 26 gender-stereotypical traits in their assessment of the perception of leader character traits.

Based on measures of previously published studies, the current study initially selected 26 items for four leadership dimensions: leader competence (7 items), leader likeability (4 items), leader masculinity/femininity (4 items), and leader professional traits (11 items) (Table 9). These measures were adjusted to reflect an academic context. A draft version of the 26 items was initially reviewed by 10 experts (two male and three female professors in the College of Education, as well as three male and two female professionals in management positions). Reviewers were asked to check the 26 items for content and clarity, as well as to identify possible alternatives and suggestions for improving these leadership measures. Based on their feedback, two items (skillful/not skillful, productive/unproductive) under the leader competence factor were dropped. Further, three items were reversed and the placement of items were randomly ordered rather than ordered under their concept. Table 9 lists the leadership items for each leadership construct.

An initial pilot study was conducted within the College of Education (see Appendix B for a detailed description of procedures and results) to examine internal consistency reliability of the four leadership factors: leader competence (5 items), leader likeability (4 items), masculinity/femininity (4 items), and leader professional traits (11 items). The final 24 leadership items were arranged on a five point semantic differential rating scale. Low scores (1) were designed to represent the negative aspects of three of the factors (i.e., incompetence, not likeable, and negative professional traits). A low score on the masculinity/femininity scale represented femininity, while the high end of the scale (5) reflected masculinity.

In addition to the 24 leadership items, four exploratory items (see Appendix C or D) were presented about the professors' behaviors in the scenario. These items asked the research participants to rate (using a 5-point Likert type scale) the professor regarding his/her potential for promotion, communication skills, level of respect, and level of likeability. The 5- point rating scale ranged from 1 (strongly agree) to 5 (strongly disagree), which were then reversed scored to reflect a high score with agreement.

Eight booklet versions were created to accommodate the eight scenario versions. In order to maximize ease of handling, a booklet was created by folding 17x11 in paper to create a four - page booklet of 8 1/2 x 11 in. The front page contained a reminder that participation was voluntary, followed by directions. Upon opening the booklet, page two contained one version of the professor – student interaction, followed on the next page (page three) by the 24 randomly ordered leadership items. Finally, the back page (page four) of the booklet contained four short comments about the professor, each paired with

a five point rating scale. Finally, background information (participant's gender, age, major, school standing, and race/ethnicity) was requested.

Table 9

Initial Items for the Four Leadership Dimensions

Leader Competence (7 items)	Masculinity/Femininity (4 items)	Leader Likeability (4 items)	Professionanl Traits (11 items)
Competent/incompetent	Timid/forceful	Critical/tolerant	Powerful/powerless
Effective/ineffective	Soft/tough	Considerate/Inconsiderate	Hardworking/lazy
Qualified/not qualified	Aggressive/not aggressive	Popular/unpopular	Persistent/gives up easily
Influential/not influential	Dominant/submissive	Likeable/not likeable	Fair/not fair
Capable/not capable			Responsible/irresponsible
Skillful/Not skillful*			Not helpful/helpful
Productive/unproductive*			Cooperative/not cooperative
			Trustworthy/Untrustworthy
			Independent/dependent
			Objective/subjective
			Unprepared/prepared

Note. Items marked with ‘*’ were dropped in the final version

Procedure

Instructors from the two colleges were contacted via email with a letter introducing the researcher and the purpose of the study and a request for an appointment (Appendix E). Most appointments secured the permission of the researcher to enter the classroom to administer the instruments. The researcher entered the classroom only by appointment and permission of the instructor at the beginning of class, or in some cases at

the end of class. The majority of data collection was conducted by the researcher. However, if an instructor wanted to hand out the booklets him/herself, he or she was given an approximate 10 minute training session in an effort to standardize data collection. Before each testing session the researcher randomly ordered the booklet forms appropriate for that class. Generally, data were collected at the beginning of class. After being introduced by the instructor as a graduate student working on her dissertation, the researcher gave her standard verbal introduction with regard to the purpose of the study, and the need for volunteer participants, and the anonymity of responses. Booklet forms were then distributed to research participants who indicated their willingness to volunteer by raising their hand. The researcher also stressed that the nature of the scenario was a matter of personal opinions on the part of the student, and that there were no right or wrong answers. Participants were reminded not to put their names on the booklet. Upon receipt of the booklet, participants were instructed to first read the front page of the booklet for instructions, then to read the scenario, and fill out all the questions including the back page of the booklet that contained general background information about the participant. While working through the booklet, participants were asked not to communicate with their fellow students, and to keep their comments or questions until all the booklets had been collected. The majority of instructors granted extra time for the researcher to answer questions that were posed after the booklets had been filled out. Generally, students took about 10 minutes to fill out the booklets.

Chapter IV

Results

This chapter is divided into three sections. Section I presents the analysis of the psychometric properties of the measures of leadership. Section II reports the results of the five hypotheses posed by this study. Section III presents exploratory findings of this study.

Section I Psychometric Analysis of the Four Leadership Dimensions

Based on results from the literature review and pilot study (Appendix B), 24 leadership items were created that used a five-point semantic differential response scale. These items were designed to measure the following four leadership dimensions: Competence (5 items), Masculinity/Femininity (4 items), Likeability (4 items), and Professional Traits (11 items).

Competence

Five items were proposed for this factor. Table 10 lists the summary statistics. For these items the low end of the five point scale reflected incompetence, while the higher end of the scale reflected competence (all five items were reversed for this purpose).

Masculinity/Femininity

This factor was represented by four items (Table 11). The lower range of the five point scale reflected femininity, while the high end of the scale reflected masculinity. Two items were reversed scored.

Table 10

Summary Statistics for Proposed Competence Factor

Item	Mean	SD	Skewness	Kurtosis
Competent/Incompetent*	3.36	0.99	-0.27	-0.21
Effective/ineffective*	2.71	1.16	0.25	-0.77
Qualified/not qualified*	3.21	0.95	0.01	-0.03
Capable/not capable*	3.45	0.96	-0.23	-0.19
Influential/not influential*	3.28	1.06	-0.25	-0.36

N=1652 Note: Items with an * were reversed scored.

Table 11

Summary Statistics for Proposed Masculinity/Femininity Factor

Item	Mean	SD	Skewness	Kurtosis
Timid/forceful	3.59	0.96	-0.46	0.02
Soft/tough	3.58	1.00	-0.46	-0.21
Aggressive/not aggressive *	3.01	1.21	-0.16	-0.67
Dominant/submissive *	3.72	0.98	-0.37	-0.40

N=1652 Note: Items with an * were reversed scored.

Likeability

Four items were developed for this factor (Table 12). A low score on the five-point response scale indicated not likeable, while the high end of the scale reflected a high degree of likeability. Three of the four items were reversed scored.

Professional Traits

This factor consisted of 11 items (Table 13). A low score on the professional traits items reflected the negative side of the trait, while a high score reflected the desired professional trait (nine items were reversed scored).

Table 12

Summary Statistics for Proposed Likeability Factor

Item	Mean	SD	Skewness	Kurtosis
Critical/tolerant *	2.64	1.16	0.26	-0.76
Considerate/inconsiderate	2.61	1.29	0.35	-1.00
Popular/unpopular *	3.18	1.14	-0.16	-0.68
Likeable/not likeable *	2.81	1.27	0.13	-1.00

N=1652 Note: Items with an * were reversed scored.

Table 13

Summary Statistics for Proposed Professional Traits Factor

Item	Mean	SD	Skewness	Kurtosis
Powerful/powerless*	3.59	0.92	-0.27	-0.11
Hardworking/lazy*	3.31	1.00	-0.07	-0.13
Persistent/gives up easily*	3.19	1.10	-0.25	-0.54
Fair/not fair*	3.10	1.00	-0.05	-0.45
Responsible/irresponsible*	3.36	1.00	-0.26	-0.48
Not helpful/helpful	2.64	1.20	0.25	-0.98
Cooperative/not cooperative*	2.82	1.20	0.08	-0.90
Trustworthy/untrustworthy*	3.21	1.00	-0.04	-0.45
Independent/dependent*	3.66	1.00	-0.52	-0.00
Objective/subjective*	3.10	0.98	-0.07	-0.13
Unprepared/prepared	3.23	1.10	-0.16	-0.67

N=1652 Note: Items with an * were reversed scored.

Principal axis factoring specifying four factors

Exploratory factor analysis was used to analyze the underlying structure of the 24 leadership items. Results of a previously conducted pilot study, in addition to the results from an extensive review of the literature, suggested a four factor solution (Competence, Masculinity/Femininity, Likeability, and Professional Traits). Based on these results, the initial exploratory factor analysis (principal axis procedure) used a four factor solution with an oblique rotation (promax). This procedure involved the whole sample (n=1652).

Initial estimates of communality were the squared multiple correlations of the correlation of each item with the remaining items.

In order to determine whether the four factor solution was reasonable, the initial unrotated eigenvalues were inspected, and compared to (a) the expected eigenvalues generated by the null model with the same parameters, (b) the Kaiser criterion of retaining eigenvalues greater than one, (c) the average eigenvalue of the four-factor solution, and (d) logical interpretation of the factors. Table 14 lists the unrotated eigenvalues, the explained proportion of that eigenvalue based on the common variance solution, and the expected eigenvalues of the null model.

Table 14

Eigenvalues for Principal Factor Extraction Method compared to Null Model

Component	Eigenvalue (n=1652)	Proportion	Null Model Eigenvalues
1	8.30	0.72	0.21
2	3.18	0.27	0.17
3	0.44	0.03	0.15
4	0.32	0.02	0.13
5	0.19	0.01	0.11
6	0.15	0.01	0.10
7	0.11	0.009	0.08
8	0.08	0.007	0.07
9	0.03	0.003	0.06
10	0.02	0.002	0.04
Average Eigenvalue=0.480			

Inspection of the eigenvalues did not support a four factor solution based on the above stated criteria. Specifically, the third and fourth eigenvalues of this four factor solution were below the Kaiser criterion of one, as well as just below the average eigenvalue of 0.480 (of the common solution). Yet, the values of the third and fourth eigenvalues were slightly higher than the eigenvalues calculated by the null model (Table 14). Since the four factors correlated moderately high with one another, both the pattern matrix (standardized regression coefficients) and structure matrix (zero-order correlations) were examined (see Table 15).

In general, items with factor loadings larger than 0.30 are considered strong factor loadings (Rummel, 1988). In light of this general rule, several of the leadership items that exhibited a high amount of cross correlations in the structure matrix (correlations), retained their item cross loading pattern when examining the standardized regression coefficients (pattern matrix in Table 15). Specifically, nine out of the 24 items exhibited moderately high cross loadings (items marked with an * in Table 15).

The factor loadings of the 24 leadership items were examined for (a) the magnitude of loadings on the factors, and (b) the magnitude of cross loadings on secondary factors. Using .30 or larger as an indication of a strong factor loading, 11 of the 24 items loaded most strongly on the first factor. Similarly, the second factor consisted of 11 items. Factor three had two items with strong loadings, while none of the 24 items loaded strongly on the fourth factor. Further, as stated earlier, nine of the 24 items had strong cross loading tendencies on at least one other secondary factor.

Item clustering of the four factor solution was compared with the originally proposed clustering of the four leadership factors (competence, masculinity/femininity, likeability, and professional traits). Factor one combined four out of the five competence items with seven of the professional traits items. Factor two combined all four masculinity/femininity items along with all four items originally suggested to belong to the likeability construct. In addition, factor two also retained three items originally proposed to belong to the professional traits cluster. Finally, two professional trait indicators loaded strongly on factor three. None of the leadership items had a strong association with factor four.

The originally proposed four dimensions of leadership were not supported by the results of the four factor solution using the current sample of 1652 research participants. Instead, these results suggested a two factor solution underlying the 24 leadership items. Thus, given these results the principal axis factor analysis with oblique rotation (promax procedure) was rerun with two factors specified.

Principal axis procedure with two factors

The two factors extracted accounted for approximately 99% of the common variance (see Table 16 for information on pattern and structure matrices). Two leadership items (independent/dependent, objective/subjective) were dropped from both factors even though their factor loadings satisfied the factor loading criterion of .30 (.34 for independent/dependent and .38 for objective/subjective). Specifically, these items were dropped because their factor loadings were substantially lower than the factor loadings of the remaining 22 leadership items. In addition, one item (critical/tolerant) was eliminated because of its strong negative factor loading (-.60) on factor two. This

item was originally proposed to belong to the likeability factor (Table 12), and reversing the direction would have seriously impeded the interpretation of this factor such that the high score would have linked likeable, popular, and considerate with critical. As a result, 21 leadership items remained in the solution. Table 16 summarizes the pattern matrix and structure matrix from the two factor solution.

Table 15

Factor Pattern and Structure Matrices for Four-Factor Solution

Item	Pattern Matrix (Standardized Regression. Coefficients)				Structure Matrix (Correlations)			
	1	2	3	4	1	2	3	4
Incompetent/competent	56	-.10	9	29	65	-.43	51	30
Ineffective/effective	64	-.24	-4	12	71	-.52	47	15
Qualified/not qualified	75	9	-2	31	69	-.24	45	9
Capable/not capable	59	6	26	7	73	-.28	63	5
Influential/not influential	18	-10	48	0	54	-.33	64	5
Critical/tolerant	-5	-.67	-3	-6	21	-.63	15	9
Considerate/inconsiderate *	41	-.61	0	-11	67	-.76	47	1
Popular/unpopular *	4	61	-32	-5	-42	71	-.50	-20
Likeable/not likeable	-19	67	-10	-14	-.53	81	-.45	-28
Timid/forceful *	2	67	32	-10	-5	58	10	-23
Soft/tough	-4	73	1	-9	-.33	76	-.26	-.25
Aggressive/not aggressive	20	75	7	-2	-.7	65	-.5	-19
Dominant/submissive	8	76	36	2	-1	61	16	-12
Powerful/powerless *	21	58	45	-3	27	35	40	-13
Hardworking/lazy	49	6	23	-1	61	-.21	53	-3
Persistent/gives up easily *	67	37	4	-11	55	10	35	-21
Fair/not fair *	40	-.45	13	2	67	-.66	54	11
Responsible/irresponsible *	68	-.9	0	31	71	-.44	50	31
Not helpful/helpful *	54	-.46	-3	-16	72	-.64	47	-8
Cooperative/not cooperative *	32	-.55	9	-11	62	-.69	48	0
Trustworthy/untrustworthy	41	-.29	13	30	61	-.57	51	36
Independent/dependent	6	29	41	5	20	12	35	1
Objective/subjective	34	1	4	1	36	-.15	26	0
Unprepared/prepared	78	21	-19	-1	57	-.5	26	-9

N=1652 Note: Items with an * items have strong cross loadings.

The rotated eigenvalue of factor 1 was 6.8, and explained 59% of the common variance in the solution. Out of the 21 leadership items, 14 items loaded primarily on this factor. This factor captured the following leadership items: (a) all five items from the competence factor (Table 10), (b) eight items from the professional traits factor (Table 13), and (c) one item from the likeability factor (Table 12). In order to name this factor appropriately, item factor loadings were judged in terms of their strength, as well as their level of cross loading on the second factor. Based on these criteria, the five original competence items loaded most prominently (unique factor loadings ranged from .62 to .80) on this factor, while at the same time exhibiting minimal cross loading for the second factor (Table 16). Next, the considerate/inconsiderate leadership item from the originally proposed likeability factor loaded strongly on this factor (factor loading of .55), while at the same time loading almost as strongly on the second factor (factor loading .47). Finally, of the eight leadership items originally proposed to belong to the professional trait factor, four items (persistent/gives up easily, fair/not fair, not helpful/helpful, and cooperative/not cooperative) exhibited strong factor loadings on both factors, while the other four professional traits items (hardworking/lazy, responsible/irresponsible, trustworthy/untrustworthy, and unprepared/prepared) had factor loadings for only the primary factor (unique factor loading ranged from .59 to .72).

The purpose of factor rotation is to obtain factors that consist of items having strong and meaningful factor loadings on only one factor. Results of this factor analysis were mixed. Specifically, out of the 14 items loading on factor 1, the five competence items (Table 10) exhibited strong and meaningful factor loadings for this factor only. Similarly, four out of the eight professional trait items also had strong primary factor

loadings and minimal secondary factor loadings. On the other hand, the remaining four professional traits indicators (persistent/gives up easily, fair/not fair, not helpful/helpful, cooperative/not cooperative), as well as the one likeability indicator (considerate/inconsiderate) had strong factor loadings on both factors (Table 16).

Table 16

Factor Pattern and Structure Matrices for the Two-Factor Solution

Item	Pattern Matrix (Standardized Regression Coefficients)		Structure Matrix (Correlations)	
	1	2	1	2
Incompetent/competent	67	-8	70	-33
Ineffective/effective	67	-18	73	-43
Qualified/not qualified	73	14	67	-13
Capable/not capable	80	16	74	-14
Influential/not influential	62	3	28	-62
Critical/tolerant	6	-60	72	-68
Considerate/inconsiderate	55	47	-55	64
Popular/unpopular	-35	50	-64	75
Likeable/not likeable	-42	59	-64	75
Timid/forceful	15	69	-11	63
Soft/tough	-18	68	-44	75
Aggressive/not aggressive	11	72	-16	68
Dominant/submissive	23	76	-6	67
Powerful/powerless	48	64	24	46
Hardworking/lazy	68	17	62	-9
Persistent/gives up easily	64	46	47	22
Fair/not fair	60	-33	73	-56
Responsible/irresponsible	72	-7	75	-34
Not helpful/helpful	62	31	73	-54
Cooperative/not cooperative	52	-40	67	-60
Trustworthy/untrustworthy	59	-26	69	-48
Independent/dependent	34	33	22	20
Objective/subjective	38	6	36	9
Unprepared/prepared	60	26	50	3

N=1652

Because of their highly ambiguous cross loadings, these five items were dropped from factor 1. The remaining nine items in the factor solution were further examined in terms of their conceptual contribution. Based on this criterion the content of the five competence items (incompetent/competent, ineffective/effective, qualified/not qualified, capable/not capable, influential/not influential) demonstrated a clear and cohesive structure. Next, of the four professional traits items that loaded highly on this factor (i.e., hardworking/lazy, responsible/irresponsible, trustworthy/untrustworthy, unprepared/prepared) one item was eliminated (unprepared/prepared) because its content did not contribute to a meaningful interpretation of this factor (it described a personal style rather than the desired general quality of leader competence).

To summarize, factor 1 originally had 14 items, of which six were eliminated from the solution. The theme of leader competence was clearly expressed in the content of these items, thus, factor 1 was named competence (Table 17). The competence factor response scale ranged from 1 (not competent) to 5 (competent).

The rotated eigenvalue of factor 2 was 4.69 and explained approximately 40% of the common variance in the solution. Of the 21 leadership items, seven items loaded primarily on this factor (see pattern matrix in Table 16). Specifically, the following leadership items identified this factor: (a) all four items originally proposed as the masculinity/femininity factor (timid/forceful, soft/tough, aggressive/not aggressive, dominant/submissive), (b) two of the items of the originally proposed likeability factor (popular/unpopular, likeable/not likeable), and one item from the original professional traits factor (powerful/not powerful). Like in factor 1 (competence), items for this factor were examined in terms of the strength of their unique factor loadings, as well as

their factor loadings on the secondary factor (Table 16). Based on these criteria, the four original items from the masculinity/femininity factor exhibited the highest factor loadings (ranging from .68 to .76), while loading minimally on the competence factor. Next, the two items from the likeability factor (popular/unpopular, likeable/not likeable) had relatively strong factor loadings on this factor, but also exhibited strong negative factor loadings on the competence factor. Similarly, the powerful/not powerful item, originally part of the professional traits cluster, showed a similarly strong double loading (Table 16).

In order to maximize factor interpretation, the seven items for this factor were examined not only for their factor loadings but also for their contribution to factor interpretability (i.e., high cross loading on the other factor and content). Based on these criteria, the four items from the masculinity/femininity scale (timid/forceful, soft/tough, aggressive/not aggressive, dominant/submissive) not only had high factor loadings on this factor, but in addition had minimal cross loadings on the competence factor. Unfortunately, the remaining three items (popular/unpopular, likeable/not likeable, powerful/not powerful) strongly loaded on this factor as well as on the competence factor, which at best would have complicated the interpretation of this factor. Therefore, these three items were dropped from this factor. As a result, Factor 2 emerged with the original core items from the masculinity/femininity factor. As a result, this factor was named masculinity/femininity. The five point scale ranged from (1) femininity to (5) masculinity. Table 17 lists the summary statistics for these two factors.

Table 17

Factor Names and Indicator Items

Factor 1 Competence (8 items)	Factor 2 Masculinity/Femininity (4 items)
Incompetent/competent	Timid/forceful
Ineffective/effective	Soft/tough
Qualified/not qualified	Aggressive/not aggressive
Capable/not capable	Dominant/submissive
Influential/not influential	
Hardworking/lazy	
Responsible/irresponsible	
Trustworthy/untrustworthy	

N=1652

Table 18

Summary Statistics for Competence and Masculinity/Femininity Factors

		Total N=1653	Males N=604	Females N=1049
Competence (8 items)	M	3.24	3.29	3.22
	SD	0.76	0.75	0.77
	Skewness	-0.01	-0.03	-0.00
	Kurtosis	-0.14	0.08	-0.26
	Alpha	.88	----	-----
Masculinity/ Femininity (4 items)	M	3.50	3.43	3.54
	SD	0.81	0.77	0.83
	Skewness	-0.10	-0.03	-0.16
	Kurtosis	-0.66	-0.49	-0.74
	Alpha	0.79	-----	-----

Note. Correlations between competence and masculinity/femininity for the total sample was -.24, for males the correlation was -.16, and the correlation for females was -.28.

Comparison of two factor results with originally proposed four factors

Originally, four leadership factors were proposed: competence (five items), masculinity/femininity (four items), likeability (four items), and professional traits (eleven items). Results from the current study revealed evidence for two factors: competence (eight items) and masculinity/femininity (four items). The competence factor included all originally proposed competence items. In addition, three leadership items from the originally proposed professional traits factor loaded strongly on this factor. Overall, the five competence items along with the three items originally attributed to professional traits represented the construct of leadership competence. The masculinity/femininity factor of the current study included all four of the original items developed for this factor. Even though two items from the original likeability factor and one item from the original professional trait factor also loaded strongly on this factor, they were dropped because of strong factor cross loadings on the competence factor.

To conclude, four factors underlying the leadership construct were initially proposed. However, results from this study did not support four factors underlying the leadership construct. Instead, leadership was represented by two factors: leader competence (eight items) and leader masculinity/femininity (four items). These two factors consisted of all the originally proposed core items. In addition, three items from the professional traits concept emerged as items clustering with the competence factor. Surprisingly, the likeability factor in its proposed form did not hold up in this study.

Overview of Confirmatory Factor Analysis Procedure (CFA)

The purpose of the CFA was to examine the fit of the two-factor solution and to evaluate sources of misfit (e.g., correlated error for the items). Figure 5 illustrates the hypothesized two factor leadership model that was developed based on the result of statistical (i.e., exploratory factor analysis) and conceptual criteria. The measurement portion of the model consists of the causal arrows that go from the latent factors (competence and masculinity/femininity) to the manifest variables that measure them. Specifically, competence had eight manifest indicators, while masculinity/femininity was represented by four indicator variables. While the two latent factors were specified to intercorrelate (as noted by the double headed arrow in Figure 5) the indicators for each factor were specified to only load on their designated factor. In order to scale both latent variables, the loading for the first item for each factor was set to 1.0. For each of the 12 indicator variables measurement error variances were estimated (independence of error was assumed).

CFA was conducted using the PROC CALIS procedure available within the SAS statistical package. The raw data were used to create a covariance matrix by using the COV option. The data were analyzed using the maximum likelihood estimation procedure. Examination of skewness and kurtosis (Tables 10 & 11) indicated that most of the variables showed some degree of kurtosis, however their levels of deviation stayed within the moderate range of +/-1. As a rule of thumb, levels of kurtosis exceeding 1.5 could compromise the validity of the maximum likelihood procedure (Bollen, 1989).

Results of the χ^2 goodness of fit test indicated a statistical significant lack of fit, $\chi^2(1615, 53) = 850.16, p < .001$. Since the χ^2 goodness of fit test is greatly influenced by sample size (Bollen, 1989), sample independent fit indices have become better indicators for model fit. Of those, one of the most frequently reported and discussed fit indices is the root mean square error of approximation (RMSEA), which is an inferential approach assessing model discrepancy relative to the degrees of freedom (MacCallum, 1986). Guidelines suggest a good fit when RMSEA $< .05$, a moderate fit to range between $.05$ and $.08$, and a poor fit when exceeding $.10$. The RMSEA of this analysis was $.09$, which indicated a moderate to poor fit. Another sample independent index routinely reported is the comparative fit index (CFI). This fit index is based on the noncentrality parameter and is scaled from 0 to 1, with values greater than $.9$ indicating an acceptable fit. The CFI result of $.90$ indicated a minimally acceptable fit. Results showed that leader competence correlated moderately negative with leader masculinity ($r = -.35$). The standardized loading estimates for the eight competence items ranged from $.61$ to $.81$, while the four masculinity/femininity items had standardized loading estimates ranging from $.67$ to $.76$.

The Lagrange/Wald test modification indices suggested that one source of model misfit was that five indicator items for the competence factor had secondary loadings on the masculinity/femininity factor, while two of the masculinity/femininity indicators had secondary loadings on the competence factor (see Table 19).

Table 19

Summary Statistics of the Seven Largest Lagrange Chisquare Change Statistics

Original Factor	Indicator Item	Reduction in χ^2	If Crossloading With Factor
Competence	Ineffective/effective	112.56	Masculinity/ Femininity
	Responsible/irresponsible	61.15	
	Influential/not influential	51.80	
	Capable/not capable	47.70	
	Hardworking/Lazy	37.90	
Masculinity/ Femininity	Timid/forceful	224.53	Competence
	Aggressive/Not aggressive	76.70	

Four Factors Revisited

The purpose of the extensive series of exploratory and confirmatory factor analytic procedures was to derive leadership factors that had a high degree of construct validity and that were internally consistent. Based on pilot testing and results from the literature review, four factors underlying the leadership construct were proposed: (a) competence (five items), (b) likeability (four items), (c) masculinity/femininity (four items), and (d) professional traits (eleven items). However, empirical evidence did not support a four factor leadership model, instead, a two factor leadership structure emerged: (a) competence (eight items), and (b) masculinity/femininity (four items). The empirically derived competence factor consisted of the originally proposed competence items. In addition, three items from the professional traits factor loaded on this construct as well. The masculinity/femininity factor that emerged as a result of testing consisted of the original four masculinity/femininity items. As for the proposed professional traits factor, three out of the eleven items loaded strongly on the competence factor.

The likeability factor as proposed did not hold as a distinct factor (see Table 12 for summary statistics). Three of the four items (considerate/inconsiderate, popular/unpopular, likeable/not likeable) had strong dual loadings on the competence and masculinity/femininity factors, while one item (critical/tolerant) was eliminated because of its lack of conceptual fit. Despite the dual loadings for the three likeability items on the competence and masculinity/femininity factors, these three items also showed moderately strong intercorrelations (Table 20). Further the internal consistency of the three items was moderately strong ($\alpha = .84$). Given this moderate high alpha and the fact that likeability was a focus of this study, the researcher decided to combine these items (likeable/not likeable, popular/not popular, considerate/inconsiderate) to form a likeability scale. The scale ranged from (1) not likeable to (5) very likable. Table 20 shows summary statistics for the three leadership factors.

Table 20

Likeability Indicator Item Intercorrelations

Item	1	2	3
1. Likeable/notlikeable	1.0		
2. Popular/not popular	.66	1.0	
3. Considerate/inconsiderate	.70	.57	1.0

N=1652

Table 21

Summary Statistics for Leadership Factors (N=1652)

Factor	Number of Items	Mean	SD	Skewness	Kurtosis	Alpha
Competence	8	3.24	0.76	-0.01	-0.14	.88
Masculinity/ Femininity	4	3.49	0.81	-0.10	-0.66	.78
Likeability	3	2.75	1.10	0.23	-0.87	.84

Note. The correlation between competence and masculinity/femininity was -.24, and between competence and likeability was .69. The correlation between masculinity/femininity and likeability was -.63.

To conclude, extensive factor analytic procedures established two internally consistent factors of leadership: competence ($\alpha = .88$) and leader masculinity/femininity ($\alpha = .78$). In addition, a three-item likeability factor was created that demonstrated reasonable psychometric properties ($\alpha = .84$).

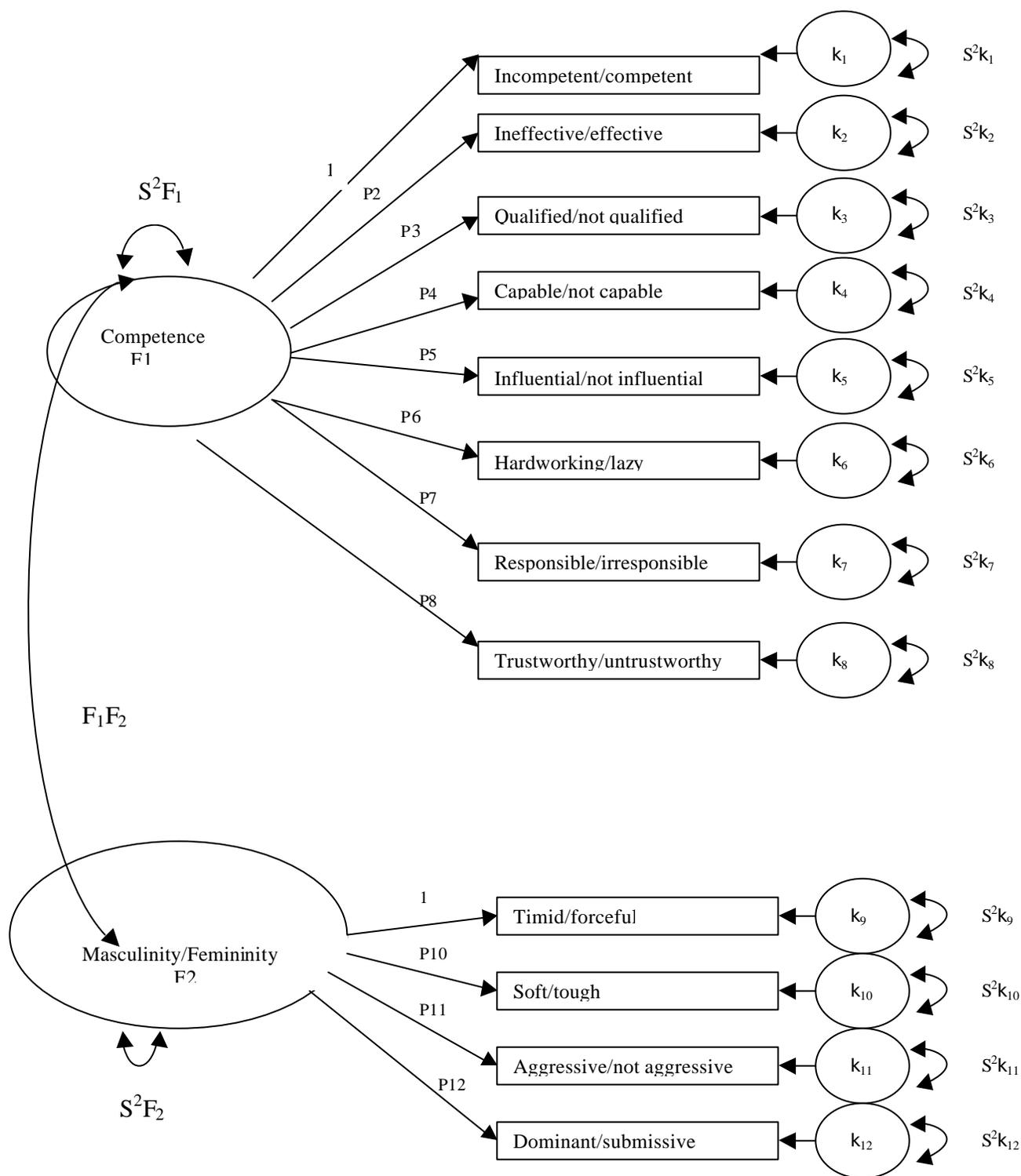


Figure 5. Two-factor confirmatory factor analysis model for leadership measures

Section II - Hypotheses Testing

The purpose of this experiment was to examine a set of five hypotheses in relation to the three leadership constructs of competence, masculinity/femininity, and likeability. The five hypotheses were tested using a 2 (professor gender) x 2 (professor behavior) x 2 (type of situation) x 2 (participant gender) between factors design. To enhance the interpretation of the statistically significant main and interaction effects, Cohen's effect sizes were calculated for statistically significant main and interaction effects as measured by Cohen's (1988) f metric:

$$f_{effect} = \sqrt{(df_{effect})F_{effect}/N}$$

Guidelines for the f metric describe effect sizes of 0.10 as small, 0.25 as medium, and 0.40 as large. Effect sizes were also calculated for follow-up pairwise comparisons using Cohen's d metric:

$$d = \frac{(\text{Sample Mean}_1 - \text{Sample Mean}_2)}{SD_{pooled}}$$

Guidelines for Cohen's d metric describe the effect sizes of 0.2 as small, 0.5 as medium, and 0.8 as large.

Finally, in order to enhance external validity the hypothesis factors were examined separately for the College of Arts and Science (N=931) and the College of Education (N=722).

Before testing hypotheses using analysis of variance, assumptions for this procedure have to be examined. The first assumption of the analysis of variance is that

the outcome scores are independent within treatment groups, as well as independent between treatment groups (Keppel, 1991). This is not only an assumption for the analysis of variance, but a requirement of an experimental design. It is reasonable to assume that this study did not violate the independence assumption because (a) test booklets were randomly assigned to students within a classroom, and (b) students completed the instruments independently and were exposed to the test for about 10 minutes.

The second assumption of the analysis of variance procedure states that individual treatment populations, from which the sample is drawn, are assumed to be randomly drawn and have a normal distribution indicated by the level of skewness (0) and kurtosis (0) (Keppel, 1991). This study examined this assumption for the three leadership outcome variables separately for the samples from the College of Arts and Science and the College of Education, as well as by booklet form and participant gender (Tables 22-33). Results show that all three leadership factors in both colleges had a level of skewness ranging from a minimum of -0.05 to 1.0 , while the level of kurtosis ranged from -0.01 to 2.23 . These ranges indicate that the normality assumption was likely violated. However, effects of leptokurtic or platykurtic distributions on the nominal alpha level (Type I error) are slight (Winer, Brown, & Michels, 1991).

Table 22

Descriptive Statistics for Male Research Participants in the College of Arts and Science for Leader Competence

Form	N	Mean	SD	Skewness	Kurtosis
1	60	3.27	0.78	0.12	-0.14
2	54	3.38	0.85	0.12	-0.59
3	61	3.03	0.80	-0.19	0.14
4	57	3.45	0.66	0.35	-0.19
5	58	3.25	0.77	-0.21	0.59
6	63	3.46	0.88	-0.10	-0.89
7	56	3.14	0.59	-0.81	2.23
8	55	3.58	0.63	0.43	-0.80

Note. Form 1 = Autocratic Female Professor in Personal Condition; Form 2 = Democratic Female Professor in Personal Condition; Form 3=Autocratic Female Professor in Task Condition; Form 4 =Democratic Female Professor in Task Condition; Form 5 = Autocratic Male Professor in Personal Condition; Form6=Democratic Male Professor in Personal Condition; Form7=Autocratic Male Professor in Task Condition; Form 8 = Democratic Male Professor in Task Condition.

Table 23

Descriptive Statistics for Female Research Participants in the College of Arts and Science for Leader Competence

Form	N	Mean	SD	Skewness	Kurtosis
1	58	3.04	0.91	0.11	-0.68
2	54	3.49	0.71	-0.59	-0.07
3	60	3.14	0.55	-0.02	0.17
4	60	3.37	0.71	0.17	-0.01
5	58	3.14	0.83	0.21	-0.24
6	59	3.68	0.71	-0.08	-0.37
7	61	3.06	0.59	0.42	0.02
8	57	3.44	0.77	-0.83	2.03

Note. Form 1 = Autocratic Female Professor in Personal Condition; Form2=Democratic Female Professor in Personal Condition; Form 3 = Autocratic Female Professor in Task Condition; Form 4 = Democratic Female Professor in Task Condition; Form 5 = Autocratic Male Professor in Personal Condition; Form 6 = Democratic Male Professor in Personal Condition; Form 7 = Autocratic Male Professor in Task Condition; Form 8 = Democratic Male Professor in Task Condition.

Table 24

Descriptive Statistics for Male Research Participants in the College of Education for Leader Competence

Form	N	Mean	SD	Skewness	Kurtosis
1	14	3.18	0.54	-0.45	-0.98
2	20	3.28	0.64	0.30	-0.33
3	14	2.80	0.68	-0.69	-0.29
4	20	3.44	0.62	0.43	-0.14
5	20	2.66	0.68	0.07	-0.22
6	16	3.42	0.78	0.18	-0.60
7	19	3.24	0.66	-0.39	2.12
8	17	3.41	0.45	0.33	1.45

Note. Form 1 = Autocratic Female Professor in Personal Condition; Form 2 = Democratic Female Professor in Personal Condition; Form 3 = Autocratic Female Professor in Task Condition; Form 4 = Democratic Female Professor in Task Condition; Form 5 = Autocratic Male Professor in Personal Condition; Form 6 = Democratic Male Professor in Personal Condition; Form 7 = Autocratic Male Professor in Task Condition; Form 8 = Democratic Male Professor in Task Condition.

Table 25

Descriptive Statistics for Female Research Participants in the College of Education for Leader Competence

Form	N	Mean	SD	Skewness	Kurtosis
1	77	3.00	0.73	0.15	-0.67
2	75	3.38	0.77	-0.45	-0.19
3	69	2.88	0.63	0.34	0.35
4	71	3.35	0.63	0.40	0.51
5	72	2.88	0.76	-0.08	-0.52
6	73	3.50	0.91	-0.31	-0.64
7	72	2.91	0.66	0.53	0.77
8	73	3.27	0.78	-0.25	-0.01

Note: Form 1 = Autocratic Female Professor in Personal Condition; Form 2 = Democratic Female Professor in Personal Condition; Form 3 = Autocratic Female Professor in Task Condition; Form 4 = Democratic Female Professor in Task Condition; Form 5 = Autocratic Male Professor in Personal Condition; Form 6 = Democratic Male Professor in Personal Condition; Form 7 = Autocratic Male Professor in Task Condition; Form 8 = Democratic Male Professor in Task Condition.

Table 26

Descriptive Statistics for Male Research Participants in the College of Arts and Science for Leader Masculinity

Form	N	Mean	SD	Skewness	Kurtosis
1	60	3.59	0.65	-0.56	0.01
2	54	3.01	0.59	0.20	-0.32
3	61	4.01	0.56	-0.37	0.35
4	57	3.37	0.68	-0.52	-0.04
5	58	3.35	0.63	0.34	-0.04
6	63	2.80	0.63	0.31	0.01
7	56	4.06	0.56	-0.02	-0.97
8	55	3.16	0.77	0.12	0.69

Note. Form 1 = Autocratic Female Professor in Personal Condition; Form 2 = Democratic Female Professor in Personal Condition; Form 3 = Autocratic Female Professor in Task Condition; Form 4 = Democratic Female Professor in Task Condition; Form 5 = Autocratic Male Professor in Personal Condition; Form 6 = Democratic Male Professor in Personal Condition; Form 7 = Autocratic Male Professor in Task Condition; Form 8 = Democratic Male Professor in Task Condition.

Table 27

Descriptive Statistics for Female Research Participants in the College of Arts and Science for Leader Masculinity

Form	N	Mean	SD	Skewness	Kurtosis
1	58	3.85	0.62	-0.38	0.32
2	54	2.78	0.67	0.60	-0.26
3	60	4.24	0.54	-0.68	0.55
4	60	3.32	0.66	0.38	-0.52
5	58	3.59	0.68	0.11	0.08
6	59	2.76	0.55	0.34	0.45
7	61	4.21	0.62	-0.83	0.45
8	57	3.37	0.73	0.16	-0.42

Note. Form 1 = Autocratic Female Professor in Personal Condition; Form 2 = Democratic Female Professor in Personal Condition; Form 3 = Autocratic Female Professor in Task Condition; Form 4 = Democratic Female Professor in Task Condition; Form 5 = Autocratic Male Professor in Personal Condition; Form 6 = Democratic Male Professor in Personal Condition; Form 7 = Autocratic Male Professor in Task Condition; Form 8 = Democratic Male Professor in Task Condition.

Table 28

Descriptive Statistics for Male Research Participants in the College of Education for Leader Masculinity

Form	N	Mean	SD	Skewness	Kurtosis
1	14	3.78	0.46	1.31	2.31
2	20	2.60	0.40	-0.65	1.91
3	14	4.37	0.62	-0.63	-1.06
4	20	3.41	0.57	-0.04	-1.26
5	20	3.38	0.81	0.14	-0.44
6	16	2.90	0.40	-0.25	-1.00
7	19	4.25	0.46	-0.05	-0.57
8	17	3.23	0.75	0.73	0.21

Note. Form 1 = Autocratic Female Professor in Personal Condition; Form 2 = Democratic Female Professor in Personal Condition; Form 3 = Autocratic Female Professor in Task Condition; Form 4 = Democratic Female Professor in Task Condition; Form 5 = Autocratic Male Professor in Personal Condition; Form 6 = Democratic Male Professor in Personal Condition; Form 7 = Autocratic Male Professor in Task Condition; Form 8 = Democratic Male Professor in Task Condition.

Table 29

Descriptive Statistics for Female Research Participants in the College of Education for Leader Masculinity

Form	N	Mean	SD	Skewness	Kurtosis
1	77	3.83	0.66	-0.63	1.11
2	75	2.71	0.72	0.38	0.08
3	69	4.24	0.60	-0.74	0.12
4	71	3.36	0.67	-0.01	-0.23
5	72	3.76	0.66	-0.30	-0.14
6	73	2.81	0.58	0.16	-0.01
7	72	4.23	0.55	-0.79	0.66
8	73	3.38	0.64	-0.11	-0.86

Note. Form 1 = Autocratic Female Professor in Personal Condition; Form 2 = Democratic Female Professor in Personal Condition; Form 3 = Autocratic Female Professor in Task Condition; Form 4 = Democratic Female Professor in Task Condition; Form 5 = Autocratic Male Professor in Personal Condition; Form 6 = Democratic Male Professor in Personal Condition; Form 7 = Autocratic Male Professor in Task Condition; Form 8 = Democratic Male Professor in Task Condition.

Table 30

Descriptive Statistics for Male Research Participants in the College of Arts and Science for Leader Likeability

Form	N	Mean	SD	Skewness	Kurtosis
1	60	2.58	1.10	0.47	-0.76
2	54	3.29	1.00	-0.23	-0.47
3	61	2.13	0.94	1.13	1.08
4	57	3.16	0.85	0.19	-0.91
5	58	2.74	1.01	0.28	-0.36
6	63	3.58	0.85	-0.46	-0.10
7	56	2.11	0.70	0.28	-0.78
8	55	3.28	0.93	-0.19	-0.48

Note. Form 1 = Autocratic Female Professor in Personal Condition; Form 2 = Democratic Female Professor in Personal Condition; Form 3 = Autocratic Female Professor in Task Condition; Form 4 = Democratic Female Professor in Task Condition; Form 5 = Autocratic Male Professor in Personal Condition; Form 6 = Democratic Male Professor in Personal Condition; Form 7 = Autocratic Male Professor in Task Condition; Form 8 = Democratic Male Professor in Task Condition.

Table 31

Descriptive Statistics for Female Research Participants in the College of Arts and Science for Leader Likeability

Form	N	Mean	SD	Skewness	Kurtosis
1	58	2.34	1.12	0.59	-0.57
2	54	3.61	0.87	-0.85	0.39
3	60	1.94	0.66	0.78	0.27
4	60	2.95	0.89	0.02	-0.04
5	58	2.58	0.89	0.63	0.22
6	59	3.97	0.88	-0.58	-0.50
7	61	1.90	0.63	0.45	-0.91
8	57	3.14	0.99	-0.37	-0.29

Note. Form 1 = Autocratic Female Professor in Personal Condition; Form 2 = Democratic Female Professor in Personal Condition; Form 3 = Autocratic Female Professor in Task Condition; Form 4 = Democratic Female Professor in Task Condition; Form 5 = Autocratic Male Professor in Personal Condition; Form 6 = Democratic Male Professor in Personal Condition; Form 7 = Autocratic Male Professor in Task Condition; Form 8 = Democratic Male Professor in Task Condition

Table 32

Descriptive Statistics for Male Research Participants in the College of Education for Leader Likeability

Form	N	Mean	SD	Skewness	Kurtosis
1	14	2.16	0.60	1.14	1.72
2	20	3.58	0.68	-0.31	0.06
3	14	1.73	0.74	0.71	-0.20
4	20	3.23	0.70	-0.42	-0.90
5	20	2.06	0.95	1.03	0.50
6	16	3.66	1.00	-0.65	-0.27
7	19	2.26	0.88	0.59	0.25
8	17	3.13	0.76	-0.99	1.22

Note: Form 1 = Autocratic Female Professor in Personal Condition; Form 2 = Democratic Female Professor in Personal Condition; Form 3 = Autocratic Female Professor in Task Condition; Form 4 = Democratic Female Professor in Task Condition; Form 5 = Autocratic Male Professor in Personal Condition; Form 6 = Democratic Male Professor in Personal Condition; Form 7 = Autocratic Male Professor in Task Condition; Form 8 = Democratic Male Professor in Task Condition

Table 33

Descriptive Statistics for Female Research Participants in the College of Education for Leader Likeability

Form	N	Mean	SD	Skewness	Kurtosis
1	77	2.11	0.77	0.81	0.25
2	75	3.56	0.87	-0.52	-0.02
3	69	1.83	0.57	0.42	-0.20
4	71	30.1	0.95	0.01	-0.50
5	72	2.16	0.85	0.21	-0.78
6	73	3.55	0.97	-0.23	-0.74
7	72	1.88	0.66	0.89	1.40
8	73	2.84	0.86	0.06	-0.68

Note: Form 1 = Autocratic Female Professor in Personal Condition; Form 2 = Democratic Female Professor in Personal Condition; Form 3 = Autocratic Female Professor in Task Condition; Form 4 = Democratic Female Professor in Task Condition; Form 5 = Autocratic Male Professor in Personal Condition; Form 6 = Democratic Male Professor in Personal Condition; Form 7 = Autocratic Male Professor in Task Condition; Form 8 = Democratic Male Professor in Task Condition

The third assumption of the analysis of variance procedure is homogeneity of variance. Violations of this assumption could impact the F test in the positive direction, especially if heterogeneity is associated with one deviant group, or when treatment groups have unequal sample sizes (Keppel, 1991). Heterogeneous variances have a very slight effect on nominal alpha, and no theoretical power value exists with heterogeneous variances (Winer et al., 1991). This study used the modified Levene's test to check the three leadership factors for homogeneity of variances by booklet form and participant gender, separately for the College of Arts and Science and Education. The homogeneity assumption held for the masculinity factor in both colleges, but was violated for the competence and likeability factors (Table 34).

Table 34

Modified Levene's Test Summary Statistics for the College of Arts and Science and Education by Booklet Form and Participant Gender

Study Sample	DF	MS	F	p
College of Arts & Science N=931				
Competence	15	1.42	2.41	<.01
Masculinity/Femininity	15	0.40	1.21	>.05
Likeability	15	3.38	2.87	<.01
College of Education N=722				
Competence	15	0.98	1.89	<.05
Masculinity/Femininity	15	0.50	1.55	>.05
Likeability	15	1.89	2.40	<.01

Hypothesis 1

Democratic professors will be rated as significantly more competent than autocratic professors

Tables 35 and 36 show the ANOVA source table for both colleges with regard to the competence factor. The prediction of the hypothesis that democratic professors would be rated more competent than autocratic professors was supported by the results of this study for the College of Arts and Science, $F(1, 915)=52.9, p<.01$, and for the College of Education, $F(1, 706)=39.7, p<.01$. Research participants from both colleges rated the democratic professor as more competent (Arts and Science: $M=3.5, SD=0.74$; Education: $M=3.4, SD=0.7$) than the autocratic professor (Arts and Science: $M=3.1, SD=0.7$; Education: $M=2.9, SD=0.7$). Estimated effect sizes for this competence main effect were medium for both colleges (Arts and Science=0.24, Education=0.23). Tables 35 and 36 list the ANOVA source information for each college.

Hypothesis 2

Autocratic professors will be rated as significantly more masculine than democratic professors.

This hypothesis was supported in this experiment in both the sample from the College of Arts and Science, $F(1, 915)=352.2, p<.01$, and the sample from the College of Education, $F(1, 706)=241.0, p<.01$. The calculated effect sizes for this main effect were large for both colleges ($f=0.60$) (see ANOVA source Tables 37 and 38). Research participants for both colleges rated the autocratic professors as more masculine (Arts and Science: $M=3.9, SD=0.7$; Education: $M=4.0, SD=0.7$), than democratic professors (Arts and Science: $M=3.1, SD=0.7$; Education: $M=3.1, SD=0.7$).

Table 35

ANOVA Source Table for the College of Arts and Science for Competence

Source	DF	MS	F	<i>f</i>
Professor Gender (PG)	1	1.4	2.5	
Professor Behavior (PB)	1	29.2	52.9**	0.24
PG x PB	1	0.3	0.7	
Situation (S)	1	1.0	1.8	
PG x S	1	0.1	0.1	
PB x S	1	0.2	0.3	
PG x PB x S	1	0.0	0.0	
Participant Gender (PaG)	1	0.1	0.2	
PG x PaG	1	0.0	0.0	
PB x PaG	1	0.5	0.8	
PG x PB x PaG	1	0.1	0.2	
S x PaG	1	0.1	0.2	
PG x S x PaG	1	0.8	1.4	
PB x S x PaG	1	3.3	6.0*	0.08
PG x PB x S x PaG	1	0.1	0.2	
Error	915	0.5		

N=931; Note: * $p < .05$, ** $p < .01$.

Table 36

ANOVA Source Table for the College of Education for Competence

Source	DF	MS	F	<i>f</i>
Professor Gender (PG)	1	0.0	0.0	
Professor Behavior (PB)	1	21.0	39.7**	0.23
PG x PB	1	0.2	0.3	
Situation (S)	1	0.0	0.0	
PG x S	1	0.9	1.8	
PB x S	1	0.1	0.1	
PG x PB x S	1	4.0	7.4**	0.10
Participant Gender (PaG)	1	0.1	0.2	
PG x PaG	1	0.0	0.0	
PB x PaG	1	0.1	0.1	
PG x PB x PaG	1	0.0	0.0	
S x PaG	1	0.8	1.6	
PG x S x PaG	1	1.2	2.3	
PB x S x PaG	1	0.0	0.1	
PG x PB x S x PaG	1	1.0	1.9	
Error	706	0.5		

N=721; Note: ** $I < .01$.

Table 37

ANOVA Source Table for the College of Arts and Science for Masculinity/Femininity

Source	DF	MS	F	<i>f</i>
Professor Gender (PG)	1	2.8	6.9**	0.09
Professor Behavior (PB)	1	144.2	352.2**	0.60
PG x PB	1	0.0	0.1	
Situation (S)	1	59.0	144.2**	0.40
PG x S	1	1.3	3.2	
PB x S	1	0.2	0.6	
PG x PB x S	1	0.8	2.0	
Participant Gender (PaG)	1	2.1	5.2*	0.07
PG x PaG	1	0.5	1.2	
PB x PaG	1	3.6	8.8**	0.10
PG x PB x PaG	1	1.1	2.6	
S x PaG	1	0.3	0.8	
PG x S x PaG	1	0.0	0.0	
PB x S x PaG	1	1.0	2.6	
PG x PB x S x PaG	1	0.1	0.2	
Error	915	0.4		

N=931; Note: * $p < .05$, ** $p < .01$.

Table 38

ANOVA Source Table for the College of Education for Masculinity/Femininity

Source	DF	MS	F	<i>f</i>
Professor Gender (PG)	1	0.2	0.5	
Professor Behavior (PB)	1	96.0	241.0**	0.60
PG x PB	1	1.2	3.1	
Situation (S)	1	38.4	96.4**	0.40
PG x S	1	0.1	0.3	
PB x S	1	0.0	0.0	
PG x PB x S	1	1.4	3.6	
Participant Gender (PaG)	1	0.3	0.8	
PG x PaG	1	0.3	0.8	
PB x PaG	1	0.0	0.1	
PG x PB x PaG	1	0.4	0.9	
S x PaG	1	0.4	1.0	
PG x S x PaG	1	0.1	0.1	
PB x S x PaG	1	0.7	1.8	
PG x PB x S x PaG	1	0.6	1.6	
Error	706	0.4		

N=722; Note: ** $p < .01$.

Hypothesis 3

Democratic professors will be rated as significantly more likeable than autocratic professors.

This hypothesis was supported in both college samples (see ANOVA source Tables 39 and 40). A statistically significant main effect was found for professor leader style for the College of Arts and Science sample, $F(1, 915)=333.7, p < .01$, and for the College of Education sample, $F(1, 706)=274.2, p < .01$. The calculated effect sizes for this main effect were equally large in both colleges ($f=0.60$). Research participants from both colleges rated the democratic professor as more likeable (College of Arts and Science:

M=3.4, SD=0.9; College of Education: M=3.3, SD=0.9), than the autocratic professor (College of Arts and Science: M=2.3, SD=0.9; College of Education: M=2.0, SD=0.8).

Table 39

ANOVA Source Table for the College of Arts and Science for Likeability

Source	DF	MS	F	<i>f</i>
Professor Gender (PG)	1	6.2	7.6**	0.09
Professor Behavior (PB)	1	273.0	333.7**	0.60
PG x PB	1	1.5	1.4	
Situation (S)	1	60.8	74.3**	0.30
PG x S	1	2.2	2.7	
PB x S	1	0.3	0.3	
PG x PB x S	1	0.1	0.0	
Participant Gender (PaG)	1	0.7	0.9	
PG x PaG	1	0.1	0.1	
PB x PaG	1	4.4	5.3*	0.08
PG x PB x PaG	1	0.0	0.0	
S x PaG	1	3.6	4.4*	0.07
PG x S x PaG	1	0.0	0.0	
PB x S x PaG	1	4.1	5.0*	0.07
PG x PB x S x PaG	1	0.1	0.1	
Error	915	0.7		

N=930; Note: * p<.05, ** p<.01.

Table 40

ANOVA Source Table for the College of Education for Likeability

Source	DF	MS	F	<i>f</i>
Professor Gender (PG)	1	0.2	0.3	
Professor Behavior (PB)	1	186.0	274.2**	0.60
PG x PB	1	0.9	1.3	
Situation (S)	1	15.0	22.0**	0.20
PG x S	1	0.1	0.2	
PB x S	1	3.2	4.7*	0.08
PG x PB x S	1	1.7	2.5	
Participant Gender (PaG)	1	1.4	2.0	
PG x PaG	1	0.4	0.6	
PB x PaG	1	0.3	0.4	
PG x PB x PaG	1	0.0	0.1	
S x PaG	1	0.9	1.3	
PG x S x PaG	1	0.6	0.9	
PB x S x PaG	1	0.0	0.0	
PG x PB x S x PaG	1	0.7	1.0	
Error	706	0.7		

N=721; Note: * $p < .05$, ** $p < .01$.

Hypothesis 4

There will be an interaction effect between participant gender, professor gender, professor leader style, and the type of situation on the ratings of leader competence.

Specifically, male research participants were hypothesized to perceive autocratic female professors to be significantly less competent than autocratic male professors.

Further, this difference was hypothesized to be most pronounced in the personal condition such that competence ratings for this condition (autocratic female professor in personal situation) were hypothesized to be the lowest compared to all other conditions. Female research participants were not influenced by professor gender in their perception of leader competence.

This hypothesis was not supported by the results from the College of Arts and Science, $F(1, 915)=0.19, p>.05$, and the College of Education, $F(1, 706)=1.90, p>.05$. The effect size estimates for these results were very small ($f=0.01$ for the College of Arts and Science and $f=0.05$ for the College of Education).

Hypothesis 5

There will be an interaction effect between participant gender, professor gender, professor leader style, and the type of situation on the ratings of leader likeability.

Specifically, male research participants were hypothesized to perceive autocratic female professors to be significantly less likeable than autocratic male professors. Further, this difference was hypothesized to be most pronounced in the personal condition such that likeability ratings for this condition (autocratic female professor in personal situation) were hypothesized to be the lowest compared to all other conditions. Female research participants were hypothesized to be influenced by professor gender in their perception of leader likeability.

This hypothesis was not supported by the results from the College of Arts and Science, $F(1, 915)=.12, p>.05$, and the College of Education, $F(1, 706)=.98, p>.05$. The effect size estimates for these results were very small ($f=0.01$ for the College of Arts and Science and $f=0.03$ for the College of Education).

Additional Findings for the Competence Factor

Interaction effects. A statistically significant three way interaction effect between professor leader style, type of situation, and participant gender was found for the College of Arts and Science sample, $F(1, 915)=5.97, p<.02$. Calculated effect size for this three way interaction effect was small ($f=0.08$). Figures 6 and 7 depict this interaction effect

separately for male and female research participants. Specifically, male and female participants did not differ in their competence ratings for the autocratic professor/ task condition (male participants: $M=3.1$, $SD=0.7$; female participants: $M=3.1$, $SD=0.6$). However, male participants rated professors in the autocratic personal condition as more competent ($M=3.3$, $SD=0.8$) than female participants ($M=3.1$, $SD=0.9$). The effect size was small ($d=0.24$). This trend was reversed in the democratic condition. That is, male participants rated professors in the democratic/task condition more competent ($M=3.6$, $SD=0.6$) than female participants ($M=3.5$, $SD=0.7$). This effect was small ($d=0.15$). However, male participants rated the democratic professor/personal condition less competent ($M=3.4$, $SD=0.8$) than female participant ratings for that condition ($M=3.6$, $SD=0.7$). The effect size estimate was small for this effect ($d=0.27$). This particular three way interaction effect was not statistically significant in the College of Education sample.

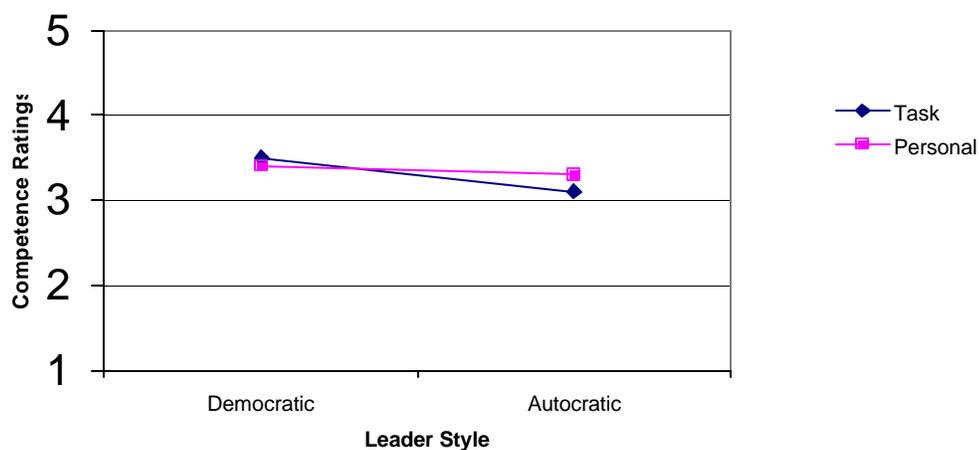


Figure 6 . Competence Ratings for the College of Arts Science: 3-Way Interaction Effects for Professor Leader Style, Situation, and Rater Gender (Male Participants)

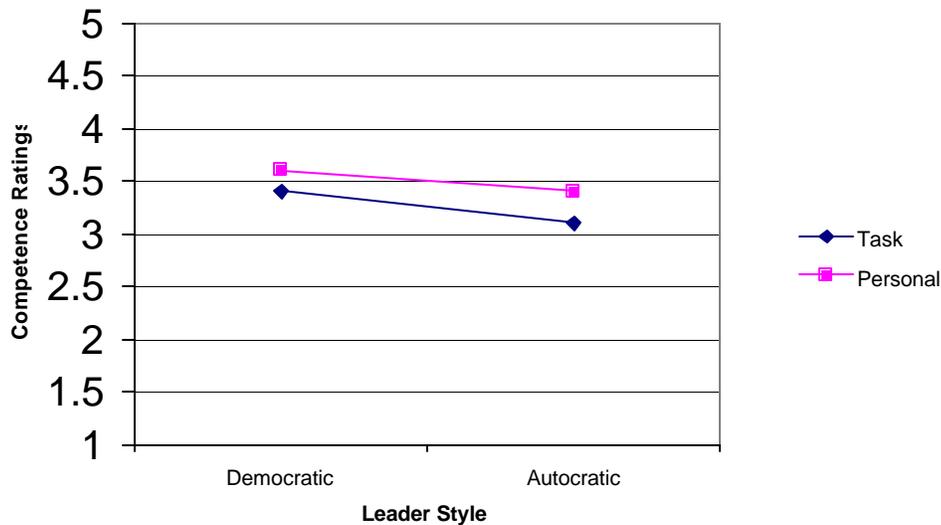


Figure 7. Competence Ratings for the College of Arts & Science: 3-way Interaction Between Professor Leader Style, Situation, and Rater Gender (Females)

Instead, results from the College of Education sample revealed a statistically significant three way interaction effect between professor gender, professor leader style, and type of situation, $F(1, 706)=7.41, p<.01$ (Figures 8 and 9). The effect size for this interaction effect was small ($f=0.10$). Research participants in the College of Education rated autocratic male professors more competent than female autocratic professors in the task condition (Male professor: $M=3.0, SD=0.7$; Female professor: $M=2.9, SD=0.6, d=0.15$) while the competence ratings were reversed for the autocratic/personal condition (Male professor: $M=2.8, SD=0.7$; Female professor: $M=3.0, SD=0.7, d=0.29$). This trend was reversed in the democratic condition. Specifically, research participants rated democratic male professor in the task condition less competent than the female counterpart (Male

professor: $M=3.3$, $SD=0.7$; Female professor: $M=3.4$, $SD=0.6$; $d=0.15$), while the democratic male professor in the personal condition was rated more competent than his female counterpart (Male professor: $M=3.5$, $SD=0.9$; Female professor: $M=3.4$, $SD=0.7$; $d=0.13$). Figures 8 and 9 illustrate this relationship.

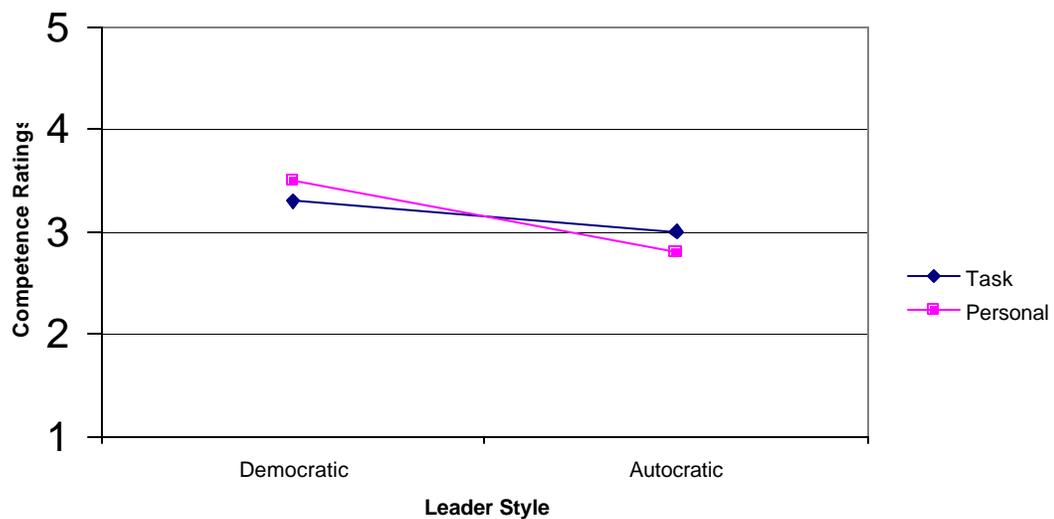


Figure 8. Competence Ratings for the College of Education: 3-way Interaction between Prof Gender, Prof Leader Style, and Situation (Ratings for Male Professors)

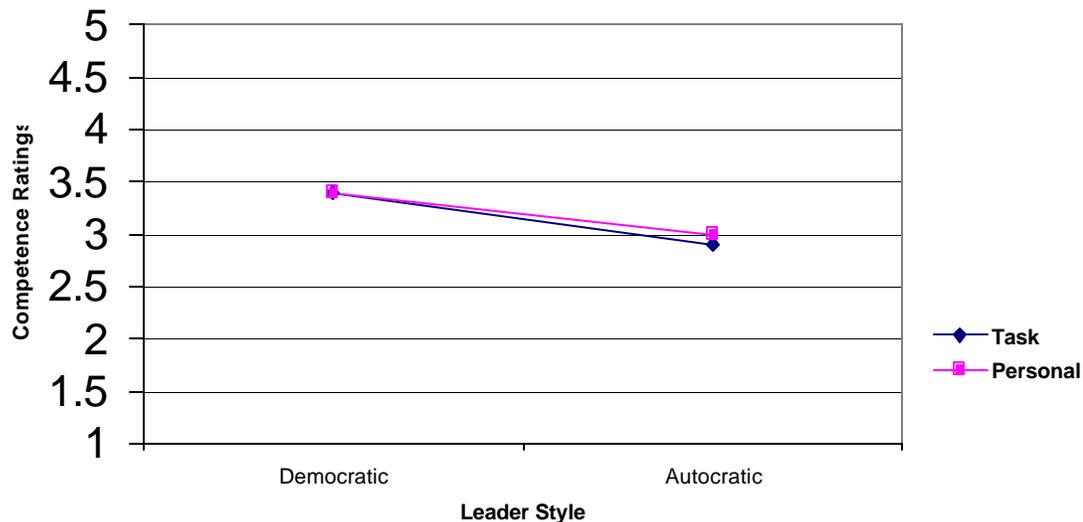


Figure 9. Competence Ratings for the College of Education: 3-way Interaction between Professor Gender, Professor Leader Style, and Situation (Ratings for Female Professors)

Additional Findings for the Masculinity Factor

A statistically significant main effect was present for the situation condition in both the College of Arts and Science, $F(1, 915)=144.2, p<.01$, and the College of Education, $F(1, 706)=96.4, p<.01$. The effect sizes for this main effect were equally large for both colleges ($f=0.40$). Research participants rated the professor as more masculine in the task condition (Arts and Science: $M=3.7, SD=0.8$; Education: $M=3.8, SD=0.8$) than in the personal condition (Arts and Science: $M=3.1, SD=0.7$; Education: $M=3.3, SD=0.8$).

Other findings suggested that there were some statistically significant differences between the two samples on the masculinity/femininity ratings. Results of this study found a statistically significant main effect for professor gender only in the College of

Arts and Science sample, $F(1, 915)=6.9, p<.01$. Calculated effect size estimate for this main effect was small ($f=0.09$). Interestingly, research participants from the College of Arts and Science perceived female professors to be more masculine ($M=3.6, SD=0.8$) than their male counterpart ($M=3.4, SD=0.8$). This main effect was not statistically significant in the College of Education sample.

Further, the main effect for participant gender was statistically significant only for the sample from the College of Arts and Science, $F(1, 915)=5.3, p<.03$. The calculated effect size for this main effect was small ($f=0.07$). Female research participants from the College of Arts and Science perceived professors in the scenarios as slightly more masculine ($M=3.5, SD=0.8$) than male research participants ($M=3.4, SD=0.8$).

Finally, a statistically significant interaction effect between professor leader style and participant gender was found in the Arts and Science sample, $F(1, 915)=8.8, p<.01$ with a small effect size ($f=0.10$) (Figure 10). This statistically significant interaction effect was not present in the College of Education sample. Male participants from the College of Arts and Science perceived autocratic professors as less masculine ($M=3.7, SD=0.7$) than female participants ($M=4.0, SD=0.7$). The effect size for this result was small to medium ($d=0.43$). Both male and female research participants did not differ in their masculinity ratings for the democratic professor (Males: $M=3.1, SD=0.7$; Females: $M=3.1, SD=0.7$). Figure 10 illustrates this interaction effect.

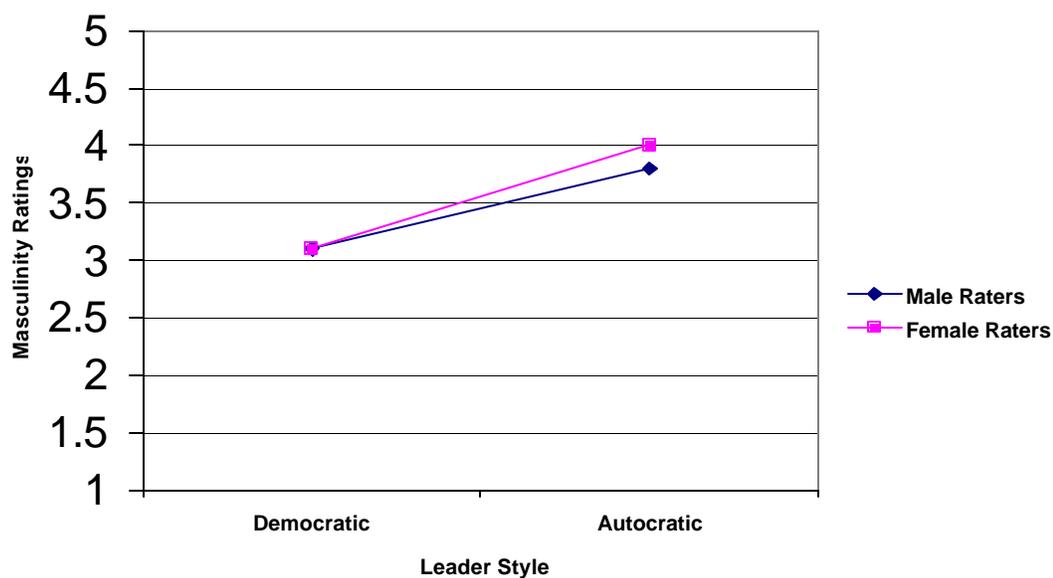


Figure 10. Masculinity Ratings for the College of Arts & Science: 2-way Interaction Effect between Professor Leader Style and Participant Gender

Additional Findings for the Likeability Factor

A statistically significant main effect for the situation condition was present in both the College of Arts and Science, $F(1, 915)=74.3, p<.01$, and the College of Education, $F(1, 706)=22.0, p<.01$. The calculated effect sizes for this main effect were medium in both colleges (Arts and Science=0.30, Education=0.20). However, further results indicated different statistically significant trends for the two colleges.

A statistically significant main effect for professor gender was obtained for the College of Arts and Science, $F(1, 915)=7.6, p<.01$. This effect was not present in the

College of Education sample. Male professors were perceived more likeable ($M=2.9$, $SD=1.0$) than female professors ($M=2.7$, $SD=1.0$).

Further, a two way statistically significant interaction effect between type of situation and participant gender was present in the College of Arts and Science, $F(1, 915)=4.4$, $p<.04$. The effect size calculated for this two way interaction effect was small ($f=0.07$). Male participants perceived professors in the task condition as more likeable ($M=2.7$, $SD=1.0$) than female participants ($M=2.5$, $SD=1.0$) ($d=0.20$), while both male and female participants were similar in their likeability rating for the professor in the personal condition ($M=3.1$, $SD=1.0$ for both genders). Figure 11 shows this relationship.

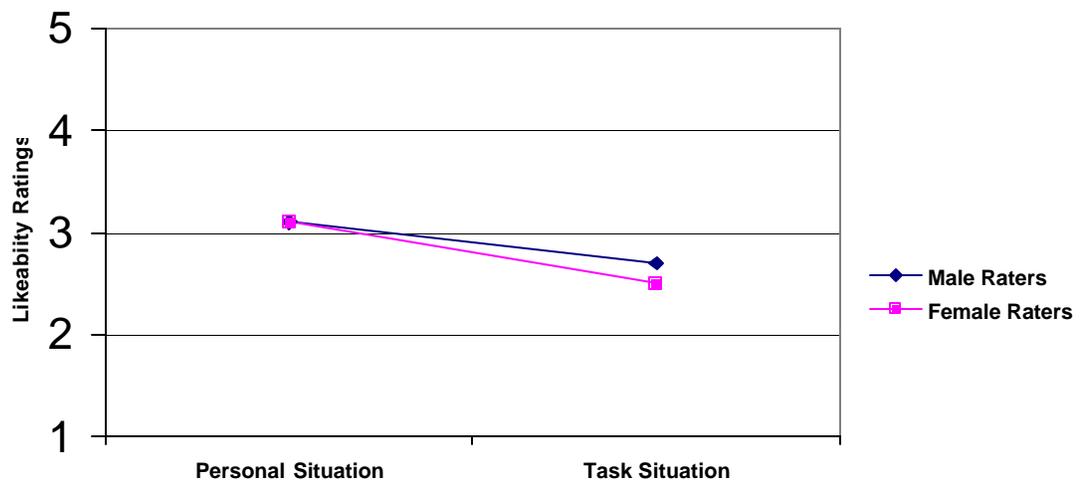


Figure 11. Likeability Ratings for the College of Arts & Science: 2-way Interaction Effect between Type of Situation and Rater Gender

A two way statistically significant interaction effect between professor leadership style and participant gender was present in the College of Arts and Science, $F(1, 915)=5.3, p<.03$. The effect size calculated for this two way interaction effect was small ($f=0.08$). Male participants perceived the autocratic professor as more likeable ($M=2.4, SD=1.00$) than female participants ($M=2.2, SD=0.9$). This result had a small effect size ($d=0.21$). This trend was reversed in the democratic condition, where male participants perceived the democratic professor as less likeable ($M=3.3, SD=0.9$) than female participants ($M=3.4, SD=1.0$) ($d=0.11$). Figure 12 is a visual presentation of this relationship.

A three way statistically significant interaction effect between professor leader style, type of situation, and participant gender was found for the College of Arts and Science, $F(1, 915)=5.0, p<.03$. The calculated effect size for this three way interaction effect was small ($f=0.07$; see Figures 13 and 14). While both male and female participants perceived the autocratic professor as the least likeable, female participants were consistently lower in their likeability ratings for the autocratic/task condition (male participants: $M=2.1, SD=0.8$; female participants: $M=1.9, SD=0.7, d=0.27$), and autocratic/personal condition (male participants: $M=2.7, SD=1.0$; female participants: $M=2.5, SD=1.0, d=0.20$) than male participants. This trend was in part reversed in the democratic condition, where male participants perceived the democratic professor in the task situation as more likeable ($M=3.2, SD=0.9$) than female research participants ($M=3.0, SD=0.9$). This effect was small ($d=0.20$). However, female research participants rated democratic professors in the personal condition more likeable ($M=3.8, SD=0.9$) than male research participants ($M=3.5, SD=0.9$). This effect was small ($d=0.33$).

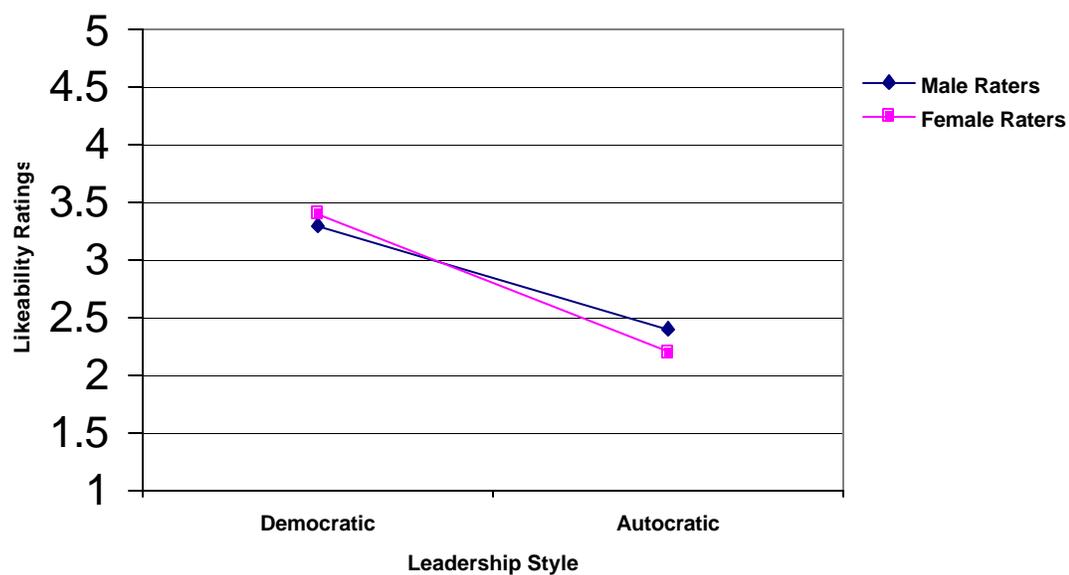


Figure 12 . Likeability Ratings for the College of Arts & Science: 2-way Interaction between Prof Leader Style and Rater Gender

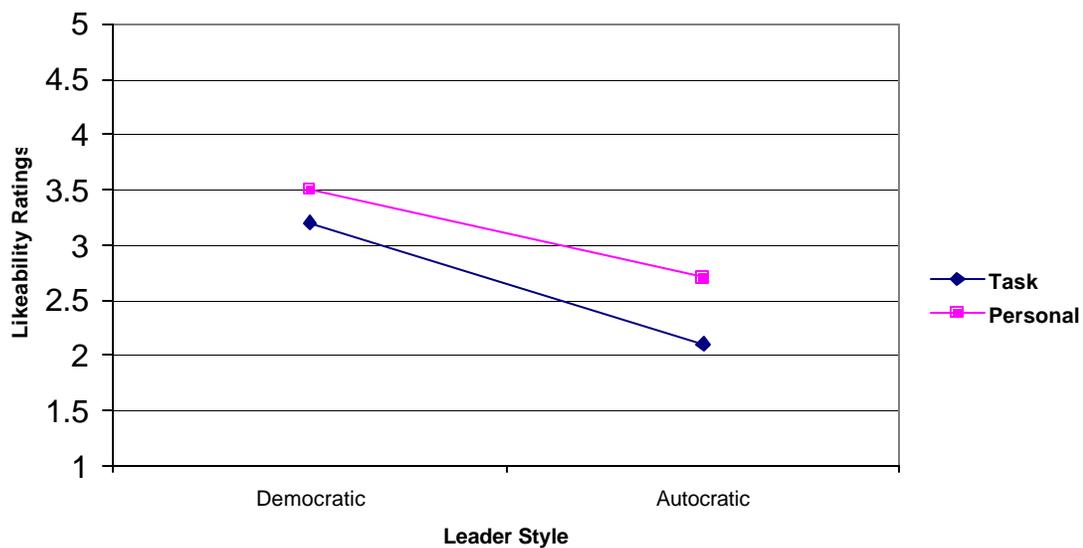


Figure 13. Likeability Ratings for the College of Arts & Science: 3-way Interaction Effect between Professor Leader Style, Situation, and Rater Gender (Male Raters)

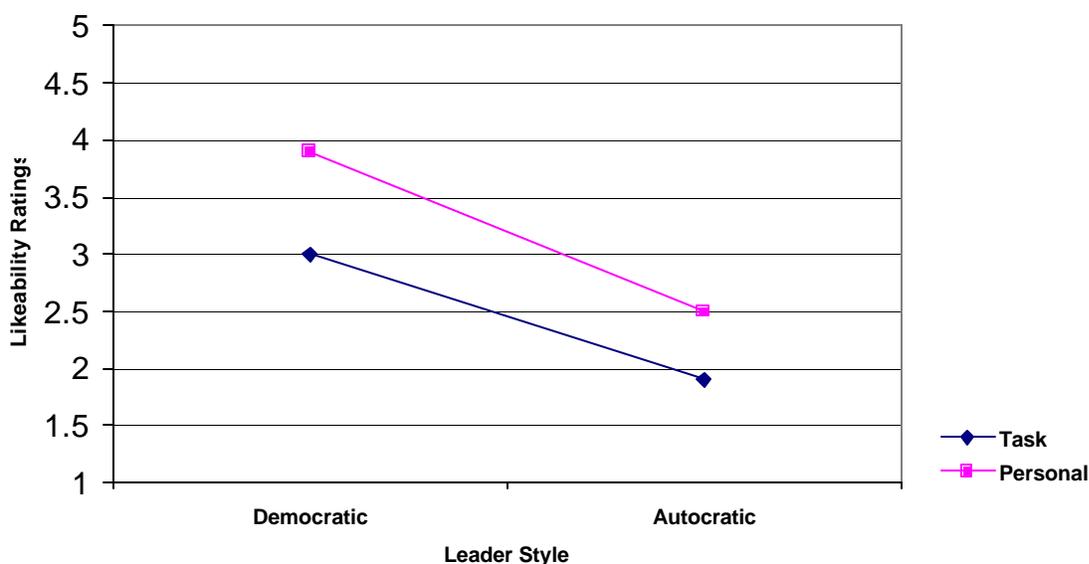


Figure 14 . Likeability Ratings for the College of Arts & Science: 3-way Interaction Effect between Prof Leader Style, Situation, and Rater Gender (Female Raters)

Finally, a two way statistically significant interaction effect for likeability between professor leader style and type of situation was found for the College of Education, $F(1, 706=4.7)$, $p<.04$. The calculated effect size for this interaction effect was small ($f=0.08$). Research participants rated autocratic professors in the task condition less likeable ($M=1.9$, $SD=0.7$) than autocratic professors in the personal condition ($M=2.1$, $SD=0.8$). This effect was small ($d=0.26$). This trend was more pronounced in the democratic condition, that is, democratic professors in the task condition were perceived less likeable ($M=3.0$, $SD=0.9$), than democratic professors in the personal condition

($M=3.6$, $SD=0.9$). The calculated effect size for this result was medium ($d=0.67$). Figure 15 illustrates this interaction effect.

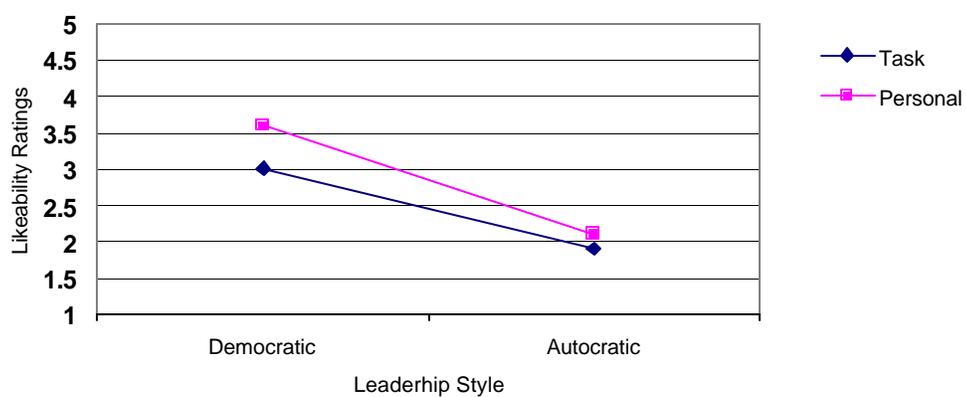


Figure 15. Likeability Ratings for the College of Education: 2-way Interaction Effect between Professor Leader Style and Type of Situation

Chapter V

Discussion

This chapter is divided into four parts. The first section summarizes the background and the purpose of the experiment. The second section discusses the psychometric properties of the leadership measures and the implications of these measures for interpreting the results of the hypothesis testing. The third section discusses the results from this experiment and how they relate to previous findings. All findings of this experiment will be discussed in light of several major gender role theories. Section four lists the limitations of this study, and discusses recommendations for future research.

Section I Summary

The purpose of this experimental study was to (a) examine the psychometric properties of the factors underlying a set of 24 leadership items and (b) test five hypotheses that were derived from two popular gender role theories as well as from empirical findings of previous studies. This study used an academic context in order to update academic leadership research.

An experimental design was created using written vignettes depicting a professor student interaction. Experimental manipulations resulting in eight booklet versions included professor gender, professor leadership style (democratic/autocratic), and the type of situation (task/personal) in which the interaction took place. In addition, the design included participant gender as a variable because female participants were

predicted to be less influenced by leader gender than male participants (as tested in hypotheses 4 and 5).

Samples were selected from the College of Arts and Science and from the College of Education in an effort to strengthen population validity. A total of 1654 undergraduate student from the College of Arts and Science (N=932) and the College of Education (N=722) volunteered for this experiment during the Fall Semester of 2001.

Section II Psychometric Properties of the Leadership Construct

One factor that has hindered research on gender differences in leadership positions is the lack of psychometrically sound instruments that measure the multidimensional nature of leadership. In view of limitations of previous research this study began with the development of a multidimensional measure of leadership. This measure was developed under the assumption that in order to advance gender role theory as it relates to leadership, sound measurement of leadership is required. In turn, good theory guides researchers in selecting important variables noted in previous research (Alumbaugh, 1995).

As an empirical strategy, leadership outcome measures that were used in previous research were selected and sorted into four leadership outcome dimensions (i.e., competence, masculinity/femininity, likeability, and professional traits). Items reflecting these dimensions were then entered into a series of exploratory and confirmatory factor analyses in order to evaluate the psychometric properties of these measures. Results of the principal axis factor analysis provided evidence of two strong factors representing competence and masculinity/femininity.

Results for the masculinity/femininity dimension revealed that all four originally proposed items retained their internal cohesiveness (moderately high internal consistency), and formed an independent factor. This factor was important because it links gender roles in Western society with general beliefs about leadership qualities (i.e., masculine attributes are associated with leadership roles, and feminine attributes are associated with subordinate roles).

Leader competence is an important leader outcome measure, and has been defined by organizational psychologists (e.g., Hunt, 1991) to be the ability of a leader to facilitate organizational goals. Eagly et al. (1995) in their meta-analytic review found that even though many of the studies evaluated leader competence, their methods of assessing this outcome varied. The current study attempted to create a leader competence factor that incorporated five competence items used by previous studies (competent/incompetent, effective/ineffective, qualified/not qualified, capable/not capable, influential/not influential). Results showed that all five items clustered with the competence factor. In addition though, three items from the professional traits (hardworking/lazy, responsible/irresponsible, trustworthy/untrustworthy) loaded on the competence factor. When pondering the qualities of what makes a leader competent, it seemed conceptually reasonable that the qualities of hard working, sense of responsibility, and level of trustworthiness were important leadership behaviors.

Although professional traits were originally proposed as a separate dimension of leadership, results from this study showed that this dimension did not hold up as an internally consistent and independent factor. This was not unexpected because the originally proposed professional traits consisted of items that had been used in previous

studies and were considered to be miscellaneous. Even though this study dropped the professional traits dimension, additional research is needed to evaluate the generalizability of this finding. For example, this study encountered difficulties with items that had a tendency to crossload on other factors. Generating a set of new items that cross load minimally, may yield improved results.

Finally, the likeability factor posed a challenge not uncommon when using analytic procedures with verbally similar yet theoretically different themes (e.g., Tracy, 1995). Even though the likeability items had high internal consistency, the likeability construct did not cluster into an independent factor. Instead, selected items tended to merge with the competence factor (1 item), and the masculinity/femininity factor (2 items). Even though the likeability items (critical/tolerant, popular/unpopular and likeable/not likeable) were theoretically defined to be a distinct dimension from the masculinity/femininity factor (timid/forceful, soft/tough, aggressive/not aggressive, dominant/submissive), results of the factor analysis provided evidence that research participants did not make that distinction. Instead, two of the likeability items loaded highly onto the masculinity/femininity factor, and one item loaded on the competence factor. It is possible that in spite of pre-testing for item quality, the descriptors for the likeability construct may have been too ambiguous to distinguish them from the other two dimensions. Even though results from this study could not clearly establish likeability as a separate factor, items from the likeability cluster were viewed as potentially adding important information not provided by the competence or masculinity/femininity factors.

To summarize, empirical evidence provided partial support for the multidimensional leadership construct. Results of this study provided evidence for at least a two factor structure (competence and masculinity/femininity), and reasonable evidence to support the likeability dimension. As a result these three factors were used in the testing of the hypotheses.

Section III Hypotheses Testing

This study tested a set of five hypotheses separately for a sample from the College of Arts and Science (N=931) and for a sample from the College of Education (N=722).

Hypothesis 1: Democratic professors will be rated as significantly more competent than autocratic professors.

Results of this study provided strong support for the notion that leadership style has an impact on the perceived competence of the leader. Research participants from both colleges viewed the democratic professor to be more competent than the autocratic professor. Medium effect sizes found in both colleges support this difference in the perception of competence between autocratic and democratic professors (Stevens, 1990). This result echoes the results of previous studies that found that in an organizational setting, democratic managers were evaluated much higher on performance and leadership abilities than autocratic managers (e.g., Luthar, 1996). The higher rating tendency of the democratic professor over the autocratic professor suggests that students within an academic setting show similar rating tendencies for a democratic leader as employees who work in business organizations. Interestingly, Luthar (1996) pointed out that this preference should be expected since North American social values have a democratic basis.

Hypothesis 2: Autocratic professors will be rated as significantly more masculine than democratic professors.

The results of this study found that research participants from both colleges perceived autocratic professors to be significantly more masculine than democratic professors. The large effect sizes indicate that research participants very clearly perceived the democratic professor to be more feminine, and the autocratic professor to be more masculine. This result seems to contradict culturally held beliefs that competent leaders display a high degree of masculine traits. The obtained effect size for this finding was large, which is rarely observed in social sciences research (Stevens, 1990), and indicates that participants clearly perceived democratic professors to be much less masculine than autocratic professors regardless of professor gender. How can this result be interpreted? First, most of the gender role theories maintain that culturally defined gender roles influence people's expectation about what makes a leader competent. The concern this study raises is whether gender roles as defined by those theories actually are still an accurate reflection of currently held gender beliefs, especially since most of the gender role theories were proposed during the 70's and 80's. On the other hand, results may reflect Kanter's (1977) notion that negative evaluation of women in leadership positions occur because of the differential distribution of men and women in managerial positions, that is, women are more visible because of their minority group membership. This study was conducted within an academic environment believed to have an approximate equal distribution of male and female professors, which could explain why the results did not show a gender difference that was hypothesized using gender role spill-over theory. The results further reflect Phillip et al.'s theory (1982) that leaders are evaluated by specific

beliefs and expectations about leader behaviors that override expectations about gender role behaviors. As already mentioned, it would be interesting to examine if cultural gender role beliefs have changed compared to the beliefs generally held about 30 years ago, and how these changes may impact gender role theories and their implications. Finally, replicating this study within an organizational environment, may help develop a more precise understanding of how and when gender matters in leader evaluation.

Hypothesis 3: Democratic professors will be rated as significantly more likeable than autocratic professors.

As expected, research participants from both colleges rated the democratic professor to be much more likeable than the autocratic professor irrespective of gender. The magnitude of this effect size was large for both samples. Past research (e.g., Bartol & Butterfield, 1976) found that female managers were evaluated to be less effective, as well as less likeable if they displayed an autocratic leadership style compared to democratic female managers. Bartol et al.'s findings support gender role congruency theory that punishes gender role incongruent behaviors of leaders. However, results of this study did not find evidence in support of the gender congruence theory. Instead, research participants rated both male and female democratic professors as more likeable than their autocratic counterparts. This study showed that research participants clearly perceived the democratic professor as much more likeable (large effect size), more competent (medium effect size), and as more feminine than their autocratic counterparts. The perception of a competent leader in this study was associated with possessing a higher degree of femininity, regardless of leader gender. Results indicated that leader style rather than leader gender was crucial in the evaluation of the leadership qualities. This finding is

consistent again with Phillips et al. (1982) who maintain that leader behavior of managers becomes more important than leader gender in the evaluation of leaders.

Hypothesis 4: There will be an interaction effect between participant gender, professor gender, professor leadership style and type of situation on the ratings of leader competence

This hypothesized four-way interaction effect was complex in that it combined predictions from gender role spill-over and gender congruency theories, while incorporating previous empirical evidence that men were more negative than women when evaluating a female leader. Specifically, it was predicted that male professors would not be subjected to the level of scrutiny when described as behaving inconsistently with their gender role expectation due to the societal association between male gender roles and leadership (gender spill-over theory). Predictions were very different for women (gender congruency theory), in that female professors who behave in a manner that is perceived to be inconsistent with their gender role were predicted to be rated less competent than gender congruent female professors. This devaluation was expected to happen with male raters (past research). Results of this study did not support this four-way interaction effect. First, male research participants did not perceive female professors who behaved in a gender incongruent manner (autocratic) while dealing with either task oriented or interpersonally oriented situations more negatively than female research participants. Further, research participants did not favor male professors over female professors when female professors behaved in a gender incongruent manner. How do the results of this study relate to the results from meta-analytic procedures? First, even though this study manipulated a condition which portrayed a female professor

behaving in a manner that is considered to be gender incongruent (autocratic) under a condition that calls for feminine-oriented skills (interpersonal situation), the manipulation may not have been strong enough to elicit the perception of gender incongruent behaviors especially in male participants since the scenarios were designed within the academic environment, which is generally believed to be more gender neutral. This is important because the magnitude of this interaction as calculated by meta-analytic research has been consistently very weak (i.e., Eagly et al., 1990; Swim et al., 1989). Eagly et al. further stressed that this differential evaluation of male and female leaders most consistently occurred within the following contexts: (a) when overall leadership measures were used (i.e., leader competence, satisfaction with leader), (b) when few independent variables were examined, and (c) when female leadership was evaluated within a male dominated field (i.e., athletic coaches), especially by male raters. The fact that the results of this study did not support a similar trend may be due to several reasons. Again, as already mentioned, the study examined leadership behaviors in an academic context that is believed to have a fairly equal gender distribution (as described by the general employment statistics of men and women within the College of Arts and Science and the College of Education as opposed to the more traditionally male oriented College of Engineering). Further, this experiment manipulated a complex set of variables simultaneously within the context of a scenario in order to allow the interplay among conditions (as opposed to manipulating only one variable). Nevertheless, the lack of an interaction effect was consistent across both samples of this study, and should not be dismissed, especially since effect sizes have consistently been very small in meta analytic research. It is crucial to also consider the publication date of studies investigating

perception of gender differences in leader positions because societal changes or shifts may occur very rapidly over time. This belief follows Gergen's (1986) notion that gender roles reflect values of society at a given point in time and must be evaluated within the context of time. For example, Rohjahn et al. (1994) quoted the 1975 research of Costrich et al. that found that men showed a tendency to evaluate female leaders who behaved in a masculine manner much less favorably than female leaders who displayed feminine behaviors. Perhaps results from this study as well as other recently published research underscore the importance of using updated research when examining the impact of societal gender roles on the evaluation of male and female leaders. Gender roles within the Western society may be changing more quickly, thus contributing to the inconsistencies in research findings (especially when current results are compared to results that are more than ten years old).

Hypothesis 5: There will be an interaction effect between participant gender, professor gender, professor leader style and type of situation on the ratings of leader likeability.

This hypothesized four-way interaction effect for likeability was based on the same premise as hypothesis 4. Again, results failed to show support for the gender congruency theory. The results of this study do show that male participants in both colleges failed to be influenced by the gender incongruent behaviors of the female professor. Instead, both male and female participants were influenced by professor leadership style when evaluating leader likeability. How does this finding compare to the findings of previous research? Bartol et al. (1976) found that male and female leaders were less liked and perceived to be less effective when they behaved in a gender

incongruent manner. More recent research (Rojahn et al., 1995) found support for the gender congruency theory only from male research participants. It appears that results of this study may indicate a weakening trend of gender role influence on the evaluation of leadership measures. Results of this study certainly seem to support the notion that it is the leadership style a leader (male or female) displays that influences how he or she is perceived in terms of competence and likeability (as well as level of masculinity). Gender of the professor clearly had a negligible influence in the perception of leader competence and likeability for male and female participants alike. This finding is encouraging because it supports a trend found in most recent research that indicates a lessening impact of social gender roles in the evaluation of male and female leaders.

Additional Findings

The results reported in this section are exploratory in nature because they were not predicted beforehand. Although these findings were interesting, it is important to point out that (a) the effects were not consistent across the two samples, and (b) the calculated effect sizes for those effects were very small. Still, reporting these results underscore the difficulties in grasping underlying mechanisms of gender differences in leadership research.

College of Arts and Science. A three-way interaction effect with regard to competence revealed that male participants rated the democratic professor in the task condition as most competent, followed by the democratic professor in the personal condition and the autocratic professor in the personal condition. The autocratic professor in the task conditions was rated the least competent. Female participants on the other hand rated the democratic professor in the personal condition as most competent,

followed by the democratic professor in the task condition. Female participants rated autocratic professors in the task and personal conditions equally low. This result is difficult to explain because it was not repeated in the College of Education sample. Nevertheless it appeared that male participants perceived democratic professors in the task condition to be most competent, while female participants rated democratic professors in the personal conditions as most competent. This could be a product of the vignette manipulation, or, perhaps that male participants felt most comfortable with the democratic task vignette, because the task situation reflected masculine qualities, while female participant felt most comfortable with the feminine overtones portrayed by the democratic personal situation

An interesting main effect with regard to the masculinity factor was found in that research participants perceived female professors to be more masculine than male professors. In addition, a two-way interaction effect revealed that female participants perceived autocratic professors to be more masculine than male participants, while all participants rated democratic professors similarly in terms of masculinity. What does this result mean? First, it needs to be pointed out that this finding was not replicated in the College of Education, and the calculated effect size was very small. Yet, the fact that female professors were perceived to be more masculine than their male counterparts may indicate that participants may have perceived both male and female professors as behaving in a somewhat gender incongruent manner. Yet, this result was tempered by the two-way interaction effect that indicated that female participants identified autocratic behavior as acutely masculine. Even though these findings were not consistent, they are intriguing because they hint at undercurrents of gender differences concerning the

inclusion and importance of external cues in the evaluation of a leader. Future research is needed to explore these issues in more depth.

With regard to the likeability factor, participants slightly preferred the male professor over the female professor. In addition, a three-way interaction effect between professor behavior, situation and participant gender indicated that female participants were more extreme in their ratings, that is, while all participants liked the autocratic professor in the task situation the least, female participants were more extreme in their low rating. Similarly, while all participants liked the democratic professor in the person-oriented situation the best, female participants also rated the democratic professor much higher than male participants on the likeability score. Again, these results were not replicated in the education sample. Still, a pattern emerges in that female participants may weigh external cues differently in the level of importance when considering leader likeability. It is recommended that future research should explore this avenue further before coming to any conclusions.

College of Education. A three-way interaction with regard to competence ratings was found between professor gender, professor behavior, and situation. The democratic male professor was rated the most competent, followed by the democratic female professor in both personal and task conditions. The lowest competence rating was given to the autocratic male professor in the personal condition, followed by the autocratic female professor in the task condition. While there was a difference in competence rating between male and female research participants in the Arts and Science sample, the results from the education sample produced a gender difference only in terms of the professor gender. This difference might have been produced due to the low numbers of male

participants available in Education (80% female participants). Yet, it is intriguing, that the democratic male professor in a personal situation received slightly higher competence ratings than his female counterpart. This result counters gender role congruency theory that punishes gender-role incongruent behavior by both genders. Luthar (1996) has noted that if the behavior of a person is perceived to be in extreme contrast to stereotypical expectation, the reaction to it could become more unpredictable. This result was not replicated in the Arts and Science sample. Nevertheless, future research should explore possible effects that extreme violations of stereotypical expectation can have on the evaluation of a leader.

Section IV Future Research

Design

In their meta-analytic review, Eagly et al. (1992) found two major research paradigms researchers used to investigate gender differences in leadership evaluation: laboratory experiments usually conducted with college students as research participants and organizational studies using employees and management to fill out rating scales. Organizational studies have the advantage of examining gender differences in a natural setting, whereas experimental studies have greater control of variable manipulation by the use of written vignettes, or by playing out a scenario with trained confederates. The use of written vignettes in experiments (a technique introduced by Rosen & Jerdee in 1973) is most common because of the relative ease of multiple variable manipulations within a scenario. The strength of an experimental design usually lies within the high degree of internal validity, that is, it is reasonable to assume that the potentially confounding effects of extraneous variables are controlled (e.g., Kennedy et al., 1985). The internal validity

of this study can be assumed to be satisfactory because of (a) the systematic manipulation of the independent variables embedded in a written scenario which remained constant and (b) the random assignment of the eight booklet version to research participants.

External validity, the process of ensuring that all members of a population have the same chance of being selected (i.e., random selection) enhances the generalizability of the results to that larger population (Kennedy et al., 1985). It was not possible to employ random selection for this study. In order to maximize external validity, two independent samples were used, one from the College of Arts and Science (N=932) and one from the College of Education (N=722). Most Introductory classes from these two colleges are attended by students with varied majors because these classes are part of the college requirement. The results of this study were replicated using two independent samples of undergraduate students who varied in terms of age, racial makeup and school standing, suggesting a moderately confident level of generalizability to the undergraduate student population.

An increasingly important issue of validity is the question of ecological validity (i.e., how appropriate is it to generalize the results of a study from one context to another?). Kerlinger (1973) described ecological validity as part of external validity, in that it concerns whether the findings derived from an experimental study conducted with a sample drawn from a particular population can be generalized to other settings or conditions. One major disadvantage of experimental designs is the byproduct of artificiality created by the experimental control. For example, Eagly et al. (1990) found that when participants in experiments were given the role of a leader, they tended to behave more stereotypically than managers observed within organizational settings.

Many researchers support the relevance or ecological validity contribution of experimental studies in the field of leadership studies. For example, Locke (1986) found that results of laboratory studies were equivalent to results found in field studies. Similar findings derived from different methodologies indeed present a cogent argument to support the notion that experimental designs have a high degree of internal validity as well as ecological validity. It is difficult to evaluate the level of ecological validity of the present study. Even though the researcher is confident that this study achieved a reasonable level of internal validity, the issues of ecological validity are more complex. Most importantly, ecological validity focuses on the realism of the experiment. This study had several strengths as well as weaknesses regarding the issue of realism. First, this study conducted pilot studies that investigated the plausibility of the content of the scenarios. Results indicated that both male and female participants rated both task and personal scenarios as plausible regardless of professor gender or professor leadership. Yet, the relationship between an employee and manager, or between a professor and student as it occurs in a realistic setting could not be simulated by one written paragraph. Thus, one scenario can only portray one incident. The realism of a relationship between leader and subordinate (professor and student) is impossible to simulate. More research is needed that focuses on how to make experimental studies more realistic without sacrificing the level of internal control. Pilot results of this study, for example, provided evidence that the scenarios were rated to be realistic, providing evidence for scenario realism. However, different methodologies may yield different results. For example, this study only used verbal descriptors of the professor. A degree of realism could be enhanced if visual images (photographs) of the professor were added to the verbal

descriptors. The impact of visual depiction may lead to different results. Extensive research, for example, has provided evidence that physically attractive people are generally evaluated more positively than their less attractive counterparts (e.g., Efran, 1974).

Finally, one of the most common sources of bias in attitudinal studies is the halo effect (Pedhazur & Schmelkin, 1991). This type of bias occurs when a general impression of the person to be rated is formed, which then systematically influences the ratings for that person on unrelated dimensions. For example, Nisbett and Wilson (1977) found that students who watched a videotape of a friendly professor rated that professor more favorably on other traits than students who watched the same videotape but with an arrogant professor. The results of this study found that democratic professors were rated to be more competent and more likeable than their autocratic counterparts. Could a systematic bias have contributed to this result? This is a difficult question to answer in hindsight. It is possible, that democratic professors may have received inflated likeability ratings due to the halo effect. However, results of this study found interaction effects with regard to competence and likeability ratings (see additional findings) that suggest a minimal halo effect on the outcome measures. Nevertheless, future research should investigate the impact of halo effects on the ratings of leadership qualities.

Implications

This study generated a list of male and female traits based on currently widely accepted theoretical frameworks of gender stereotypes (e.g., Eagly et al., 1984). These gender stereotypes (e.g., dominant and aggressive for men, submissive and friendly for women) have changed very little since Halpin's (1957) research on gender differences in

leadership behavior. It may be appropriate to question whether, for example, men are still believed to be more self assertive and master their environment more so than women, while women on the other hand are still believed to be more selfless and concerned with others. Stereotypes people have about gender and leadership roles may have significantly changed since the 70's and 80's to the point of re-evaluating those beliefs. Similarly, research should update theoretical frameworks of gender roles (e.g., gender role spillover, gender congruency theory) in order to enhance the predictive power of the model. Gender role theories are very popular, yet, meta-analytic results revealed weak evidence at best in their support.

Results of this experiment showed that male and female participants rated democratic professors more competent than autocratic professors regardless of gender and situation. This information could be helpful for tenure track seeking instructors who depend in part on student ratings for their job promotions. Simply being aware that a democratic teaching style can positively impact teacher evaluations could help improve teaching skills as well as increase their chances for promotions. Further, teaching democratic leadership behaviors seem crucial in developing future leaders within academic and organizational environments. This study provided evidence that democratic behaviors are conveyed verbally as well as nonverbally. Democratic nonverbal behaviors (e.g., leaning forward, smiling, nodding, moderately high eye contact) should be included when teaching future leaders to achieve democratic leadership style. In addition, this study provided evidence that democratic leadership styles are equally effective in task as well as personal situations. This information is important especially for female leaders who may feel pressured to adopt an autocratic style in personal situations so as not to be

perceived as weak. Current and future leaders within academic or organizational arenas should consider de-emphasizing an autocratic style and instead emphasize leadership behaviors that reflect a democratic style.

References

- Andersen, P. A., & Coussoule, A. R. (1980). The perceptual world of the communication apprehensive: The effect of communication apprehension and interpersonal gaze on interpersonal perception. *Communication Quarterly*, 28, 44-54.
- Alimo-Metcalfe, B. (1995). Leadership and Assessment. In A. Buckley, *The essence of women in management*, (pp. 92-109). Englewood Cliffs, NJ: Prentice Hall.
- Ashmore, R. D., & Del Boca, F. K. (1979). Sex stereotypes and implicit personality theory: Toward a cognitive-social psychological conceptualization. *Sex Roles*, 5, 219-248.
- Bales, R. F. (1950). *Interaction process analysis: A method for the study of small groups*. Reading, MA: Addison-Wesley.
- Baron, R. M., & Kenny. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173-1182.
- Bartol, K. M. & Martin, D. C. (1986). Women and men in task groups. In R. D. Ashmore & R. K. Del Boca (eds). *The social psychology of female-male relations: A critical analyses of central concepts* (pp. 259-310)
- Bass, B. M. (1990). *Bass & Stogdill's handbook of leadership: Theory, research, and managerial applications* (3rd ed.). New York: Free Press.
- Bass, B. M. (1981). *Stogdill's handbook of leadership: A survey of theory and research* (rev. ed.). New York: Free Press.
- Berger, J., Webster, M., Jr., Ridgeway, C. L., & Rosenholtz, S. J. (1986). Status cues, expectations and behaviors. In E. Lawler (Ed.), *Advances in group processes*, (p 3, 1-22). Greenwich, CT: JAI Press.
- Berger, J., Wagner, D. G., & Zelditch, M., Jr. (1985). Introduction: Expectation states theory. In J. Berger & M. Zelditch, Jr. (Eds.), *Status, rewards, and influence* (pp. 1-72). San Francisco: Jossey-Bass.
- Berger, J., Rosenholtz, S. J., & Zelditch, M., Jr., (1980). Status organizing processes. *Annual Review of Sociology*, 6, 479-508.

Bergmann, W. (ed.) (1988). Error without trial. In *Psychological Research on Antisemitism*. New York, NY: Walter de Gruyter.

Briton, N. J., & Hall, J. A. (Spring, 1995). Gender-based expectancies and observer judgements of smiling. *Journal of Nonverbal Behavior*, 19(1), 49-65.

Broverman, I. K., Vogel, S. R., Broverman, D. M., Clarkson, F. E., & Rosenkrantz, P. S. (1972). Sex-role stereotypes: A current appraisal. *Journal of Social Issues*, 28(2), 59-78.

Burgoon, J. K., Coker, D. A., & Coker, R. A. (1986). Communicative effects of gaze behavior: A test of two contrasting explanations. *Human Communication Research*, 12(4), 495-524.

Burgoon, J. K. (1985). Nonverbal signals. In M. L. Knapp & G. R. Miller (Eds.) *Handbook of interpersonal communication*. Beverly Hills, CA: Sage.

Burgoon, J. K., Buller, D. B., Hale, J. L., & deTurck, M. A. (1984). Relational messages associated with nonverbal behaviors. *Human Communication Research*, 10, 351-378.

Burgoon, J. K., & Jones, S. B. (1976). Toward a theory of personal space expectations and their violations. *Human Communication Research*, 2, 131-146.

Carli, L. L., LaFleur, S. J., & Loeber, C. C. (1995). Nonverbal behavior, gender, and influence. *Journal of Personality and Social Psychology*, 68(6), 1030-1041.

Carli, L. L., Martin, C., Leatham, G., Lyons, K., & Tse, I. (1993). *Perceptions of nonverbal behavior*. Paper presented at the annual meeting of the American Psychological Association, Toronto, Ontario, Canada.

Colwill, N. L. (1995). Sex differences. In Buckley, A. (1995). *The essence of women in management*, (pp. 20-34). Englewood Cliffs, NJ: Prentice Hall.

Denmark, F. L. (1993). Women, leadership, and empowerment. *Psychology of Women Quarterly*, 17, 343-356.

Deaux, & Major. (1987).

Deutsch, F. M. (September, 1990). Status, sex and smiling: The effect of role on smiling in men and women. *Personality and Social Psychology Bulletin*, 16(3), 521-540.

Dorenkamp, A. G., McClymer, J. F., Moynihan, M. M., & Vadum, A.C. (1995). *Images of women in American popular culture* (2nd ed). Fort Worth, TX: Harcourt Brace & Co.

Dovidio, J. F., Keating, C. F., Heltman, K., Ellyson, S. L., & Brown, C. E. (1988). The relationship of social power to visual displays of dominance between men and women. *Journal of Personality and Social Psychology*, *54*(2), 233-242.

Dovidio, J. F., Heltman, K., Brown, C. E., Ellyson, S. L., & Keating, C. F. (1988). *Journal of Personality and Social Psychology*, *55*(4), 580-587.

Eagly, A. H., Karau, S. J., Makhijani, M. G. (1995). Gender and the effectiveness of leaders: A meta-analysis. *Psychological Bulletin*, *117*(1), 125-145.

Eagly, A. H., & Johnson, B. T. (1990). Gender and leadership style: A meta-analysis. *Psychological Bulletin*, *108*(2), 233-256.

Eagly, A. H., Makhijani, M. G.; & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, *111*(1), 3-22.

Eagly, A. H., & Wood, W. (1991). Explaining sex differences in social behavior: A meta-analytic perspective. *Journal of Personality and Social Psychology*, *17*(3), 306-315.

Eagly, A. W., & Mladinic, A. (1989). Gender stereotypes and attitudes toward women and men. *Personality and Social Psychology Bulletin*, *15*(4), 543-558.

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale, NJ: Erlbaum.

Eagly, A. H., & Steffen, V. J. (1986).

Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, *46*, 735-754.

Ellsworth, P. C., & Langer, E. J. (1976). Staring and approach: An interpretation of the stare as a nonspecific activator. *Journal of Personality and Social Psychology*, *33*, 117-122.

Ellyson, S. L., & Dovidio, J. F. (Eds.) (1985). *Power, dominance, and nonverbal behavior*. New York, NY: Springer-Verlang.

Exline, R. V., Ellyson, S. L., & Long, B. (1975). Visual behavior as an aspect of power role relationship. In P. Pliner, L. Krames, & T. Alloway (Eds.), *Nonverbal communication of aggression* (pp. 21-52). NY: Plenum Press.

Exline, R. V., & Winters, L. C. (1965). Affective relations and mutual glances in dyads. In S. S. Tomkins, & C. W. Izard (Eds.) *Affect, cognition and personality*. NY: Springer.

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.

Fiske, S. T., Bersoff, D. N., Borgida, E., Deaux, K., & Heilman, M. E. (1991) Social science research on trial: The use of sex stereotyping research in Price Waterhouse vs. Hopkins. *American Psychologist*, 46, 1049-1060.

Friesen, E. P., & Ellsworth, P. (1972). *Emotion in the human face*. NY: Pergamon Press

Gergen, K. J. (1986). Correspondence versus autonomy in the language of understanding human action. In D. W. Fiske & R. A. Shewder (Eds.) *Metatheory in Social sciences*. Chicago: University of Chicago Press.

Goffman, E. (1976). Gender advertisements. *Studies in the Anthropology of Visual Communication*, 3, 69-154.

Gutek, B. A. (1993). Changing the status of women in management. *Applied Psychology: An International Review*, 42, 301-311.

Gutek, B. A., & Morasch, B. (1982). Sex-ratios, sex-role spillover, and sexual harassment of women at work. *Journal of Social Issues*, 38(4), 55-74.

Halberstadt, A. G., Hayes, C. W., & Pike, K. M. (1988). Gender and gender role differences in smiling and communication consistency. *Sex Roles*, 19(9-10), 589-605.

Halberstadt, A. G., & Saitta, M. B. (1987). Gender, Nonverbal behavior, and perceived dominance: A test of the theory. *Journal of Personality and Social Psychology*, 53(2), 257-272.

Hall, J. A., & Halberstadt, A. G. (1986). Smiling and gazing. In J. S. Hyde & M. Linn (Eds.), *The psychology of gender: Advances through meta-analysis*. Baltimore, MD: The John Hopkins University Press.

Hall, J. A. (1984). *Nonverbal sex differences: Communication accuracy and expressive style*. Baltimore, MD: John Hopkins University Press.

Harris, M. B. (1976). The effects of sex, sex-stereotyped descriptions, and institution on evaluation of teachers. *Sex Roles*, 2, 15-21.

- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press
- Heilman, M. E., Block, C. J., Martell, R. F., & Simon, M. C. (1989). Has anything changed? Current characterizations of men, women, and managers. *Journal of Applied Psychology, 74*, 935-942.
- Henley, N. M. (1997). *Body politics: Power, sex, and nonverbal communication*. Englewood Cliffs, NJ: Prentice Hall.
- Henley, N. M. (1973). Status and sex: Some touching observations. *Bulletin of the Psychonomic Society, 2*, 91-93.
- Hollander, E. P. (1992). The essential interdependence of leadership and followership. *Current Directions in Psychological Science, 1*, 71-75.
- Hooff, J. A. R. A. M. van (1962). Facial expressions in higher primates. *Symposia of the Zoological Society of London, 8*, 97-125.
- House, R. J. (1971). A path goal theory of leader effectiveness. *Administrative Science Quarterly, 16*, 321-338.
- Hunt, J. G. (1991). *Leadership: A new synthesis*. Newbury Park, CA: Sage.
- Johnson, C. (1994). Gender, legitimate authority, and leader-subordinate conversations. *American Sociological Review, 59*, 122-135.
- Kanter, R. M. (1977). *Men and women of the corporation*. New York: Basic Books.
- Kaschak, E. (1978). Sex bias in student evaluations of college professors. *Psychology of Women Quarterly, 2*(3), 235-43.
- Kierstead, D., D'Agostino, P., & Dill, H. (1988). Sex role stereotyping of college professors: Bias in students' ratings of instructors. *Journal of Educational Psychology, 80*(3), 342-344.
- Kleinke, C. L. (1980). Interaction between gaze and legitimacy of request on compliance in a field setting. *Journal of Nonverbal Behavior, 5*, 3-10.
- Klenke, K. (1996). *Women and leadership: A contextual perspective*. NY, NY: Springer.
- Kollock, P., Blumsteing, P. & Schwartz, P. (1985). Sex and power in interaction: Conversational privileges and duties. *American Sociological Review, 52*, 34-46.

Kraut, R. E., & Johnston, R. E. (1979). Social and emotional message of smiling: An ethological approach. *Journal of Personality and Social Psychology*, 37, 9, 1539-1553.

Kruse, L., & Wintermantel, M. (1986). Leadership Ms.-qualified: I. The gender bias in everyday and scientific thinking. In C. F. Graumann & S. Moscovici (Eds.) *Changing conceptions of leadership*, pp. 171-197. New York: Springer-Verlag.

Levin, K., & Lippitt, R. (1938). An experimental approach to the study of autocracy and democracy: A preliminary note. *Sociometry*, 1, 292-300.

Lindsey, L. (1990). *Gender roles: A sociological perspective*. Englewood Cliffs, NJ: Prentice-Hall.

Loden, M. (1985). *Feminine leadership or how to succeed in business without being one of the boys*. New York: Times Books.

MacCoby, E. E., & Jacklin, C. N. (1974). *The psychology of sex differences*. Stanford, CA: Stanford University Press.

Martin, E. (1984). Power and authority in the classroom: Sexist stereotypes in teaching evaluations. *Journal of Women in Culture and Society*, 9, 482-492.

MacCallum, R. (1986). Specification searches in covariance structure modeling. *Psychological Bulletin*, 100, 107-120.

Mehrabian, A. (1967). Orientation behaviors and nonverbal attitude communication. *Journal of Communication*, 16, 324-332.

Nieva, V. E., & Gutek, B. A. (1980). Sex effects on evaluation. *Academy of Management Review*, 5, 267-276.

Payne, K. E., Fuqua, H. E., & Canegami, J. P. (1997). Women as Leaders. In Rosenbah, W. E., & Taylor, R. L. (1998). *Contemporary issues in leadership*, p. 145-167. Boulder, CO: Westview Press.

Petty, M. M., & Lee, G. K. (1975). Moderating effects of sex of supervisor and subordinate on relationships between supervisory behavior and subordinate satisfaction. *Journal of Applied Psychology*, 60, 624-628.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, Design, and Analysis: An Integrated Approach*. Hillsdale, NJ: Lawrence Erlbaum.

Phillips, J. S., & Lord, R. G. (1982). Schematic information processing and perceptions of leadership in problem-solving groups. *Journal of Applied Psychology*, *67*, 486-492.

Powell, G. N. (1988). *Women and men in management*. Newbury Park, CA: Sage.

Ragins, B. R., & Sundstrom, E. (1989). Gender and power in organizations: A longitudinal perspective. *Psychological Bulletin*, *105*, 51-88.

Remland, M., Jacobson, C., & Jones, T. (1983). Effects of psychological gender and sex-incongruent behavior on evaluations of leadership. *Perceptual and Motor Skills*, *57*, 783-789.

Rice, R. W., Instone, D., & Adams, J. (1984). Leader sex, leaders success, and leadership process: Two field studies. *Journal of Applied Psychology*, *69*, 12-31.

Ridgeway, C. (1987). Nonverbal behavior, dominance, and the basis of status in task groups. *American Sociological Review*, *52*, 683-694.

Ridgeway, C. L., Berger, J., & Smith, L. (1985). Nonverbal cues and status: An expectation states approach. *American Journal of Sociology*, *90*, 955-978.

Ridgeway, C. L., (1982). Status in groups: The importance of motivation. *American Sociological Review*, *47*, 76-88.

Rojahn, K., & Willemssen, T. M. (1994). The evaluation of effectiveness and likability of gender-role congruent and gender-role incongruent leaders. *Sex Roles*, *30*(1-2), 109-119.

Sargent, A. G. (1981). *The androgynous manager*. New York: Amacom.

Strauss, H. A., & Bergman, W. (1988). *Current research on antisemitism*. New York, NY: Walter de Gruyter.

Swim, J., Borgida, E., Maruyama, G., & Myers, D. G. (1989). Joan McKay versus John McKay: Do gender stereotypes bias evaluations? *Psychological Bulletin*, *105*, 409-429.

Tracy, L. (1987). Consideration and initiating structure: Are they basic dimensions of leader behavior. *Social Behavior and Personality*, *15*(1), 21-33.

Tocqueville, A. de. (1840). *Democracy in America: Part the Second. The Social Influence of Democracy* (trans. Henry Reeve). NY: J & H. G. Langley.

Unger, R. (1979). Sexism in teacher evaluation: The comparability of real life to laboratory analogs. *Academic Psychology Bulletin*, *1*, 163-171.

Vinicombe, S. & Colwill, N. L. (1995). Managers and Secretaries. In Buckley, A. (1995). *The essence of women in management*, (pp. 59-73). Englewood Cliffs, NJ: Prentice Hall.

Welsh, M. C. (1979). Attitudinal measures and evaluation of males and females in leadership roles. *Psychological Reports*, *45*, 19-22.

Willson, A., & Lloyd, B. (1990). Gender vs. power: Self-posed behavior revisited. *Sex Roles*, *23*(1-2), 91-98.

Wittrock, M. C., & Lumsdaine, A. A. (1977). Instructional psychology. In M. R. Rosenzweig & L. W. Porter (Eds.) *Annual Review of Psychology*, *28*, 417-459.

Wolfgang, A. (Ed.) (1984). *Nonverbal behavior: Perspective, applications, intercultural insights*. Lewiston, NY: C. J. Hogrefe.

Wood, W. (1987). Meta-analytic review of sex differences in group performance. *Psychological Bulletin*, *102*, 53-71.

Appendices

Appendix A. Pilot I: Validation of Scenario Manipulation

The purpose of this pilot was to statistically validate the experimental manipulations of two leadership dimensions (task-oriented and person-oriented), and leader behavior styles (democratic and autocratic), leader gender (male/female) of a set of eight written scenarios depicting an academic environment. Specifically, this pilot study examined whether research participants perceived the experimental conditions as they were designed to be perceived. Thus, the task-oriented situation was expected to be perceived as less emotional and more objective than the person-oriented situation. Similarly, the democratic communication style was expected to be perceived as friendly and supportive, while the autocratic communication style was expected to be rated as non-supportive and cold.

Dependent Measures

Using a seven-point semantic differential scale, three dependent descriptors were used to describe each of the two leadership dimensions (task/person, and democratic/autocratic). Specifically, the task/person dimension was measured using the following anchors: personal/impersonal, emotional/unemotional, and objective/subjective. The democratic/autocratic communication style was measured using the following anchors: warm/cold, non-supportive/supportive, and uncaring/caring. Finally, this study examined whether the content of the scenario was perceived to be plausible. The measures for scenario plausibility were measured using the following anchors: plausible/implausible, realistic/unrealistic, and clear/unclear.

Appendix A. (Continued)

Procedure

An experimental design was used, in which research participants were randomly assigned to one of the eight scenario versions. Data were collected during undergraduate and graduate classes within the Education department of a large urban university. A total of 18 classes were used for this pilot with an average class size of 20. At the beginning of class, the author or the class instructor explained the purpose of the study (judgment of professor/student interaction). It was stressed that there were no correct answers, and that responses were anonymous. Participation was voluntary, and no incentives were offered for participation. Participants completed the experiment within approximately 10 minutes.

Data Analysis

Three factorial Analyses of Variances (ANOVAs) were used to examine how students perceived the two leadership dimensions (task/person and democratic/autocratic), as well as students' perception of vignette plausibility. For each dimension, the dependent variables were averaged to form one factor score estimate. Internal reliability estimates were calculated using Cronbach alphas. Finally, effect size estimates using Keppel's Omega Squared (Keppel, 1991).

Appendix A. (Continued)

*Results**Sample*

A total of 356 students participated in this pilot. One hundred forty-eight (42%) were male, and 208 (58%) were female. Mean age for female students was 27.4 years (range= 17-18 years, SD= 8.7), and mean age for male students was 29.8 (range= 18-59 years, SD= 10.6). Seventy-one percent (n=266) of the sample was Caucasian. 6.4% (n=22) were African American, 4.6% (n=16) were of Hispanic origin, and 8.1% (n=28) were of Asian/Pacific background. 65% (n=225) were undergraduate students, and 28% (n=98) were graduate students.

Ratings of Task-Oriented and Person-Oriented Dimension (Scenario Situation)

Three items (personal/impersonal, emotional/unemotional, objective/subjective) rated on a seven- point semantic differential type scale were used. The higher number represents the descriptors personal, emotional, and subjective. Because two descriptors (personal/impersonal and emotional/unemotional) were reversed to avoid response set during the rating process (Crocker & Algina, 1986), they were reversed during statistical analysis. Further, these three items were combined to form one dependent factor score estimate representing the task/person dimension. Calculated internal consistency estimates as measured by Cronbach alpha for the task/person dimension was .95. This resulting factor score estimate was then analyzed using a 2 (gender of research participant) X 2 (task/person situation) X 2 (professor gender) X 2 (professor democratic/autocratic communication style) univariate analysis of variance (ANOVA). Results (see Table 2 for ANOVA summary) revealed a main effect for participant gender,

Appendix A. (Continued)

$F(1, 353) = 4.40, p < .04$). The relative treatment magnitude as measured by omega squared (w^2) was .003. Cohen's guideline for treatment magnitude measure w^2 suggests that a magnitude of .01 should be considered a 'small' effect, an obtained w^2 of .06 should be labeled a 'medium' effect, and obtaining w^2 of .15 or greater should be considered a 'large' treatment magnitude (Cohen, 1988). Compared to Cohen's standard, the obtained treatment magnitude of .003 was very small. Another type of effect size estimation for analyses of variance procedures is the f statistic. This measure represents the standard deviation of the standardized means (Stevens, 1990), and characterizes an f around .1 to be a small effect size, an f around .25 to be a medium effect size, and an f larger than .4 to be a large effect size (Cohen, 1988). Using this measure, the effect size estimate for the participant gender main effect was small ($f = .11$).

Appendix A. (Continued)

Table 2

Obtained ANOVA Source Table for Task/Person Dimension

Source	df	MS	F	w ²	f
Participant Gender (PaG)	1	4.94	4.40*	.003	.11
Professor Gender (PrG)	1	0.55	0.49		
Democratic/Autocratic (Pr) Style	1	0.20	0.18		
Task/Person Situation (Situation)	1	704.1	627.88**	.06	1.33
PaG x PrG	1	0.34	0.30		
PaG x Pr Style	1	0.01	0.01		
PrG x PrStyle	1	0.91	0.81		
PaG x Situation	1	0.73	0.65		
PrStyle x Situation	1	3.20	2.85		
PaG x PrG x PrStyle	1	0.28	0.25		
PaG x PrG x Situation	1	0.10	0.08		
PaG x PrStyle x Situation	1	1.54	1.38		
PaG x PrG x PrStyle x Situation	1	0.76	0.67		
Error	338	1.12			

*p < .04

**p < .0001

Results further revealed that research participants rated a task-oriented situation as more objective and less personal than the person-oriented situation (see Table 3 for summary statistics) $F(1,353)= 627.88, p< .0001$). The calculated treatment magnitude for the situation main effect was medium ($w^2=.06$), and the effect size estimate was large ($f=1.33$).

Appendix A. (Continued)

Table 3

Summary Statistics for Task/Person Factor Score^a

Scale Dimension	Coefficient Alpha	<u>Mean</u>		<u>SD</u>	
		Males	Females	Males	Females
Task Situation (Objective)	.95	3.52	3.82	1.38	1.33
Person Situation (Subjective)	.95	6.56	6.74	0.82	0.63

^aN = 356*Ratings of the Democratic/Autocratic Dimension**(Professor Nonverbal Communication Style)*

The democratic/autocratic dimension was measured using three items on a seven-point semantic differential anchored by warm/cold, non-supportive/supportive, and uncaring/caring. A lower score represents an autocratic style, while a higher rating score represents a democratic style. One item (warm/cold) was reversed for statistical analysis. All three items were combined to create a factor score. Internal consistency estimates as calculated by Cronbach alpha for was .95. This factor score estimate was then analyzed using a 2(gender of research participant) X 2 (task/person situation) X 2 (professor gender) X 2 (professor democratic/autocratic communication style) univariate analysis of variance (ANOVA). Results (see Table 4 for ANOVA summary) revealed a statistically significant main effect for professor gender $F(1, 355) = 6.07, p < .004$. The calculated

Appendix A. (Continued)

treatment magnitude as measured by w^2 was very small (.004), and the calculated effect size was very small ($f=.13$). Further, the main effect for task/person was statistically significant $F(1, 355) = 18.16, p < .0001$. Calculated treatment magnitude for this main effect was small ($w^2=.01$), and its calculated effect size was small ($f=.22$). Further, the democratic/autocratic nonverbal communication style dimension was statistically significant $F(1, 355) = 532.43, p < .0001$. The calculated treatment magnitude was medium ($w^2=.5$), and its effect size was large ($f=1.22$). Finally, there was a three way interaction effect between participant gender, professor nonverbal communication style, and task/person situation dimension $F(1, 355) = 11.82, P < .0007$. The calculated treatment magnitude was small ($w^2=.01$), and the effect size estimate was small ($f=.19$). Table 5 shows how men and women rated the scenarios under different conditions of communication style and task/person situation. It appears that the perception of democratic behavior (e.g., friendly) may depend on the type of situation, as well as the gender of the rater.

Appendix A. (Continued)

Table 4

ANOVA Source Table for Democratic/Autocratic Communication Style

Source	Df	MS	F	w ²	f
Participant Gender (PaG)	1	0.00	0.00		
Professor Gender (PrG)	1	6.88	6.07*	.004	.13
Professor Communication Style (PrStyle)	1	603.14	532.43**	.50	1.22
Task/Person Situation (Situation)	1	20.57	18.16*	.02	.22
PaG x PrG	1	0.54	0.48		
PaG x Pr Style	1	2.47	2.18		
PrG x PrStyle	1	2.61	2.31		
PaG x Situation	1	0.03	0.03		
PrG x Situation	1	2.24	1.98		
PrStyle x Situation	1	0.10	0.96		
PaG x PrG x PrStyle	1	0.56	0.50		
PaG x PrG x Situation	1	1.88	1.66		
PaG x PrStyle x Situation	1	13.40	11.82**	.01	.19
PaG x PrStyle x Situation	1	1.42	1.26		
PaG x PrG x PrStyle x Situation	1	0.72	0.68		
Error	340	385.15			

*p < .01

**p < .0001

Appendix A. (Continued)

Table 5

Summary Statistics of Democratic/Autocratic Nonverbal Communication Style Broken Down by Task/Person Situation and Participant Gender^w

Professor Communication Style (a .95)	Task/Person Situation	<u>Mean</u>		<u>SD</u>	
		Males	Females	Males	Females
Democratic	Task	4.57	4.31	1.57	0.74
Democratic	Person	4.73	5.33	1.26	1.10
Autocratic	Task	1.73	1.88	0.82	0.98
Autocratic	Person	2.47	1.91	1.31	.98

N = 357

Ratings of Scenario Plausibility

This pilot study examined whether the written scenarios were perceived as plausible events within an academic setting. Scenario plausibility was measured using two items on a seven-point semantic differential anchored by plausible/implausible, and realistic/unrealistic. Higher ratings represent the implausible dimension, while a lower rating represents the realistic aspect of the dimension. Both items were combined to create a factor score. Calculated internal consistency for this factor score as estimated by Cronbach alpha was a stable .82. This factor score estimate was then analyzed using a 2 (gender of research participant) X 2 (task/person situation) X 2 (professor gender) X 2 (professor democratic/autocratic communication style) univariate analysis of variance (ANOVA). Results (see Table 6 for ANOVA summary source table). Results revealed a significant main effect for task/person situation dimension $F(1, 352) = 48.90, p < .0001$.

Appendix A. (Continued)

The calculated treatment magnitude was large ($w^2=.11$), along with a medium effect size estimate ($f=.37$). Further, a significant interaction between participant gender and professor gender was found $F(1, 352) = 6.58, p < .01$. The corresponding treatment magnitude this interaction effect was small ($w^2=.01$), and the effect size estimate was small ($f=.17$). The mean rating for a task-oriented scenario was 2.07 (SD= 1.01) compared to the mean rating of 3.06 (SD=1.43) for the person-oriented scenario. Even though this difference produced a statistical significant difference, the value of 3.06 for the person-oriented situation still is well within the range of being considered as realistic and plausible (scale ranged from 1(realistic) to 7 (unrealistic)). Table 7 depicts the means and standard deviations produced by the interaction effect between participant gender and task/person situation dimension. It appears that male students rated the task situation as less realistic (see Table 7) than female students. However, both men and women rated the person-oriented vignette comparably with regard to plausibility.

Appendix A. (Continued)

Table 6

ANOVA Source Table for Ratings of Scenario Plausibility

Source	Df	MS	F	w ²	f
Participant Gender (PaG)	1	0.00	0.00		
Professor Gender (PrG)	1	2.74	1.72		
Professor Communication Style (PrStyle)	1	0.19	.012		
Task/Person Situation (Situation)	1	77.92	48.90*	.11	.37
PaG x PrG	1	10.48	6.58**	.01	.17
PaG x Pr Style	1	4.40	2.76		
PrG x PrStyle	1	3.86	2.42		
PaG x Situation	1	1.51	.095		
PrG x Situation	1	0.02	0.02		
PrStyle x Situation	1	0.64	0.40		
PaG x PrG x PrStyle	1	0.77	0.48		
PaG x PrG x Situation	1	0.48	0.30		
PaG x PrStyle x Situation	1	0.05	0.04		
PaG x PrStyle x Situation	1	1.11	0.70		
PaG x PrG x PrStyle x Situation	1	0.78	0.49		
Error	337	1.59			

*p < .01

**p < .0001

Appendix A. (Continued)

Table 7

Summary Statistics of Plausibility Ratings of Male and Female Students Under the Scenario Dimension of Task/Person

Task/Person Dimension	<u>N</u>		<u>Mean</u>		<u>SD</u>	
	Male	Female	Males	Females	Males	Females
Task Situation	61	91	2.20	1.98	1.04	0.99
Person Situation	78	117	3.02	3.10	1.53	1.36

Summary

The purpose of this pilot study was to verify the experimental manipulation. Results of the experiment revealed that the manipulations of the two leadership dimensions of task/person, and democratic/autocratic were perceived as expected. Further, the depictions of scenarios were perceived by men and women to be realistic and plausible.

Appendix A. (Continued)

Scenario of Student - Professor Interaction

If at any point you become uncomfortable with this attitude scale, you may discontinue.

DO NOT write your name on this form. All your responses are anonymous.

DIRECTIONS:

Please read inside this booklet the description of an interaction between a student and professor. After reading the scenario go to the next page where you will see a list of descriptors for the professor and the scenario. Each descriptor consists of a pair of extreme opposites separated by seven spaces. Consider each descriptor and mark one of the spaces between each pair of opposites that is closest to your perceptions. Please mark every descriptor **using your best judgement**.

EXAMPLES

If you feel the person *is probably* competent, you would mark a space near the word competent.

Competent ___ **T** ___ ___ ___ ___ ___ Incompetent

If you feel the person *is definitely not* competent, you would mark a space nearest the word incompetent.

Competent ___ ___ ___ ___ ___ ___ **T** Incompetent

In addition, please provide the biographical information requested on the last page.
Thank you.

Appendix A. (Continued)

Scenario (sample 1 out of 8)

The student knocks on Professor Smith's door. "Do you have a minute?" Sitting at her computer, Professor Smith gets up from her desk, nods an encouraging hello to the student and says, "Come in, come in." Professor Smith sits in a chair opposite the student and asks, "What seems to be the problem?" The student slumps into a chair and looks down on the ground. Professor Smith leans forward toward the student, folds her arms in her lap and patiently waits for the student. The student says, "My long-term relationship just fell to pieces and I am so upset over it that I really don't know what to do." Looking compassionately at the student, Professor Smith gently replies, "I don't know how to help you." Professor Smith remains seated across from the student while she lets some time pass. "I can see this will take more time than we have right now. Let's make an appointment."

Appendix A. (Continued)

Mark the space between each pair of opposites that is closest to your perception

How would you describe the professor-s behavior?

warm	_ _ _ _ _	cold
nonsupportive	_ _ _ _ _	supportive
uncaring	_ _ _ _ _	caring

**The scenario consists of a student presenting a problem.
How would you describe the nature of this problem ?**

personal	_ _ _ _ _	impersonal
emotional	_ _ _ _ _	unemotional
objective	_ _ _ _ _	subjective

How would you describe the scenario presented here?

plausible	_ _ _ _ _	implausible
realistic	_ _ _ _ _	unrealistic
clear	_ _ _ _ _	unclear

Please turn over **L**

Appendix B. Pilot II: Manipulation Check on Dependent Measures

The purpose of the second pilot was to statistically evaluate the internal consistencies of the dependent measures, as well as evaluate the effectiveness of the data collection procedure.

Dependent Measures

Each of the eight scenario versions (see Appendix A) was rated for leadership effectiveness, leader likeability, and leader character attributes.

Leader effectiveness. Using a five-point semantic differential scale, three effectiveness descriptors anchored with the following descriptors were used: competent/incompetent, effective/ineffective, and well qualified/not at all qualified. In order to have a higher number reflect a higher degree of effectiveness, the scale direction was reversed during statistical analysis to follow that logic.

Leader likeability. Four descriptors arranged on a five-point semantic differential scale were used anchored by the following descriptors: critical/tolerant, considerate/inconsiderate, popular/unpopular, and likeable/not likeable. A higher rating reflected a higher degree of likeability, thus, some of the items re-scaled during statistical analysis to fit that logic.

Leader character attributes. Fifteen descriptors were arranged on five-point semantic differential reflecting male and female stereotypes. These 15 character attributes were anchored with the following positive and negative anchors: powerful/powerless, timid/forceful, hardworking/lazy, persistent/gives up easily, soft/tough, fair/not fair, responsible/irresponsible, not helpful/helpful, cooperative/not cooperative, trustworthy/untrustworthy, aggressive/not aggressive, independent/dependent,

Appendix B (Continued)

dominant/submissive, objective/subjective, and unprepared/prepared. A lower rating on this scale corresponded to a negative character evaluation, while a higher rating reflected a positive character attributes.

Additional questions. In addition, six questions were posed to evaluate professor appropriateness, communication skills, popularity with students, as well as questions referring to professor's chances for future promotions.

Procedure

An experimental design was used, in which research participants from two undergraduate and one graduate class recruited from the Education department of a large urban university. After one of eight scenario versions was randomly assigned to participating students, they were asked to complete the forms. Participation was voluntary, and no incentives were offered for participation. On average it took about 8 minutes to complete the experiment. Further, students who participated in the study were encouraged to give verbal or written feedback with regard to the experiment. It was stressed that all responses were anonymous.

It was stressed that all responses were anonymous.

Data Analysis

Internal reliability estimates were calculated using Cronbach alphas for the descriptors of leader effectiveness, leader likeability, and leader character attributes. Student comments were evaluated and used to make design improvements.

Appendix B (Continued)

*Results**Sample*

Sixty students participated in this pilot. Twenty-nine (48.3%) were male, and 31 (51.7%) were female. Mean age for female students was 32.13 years (SD= 7.5 years, range= 21-48 years), and the mean age for male students was 35.1 (SD= 8.6 years, range= 22-57 years). For this sample, 79.3% (n=37) were Caucasian, 8.6% (n=5) were African American, 3.5% (n=2) were Hispanic, and 3.5% (n=2) marked an Asian/Pacific background. 37% (n=23) of the sample were undergraduate students, and 63% (n=37) were graduate students.

Reliability Estimates (Cronbach Alpha's) for the Dependent Measures

Leader effectiveness. The three effectiveness indicators (competent/incompetent, effective/ineffective, well qualified/not at all qualified) were combined equally in order to receive a leader effectiveness factor score estimate. Calculated internal consistency estimates as measured by Cronbach alpha was .84.

Leader likeability. The four indicators for leader likeability (critical/tolerant, considerate/inconsiderate, popular/unpopular, likeable/not likeable) were combined resulting in one likeability factor score estimate with a calculated internal consistency estimate of .86 (Table 8).

Appendix B (Continued)

Table 8.

Means and Standard Deviations for Factor Scores by Booklet Form

Booklet Form	Factor Scores								
	Effectiveness (a = .84)		Likeability (a = .86)		Professional Traits (a = .82)		Dominance (a = .82)		
	N	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Personal Situation/Autocratic Female	8	3.13	1.5	2.85	1.0	3.24	0.4	3.38	1.0
Personal Situation/Democratic Female	12	3.72	0.9	3.93	0.8	3.79	0.7	3.12	0.6
Task Situation/Democratic Female	5	3.60	0.5	2.95	0.8	3.49	0.7	3.65	0.8
Task Situation/Autocratic Female	6	2.60	0.8	2.13	0.6	3.23	0.5	4.38	0.7
Personal Situation/Democratic Male	7	3.52	0.7	4.02	1.0	3.32	0.7	2.64	0.6
Personal Situation/Autocratic Male	7	3.24	0.9	2.64	0.8	3.29	0.4	3.96	0.5
Task Situation/Autocratic Male	7	3.14	1.1	1.79	0.9	2.90	0.5	4.14	0.4
Task Situation/Democratic Male	8	3.33	0.8	3.25	0.7	3.52	0.7	3.50	0.4

Leader character traits. Initially, all 15 character traits (powerful/powerless, timid/forceful, hardworking/lazy, persistent/gives up easily, soft/tough, fair/not fair, responsible/irresponsible, not helpful/helpful, cooperative/not cooperative, trustworthy/untrustworthy, aggressive/not aggressive, independent/dependent, dominant/submissive, objective/subjective, unprepared/prepared) were combined to obtain one factor score estimate. The calculated internal consistency estimate for this factor was .72. After closer inspection, the following four character attributes (timid/forceful, soft/tough, aggressive/not aggressive, and dominant/submissive) contributed negatively to the item total correlation. As a result, these four items were

Appendix B (Continued)

separated from the rest of the items and treated as a separate factor. Internal consistency estimate for the four item dominant factor score was calculated to be .82. The remaining eleven items were combined to form the leader professional traits factor score and the calculated internal consistency estimate for this factor score was .82. Table 8 shows the means and standard deviations for all four factor scores for each of the eight booklet forms.

Evaluation of Student Feedback

Overall, the directions provided by the instructor/author were adequate, that is, students who participated in this study, followed the directions without difficulty. However, the most frequently mentioned comment students had was that the content of the scenario did not hold enough information in order for them to for example answer questions pertaining to professor future promotions etc. As a result of these comments, the author will add a neutral statement with regard to the professor's career at the University.

Summary

The purpose of this pilot study was to adjust for possible procedural difficulties in data collection, as well as examining internal consistency estimates after combining proposed items for leader effectiveness, leader likeability, leader work attitude, and leader dominance into factor score estimates. Results of this study revealed that the calculated internal consistency estimates (Cronbach alpha's) were stable. Further, based

Appendix B (Continued)

on the results of student feedback, the scenarios were adjusted to include sufficient information about the professor in order to make judgments about his/her future.

Appendix C. Task Vignette Booklets

Scenario of Student - Professor Interaction

If at any point you become uncomfortable with this attitude scale, you may discontinue.

DO NOT write your name on this form. All your responses are anonymous.

DIRECTIONS:

Please read inside this booklet the description of an interaction between a student and professor. After reading the scenario go to the next page where you will see a list of descriptors for the professor and the scenario. Each descriptor consists of a pair of extreme opposites separated by seven spaces. Consider each descriptor and mark one of the spaces between each pair of opposites that is closest to your perceptions. Please mark every descriptor **using your best judgement**.

EXAMPLES

If you feel the person *is probably* competent, you would mark a space near the word competent.

Competent ___ I ___ ___ ___ ___ ___ Incompetent

If you feel the person *is definitely not* competent, you would mark a space nearest the word incompetent.

Competent ___ ___ ___ ___ ___ ___ I Incompetent

In addition, please provide the biographical information requested on the last page.
Thank you.

Appendix C (Continued)

<p style="text-align: center;">Scenario Autocratic Female Professor in Task Situation</p>

The student knocks on Professor Smith's door. "Do you have a minute?"

Professor Smith frowns while looking at her watch, "All right, what is it?" The student steps into her office, "Do you remember last class you encouraged our group to participate more during class?" Professor Smith, sitting straight in her chair, continues to stare at the student. "Well," the student continues, "I think the problem is that we are so busy writing down information that we cannot concentrate on your lecture. So, if we could make copies of your overheads before each class it would make it much easier for us to participate." Leaning back in her chair, Professor Smith lowers her eyebrows. Slowly she shakes her head, points her finger at the student and declares, "You want me to copy my notes for you before each class?" "Oh no," the student replies, "One of the students would be responsible for all the copies." Still staring at the student, Professor Smith stands up abruptly and motions the student to the door. "I'll let you know by Monday."

Appendix C (Continued)

<p style="text-align: center;">Scenario Democratic Female Professor in Task Situation</p>

The student knocks on Professor Smith's door. "Do you have a minute?"

Professor Smith smiles and nods while glancing at her watch, "All right, what is it?" The student steps into her office, "Do you remember last class you encouraged our group to participate more during class?" Professor Smith leans forward in her chair, nods and smiles encouragingly. "Well," the student continues, "I think the problem is that we are so busy writing down information that we cannot concentrate on your lecture. So, if we could make copies of your overheads before each class it would make it much easier for us to participate."

Professor Smith gently tugs at her collar and nods her head while she is listening to the student. Then she asks, "You want me to copy my notes for you before each class?" "Oh no," the student replies, "One of the students would be responsible for all the copies." Professor Smith now smiles and gets up to accompany the student to the door while she says, "I'll let you know by Monday."

Appendix C (Continued)

Scenario**Autocratic Male Professor in Task Situation**

The student knocks on Professor Smith's door. "Do you have a minute?"

Professor Smith frowns while looking at his watch, "All right, what is it?" The student steps into his office, "Do you remember last class you encouraged our group to participate more during class?" Professor Smith, sitting straight in his chair, continues to stare at the student "Well," the student continues, "I think the problem is that we are so busy writing down information that we cannot concentrate on your lecture. So, if we could make copies of your overheads before each class it would make it much easier for us to participate." Leaning back in his chair, Professor Smith lowers his eyebrows. Slowly he shakes his head, points his finger at the student and declares, "You want me to copy my notes for you before each class?" "Oh no," the student replies, "One of the students would be responsible for all the copies." Still staring at the student, Professor Smith stands up abruptly and motions the student to the door, "I'll let you know by Monday."

Appendix C (Continued)

Scenario**Democratic Male Professor in Task Situation**

The student knocks on Professor Smith's door. "Do you have a minute?"

Professor Smith smiles and nods while glancing at his watch, "All right, what is it?" The student steps into his office, "Do you remember last class you encouraged our group to participate more during class?" Professor Smith leans forward in his chair nods and smiles encouragingly. "Well," the student continues, "I think the problem is that we are so busy writing down information that we cannot concentrate on your lecture. So, if we could make copies of your overheads before each class it would make it much easier for us to participate."

Professor Smith gently tugs at his collar and nods his head while he is listening to the student. Then he asks, "You want me to copy my notes for you before each class?" "Oh no," the student replies, "One of the students would be responsible for all the copies." Professor Smith now smiles and gets up to accompany the student to the door while he says, "I'll let you know by Monday."

Appendix C (Continued)

DEPENDENT MEASURES**How would you describe the professor?**

competent	___	___	___	___	___	incompetent
effective	___	___	___	___	___	ineffective
well qualified	___	___	___	___	___	not at all qualified
critical	___	___	___	___	___	tolerant
considerate	___	___	___	___	___	inconsiderate
popular	___	___	___	___	___	unpopular
likeable	___	___	___	___	___	not likeable
powerful	___	___	___	___	___	powerless
timid	___	___	___	___	___	forceful
hardworking	___	___	___	___	___	lazy
persistent	___	___	___	___	___	gives up easily
soft	___	___	___	___	___	tough
fair	___	___	___	___	___	not fair
responsible	___	___	___	___	___	irresponsible
not helpful	___	___	___	___	___	helpful
cooperative	___	___	___	___	___	not cooperative
trustworthy	___	___	___	___	___	untrustworthy
aggressive	___	___	___	___	___	not aggressive
independent	___	___	___	___	___	dependent
dominant	___	___	___	___	___	submissive
objective	___	___	___	___	___	subjective
unprepared	___	___	___	___	___	prepared

In your opinion the student in this scenario is

Male

Female

(Please circle one)

Appendix C (Continued)

Directions: Below are statements describing the professor. Please indicate the extent of your agreement or disagreement by circling your response in the box which best corresponds to your beliefs.

		Strongly Agree	Agree	Uncertain	Disagree	Strongly Disagree
1.	The professor is handling the situation well.	SA	A	U	D	SD
2.	The professor has potential for future promotions.	SA	A	U	D	SD
3.	The professor's reaction to the student is appropriate.	SA	A	U	D	SD
4.	The professor has great communication skills.	SA	A	U	D	SD
5.	Students respect this professor.	SA	A	U	D	SD
6.	This professor is well liked by students.	SA	A	U	D	SD

Please check the following information about yourself.

Gender M F
Age _____

School Standing Freshman Sophomore
 Junior Senior
 Graduate Other (specify)

Major: _____
(Please write in)

Race/Ethnicity: American Indian Asian/Pacific Islander
 Black (African American) Hispanic
 White (Non Hispanic) Other (please specify): _____

Appendix D. Personal Vignette Booklets

Scenario of Student - Professor Interaction

If at any point you become uncomfortable with this attitude scale, you may discontinue.

DO NOT write your name on this form. All your responses are anonymous.

DIRECTIONS:

Please read inside this booklet the description of an interaction between a student and professor. After reading the scenario go to the next page where you will see a list of descriptors for the professor and the scenario. Each descriptor consists of a pair of extreme opposites separated by seven spaces. Consider each descriptor and mark one of the spaces between each pair of opposites that is closest to your perceptions. Please mark every descriptor **using your best judgement**.

EXAMPLES

If you feel the person *is probably* competent, you would mark a space near the word competent.

Competent ___ I ___ ___ ___ ___ ___ Incompetent

If you feel the person *is definitely not* competent, you would mark a space nearest the word incompetent.

Competent ___ ___ ___ ___ ___ ___ I Incompetent

In addition, please provide the biographical information requested on the last page.
Thank you.

Appendix D (Continued)

<p style="text-align: center;">Scenario Autocratic Female Professor in Personal Situation</p>

The student knocks on Professor Smith's door. "Do you have a minute?" Sitting at her computer, Professor Smith does not take her eyes off her computer screen, but continues working on it, as she impatiently says, "Come in, come in." Still working on the computer she asks, "What seems to be the problem?" The student slumps into a chair and looks down on the ground. Professor Smith glances over at the student, sighs, crosses her arms over her chest and remains silent. The student says, "My long-term relationship just fell to pieces and I am so upset over it that I really don't know what to do." Staring at the student, Professor Smith finally says, "I don't know how to help you." Professor Smith checks to see whether her files on the computer are saved. To be sure, she hits the save button again. Then she turns back to the student and remarks, "I can see this will take more time than we have right now. Let's make an appointment."

Appendix D (Continued)

<p style="text-align: center;">Scenario Democratic Female Professor in Personal Situation</p>

The student knocks on Professor Smith's door. "Do you have a minute?" Sitting at her computer, Professor Smith gets up from her desk, nods an encouraging hello to the student and says, "Come in, come in." Professor Smith sits in a chair opposite the student and asks, "What seems to be the problem?" The student slumps into a chair and looks down on the ground. Professor Smith leans forward toward the student, folds her arms in her lap and patiently waits for the student. The student says, "My long-term relationship just fell to pieces and I am so upset over it that I really don't know what to do." Looking compassionately at the student, Professor Smith gently replies, "I don't know how to help you." Professor Smith remains seated across from the student while she lets some time pass. "I can see this will take more time than we have right now. Let's make an appointment."

Appendix D (Continued)

<p style="text-align: center;">Scenario Autocratic Male Professor in Personal Situation</p>

The student knocks on Professor Smith's door. "Do you have a minute?" Sitting at his computer, Professor Smith does not take his eyes off his computer screen, but continues working on it as he impatiently says, "Come in, come in." Still working on the computer he asks, "What seems to be the problem?" The student slumps into a chair and looks down on the ground. Professor Smith glances over at the student, sighs, crosses his arms over his chest and remains silent. The student says, "My long-term relationship just fell to pieces and I am so upset over it that I really don't know what to do." Staring at the student, Professor Smith finally says, "I don't know how to help you." Professor Smith checks to see whether his files on the computer are saved. To be sure, he hits the save-button again. Then he turns back to the student and remarks, "I can see this will take more time than we have right now. Let's make an appointment."

Appendix D (Continued)

<p style="text-align: center;">Scenario Democratic Male Professor in Personal Situation</p>

The student knocks on Professor Smith's door. "Do you have a minute?" Sitting at his computer, Professor Smith gets up from his desk, nods an encouraging hello to the student and says, "Come in, come in." Professor Smith sits in a chair opposite the student and asks, "What seems to be the problem?" The student slumps into a chair and looks down on the ground. Professor Smith leans forward toward the student, folds his arms in his lap, and patiently waits for the student. The student says, "My long-term relationship just fell to pieces and I am so upset over it that I really don't know what to do." Looking compassionately at the student, Professor Smith gently replies, "I don't know how to help you?" Professor Smith remains seated across from the student while he lets some time pass. He finally says in a comforting voice, "I can see this will take more time than we have right now. Let's make an appointment."

Appendix D (Continued)

INDEPENDENT MEASURES**How would you describe the professor?**

likable	___	___	___	___	___	not likeable
soft	___	___	___	___	___	tough
competent	___	___	___	___	___	incompetent
trustworthy	___	___	___	___	___	untrustworthy
responsible	___	___	___	___	___	irresponsible
effective	___	___	___	___	___	ineffective
critical	___	___	___	___	___	not likeable
unprepared	___	___	___	___	___	prepared
aggressive	___	___	___	___	___	not aggressive
well qualified	___	___	___	___	___	not at all qualified
persistent	___	___	___	___	___	gives up easily
objective	___	___	___	___	___	subjective
hardworking	___	___	___	___	___	lazy
considerate	___	___	___	___	___	inconsiderate
dominant	___	___	___	___	___	submissive
not helpful	___	___	___	___	___	helpful
cooperative	___	___	___	___	___	not cooperative
powerful	___	___	___	___	___	powerless
timid	___	___	___	___	___	forceful
independent	___	___	___	___	___	dependent
fair	___	___	___	___	___	not fair
popular	___	___	___	___	___	unpopular
capable	___	___	___	___	___	not capable
influential	___	___	___	___	___	not influential

In your opinion the student in this scenario is**Male****Female**

Appendix D (Continued)

Directions: Below are statements describing the professor. Please indicate the extent of your agreement or disagreement by circling your response in the box which best corresponds to your beliefs.

		Strongly Agree	Agree	Uncertain	Disagree	Strongly Disagree
1.	The professor is handling the situation well.	SA	A	U	D	SD
2.	The professor has potential for future promotions.	SA	A	U	D	SD
3.	The professor's reaction to the student is appropriate.	SA	A	U	D	SD
4.	The professor has great communication skills.	SA	A	U	D	SD
5.	Students respect this professor.	SA	A	U	D	SD
6.	This professor is well liked by students.	SA	A	U	D	SD

Please check the following information about yourself.

Gender M F
Age _____

School Standing Freshman Sophomore
 Junior Senior
 Graduate Other (specify)

Major: _____
(Please write in)

Race/Ethnicity: American Indian Asian/Pacific Islander
 Black (African American) Hispanic
 White (Non Hispanic) Other (please specify): _____

Appendix E. Survey Introduction Letter

Monday, August 06, 2001

My name is Michela LaRocca, and I am a Doctoral Candidate in the Department of Measurement and Research within the College of Education at the University of South Florida. I have successfully defended my dissertation proposal, and have received Institutional Review Board (IRB) approval to conduct an experiment designed to examine how male and female students perceive leadership traits in male and female professors.

The actual experiment consists of versions of a written scenario, which is academic in nature. Leader in this case is defined in the role of the professor. Versions were created in order to vary professor gender, professor behavior, and the situation in which an interaction is described between the professor and one of the students in the class. Extensive pilot testing was conducted to verify the manipulations of the variables, as well as the outcome measures.

Student participation is voluntary. Pilot studies indicated that most students were able to follow the instructions on the front page of the booklet without incidents. The least disruptive method was to distribute the booklets before class. Students will likely take about 8-10 minutes reading the scenario, and filling out the questions.

I should be able to have results available for instructors interested by the summer of 2002.

This research is very important for understanding how males and females are perceived in leadership roles.

Thank you so much for your consideration, and if you have any questions, please feel free to contact me via phone: (813) 792-8501, or via email: michela1@tampabay.rr.com

About the Author

Michela A. LaRocca earned a Bachelor of Arts degree in Psychology from University of Kentucky in 1986, and a Master of Arts in Clinical Psychology from Western Kentucky University in 1989. While earning the PhD. in measurement and research, Ms LaRocca worked as a statistical consultant to other researchers.

Ms. LaRocca co-authored three publications (she was first author on one publication). Her research focuses on the development of attitude scales. In addition, Ms. LaRocca examines theoretical aspects of gender roles as they pertain to current gender issues.