This is to certify that the
dissertation entitled

GENDER STEREOTYPES AND 360-DEGREE LEADERSHIP FEEDBACK RATINGS: AN APPLICATION OF ROLE CONGRUITY THEORY

presented by

CORI A. DAVIS

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Psychology

J. Kevin Ford
Major Professor's Signature

August 22 2003
Date

GENDER STEREOTYPES AND 360-DEGREE LEADERSHIP FEEDBACK RATINGS: AN APPLICATION OF ROLE CONGRUITY THEORY

By

Cori A. Davis

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Psychology

2003

ABSTRACT

GENDER STEREOTYPES AND 360-DEGREE LEADERSHIP FEEDBACK RATINGS: AN APPLICATION OF ROLE CONGRUITY THEORY

By

Cori A. Davis

This study examined gender differences in a 360-degree leadership development assessment. First, consistent with Role Congruity Theory (Eagly & Karau, 2002), the context of the work environment was expected to impact the relationship between leader gender and leader ratings such that female leaders would be rated as being less effective and needing more development than their male counterparts in male-dominated work contexts. The types of context variables explored were the level of male domination of the industry and functional areas within which the leaders worked, and the proportion of male raters for each leader. Second, gender differences in ratings were expected for different types of skills being assessed. In this study, both communal and agentic leadership skills were included in the 360 instrument. It was expected that women would generally receive ratings indicating higher effectiveness and less need for development with regard to communal, as opposed to agentic, leadership skills. Finally, the impacts of rater gender, rater level relative to the leader (i.e. boss, peer, subordinate), and leader level on leadership ratings (including self-ratings) were explored.

No gender differences were found for any type of rating: overall leadership effectiveness, agentic development needs, or communal development needs. Further, gender differences in ratings were not impacted by the organizational context variables included in this study, thereby failing to support Role Congruity Theory. Finally, no gender differences were found as a function of rater gender, rater level, or leader level.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

LIST OF TABLES

LIST OF TABLES (CONTINUED)

LIST OF FIGURES

INTRODUCTION

Over the past three decades, the percentage of female managers in the workforce has increased from 18% to 45% (U. S. Bureau of Labor Statistics, 2001a). In addition, women are earning 56% of Bachelors degrees and 45% of advanced degrees including 42% of PhDs and 43% of professional degrees (e.g. law and medicine; U. S. Bureau of the Census, 2000). Yet, despite the growing number of educated women in the workforce, women constitute only 4% of the top officers in Fortune 500 companies, 3% of the most highly paid officers, and 0.4% of CEOs (Catalyst, 2000).

The "pipeline" problem is a common response to the huge disparity between the number of educated, working women and the number of women in leadership positions. The pipeline explanation states that there are simply not enough women with the appropriate education and background to fill top leadership positions. However, this explanation is obviously inconsistent with the statistics on the current number of educated, female managers in the United States (Carli & Eagly, 2001).

Heilman (1995) also suggests that the pipeline explanation is insufficient. She notes that research indicates that women who moved into management in the 1970's have not progressed through the ranks as quickly as males who entered management at around the same time. In fact, a report by the Feminist Majority Foundation states that at the present rate of advancement, it will take until the year 2465 for women to be as equally represented as men in executive suites (1991; cited in Morrison, White, & Van Velsor, 1992) despite research evidence that there are few, if any, actual gender differences in qualifications to lead (see Dobbins & Platz, 1986; Morrison & Von Glinow, 1990; Powell, 1993; Stroh, Brett, & Reilly, 1992; Cleveland, Stockdale, & Murphy, 2000).

This indicates that despite progression in education and movement to management positions, women are still experiencing problems scaling the hierarchy (Bartol, 1999). This holds true even for women in female-dominated fields. In fact, some research has shown that men are promoted even more quickly in female-dominated as opposed to in male-dominated fields (e.g. Maume, 1999; Ott, 1989; Williams, 1992, 1995; Yoder & Sinnett, 1985).

If the pipeline explanation cannot explain why there are so few women in leadership positions, what can? There have been many explanations proposed in response to this question. Kaufmann, Isaksen, and Lauer (1996), for example, tested three different explanations for this phenomenon which they called Experience Theory, Personality Theory, and Stereotype Theory. The first two theories state that women simply have not yet built up the qualifying experiences or do not possess the traits necessary to becoming successful leaders. The last theory, Stereotype Theory, states that the glass ceiling is product of attitudes toward women. They found the most support for Stereotype theory.

In fact, most explanations for why more women are not in the executive suite stem from Stereotype Theory. The theory, when applied to gender, generally states that individuals have a preconceived notion about what actions are appropriate for men and women. Although some may argue that gender stereotypes are diminishing, particularly in the workplace, recent Gallop Poll responses confirm that stereotypical attitudes regarding leadership and gender still exist. For instance, when it comes to choosing a boss, Americans still prefer men (Simmons, 2001: January 11). In addition, Americans

see women as emotional and affectionate, and men as more aggressive (Newport, 2001: February 21).

Powell, Butterfield, and Parent (2002) performed a recent study which assessed attitudes about women managers. They gathered data from 348 undergraduate and part-time graduate business students. They found that although stereotypes about managers place less emphasis on masculine characteristics than in earlier studies, a good manager is still perceived as having a predominately masculine skill set (i.e. autocratic, decisive, assertive, etc.) (Powell et al., 2002).

Heilman (2001) has noted that the glass ceiling is a natural consequence of gender stereotypes and the subsequent expectations stereotypes produce about what women are like and how they should act. She also proposed that stereotypes in the workplace often manifest in evaluation bias. She claims that bias in performance evaluation is the primary cause for the glass ceiling because when women receive inadequate leadership performance ratings, they are less likely to be selected for leadership positions. Bartol (1999) also noted that an important area of study relating to the advancement of women in organizations is whether there are systematic gender differences in performance evaluations. Unfortunately, there have been few consistent findings with regard to the direct effect of leader gender on performance ratings.

Several laboratory studies have suggested that ratee gender has little to no influence on performance evaluations (e.g., Frank & Drucker, 1977; Giannantonio, Olian, & Carroll, 1995; Hall & Hall, 1976), whereas others have found significant gender effects sometimes in favor of women (e.g., Bigoness, 1976; Hamner, Kim, Baird, & Bigoness, 1974), and sometimes in favor of men (e.g., Pazy, 1986; Woehr & Roch, 1996). As with

these lab studies, field studies have not led to a clear conclusion about the impact of leader gender on performance evaluations.

Bartol (1999) has noted that there are relatively few field studies considering ratee gender effects in official performance appraisals, particularly at the managerial levels. Some field studies found that men were evaluated more positively than women (e.g. Deaux, 1979; Eagley, Makhijani, & Klonsky, 1992; Pulakos, White, Oppler, & Borman, 1989), others found that women were rated more highly than were men (e.g. Bennett, 1982; Mobley, 1982; Pulakos & Wexley, 1983), and still others found equal ratings between men and women (e.g. Kirnan, Farley, & Geisinger, 1989; Yammarino & Dubinsky, 1988). The inconsistent results of these studies may suggest that more complex issues are involved (Bartol, 1999).

Some researchers have noted that gender differences in performance evaluation do not always explain the disparity between men and women in the top ranks. There is also evidence that female employees receive less developmental opportunities than their male counterparts. For instance, females are less likely to have mentors, and more likely to experience barriers to gaining a mentor (Ragins & Cotton, 1991). In addition, women are less likely to be offered challenging experiences that stretch their skills (Lyness, 2002). Lyness (2002) notes that high-quality organizational career management and assessment systems, including coaching and feedback, can help to offset some of the developmental disadvantages women face. However, if bias exists in these sorts of assessment systems, then they are not serving their purpose.

Given the already disparate developmental opportunities afforded women, it is important to understand how gender stereotypes impact current leadership development

systems. An increasing number of organizations are moving toward a leadership development model that includes some form of multisource performance assessment, often referred to as 360-degree feedback (Atwater & Waldman, 1998; Becker, Ayman, & Korabik, 2002; London & Smither, 1995). Such systems usually incorporate performance and development evaluation feedback from a variety of sources, including subordinates, peers, and bosses.

Antonioni (1996) reported that approximately 25% of companies use some type of multirater or 360 degree feedback process. One of the reasons for its increased popularity is the belief that developmental feedback will improve leadership performance. Studies have, in fact, found improvements in overall performance as a result of 360 feedback (e.g. Atwater, Roush, & Fischthal, 1995; Reilly, Smither, & Vasilopoulos, 1996).

Although there is a growing body of literature on 360 feedback, little is known about whether gender differences that are said to occur in performance appraisals carry over into 360 developmental feedback assessments. Do these same differences exist when individuals receive negative 360 feedback that is confidential, developmental, and not used for reward or promotion decisions? In addition, relatively little is known regarding gender effects related to evaluations provided by sources other than supervisors within a 360 feedback system (Bartol, 1999).

Given that a 360 assessment is essentially an informal performance evaluation from every type of person, using this type of developmental tool raises potential questions about possible gender bias. There is an opportunity here to explore how past findings regarding gender differences in performance evaluation can be generalized to

360 assessment systems. In addition, we need to explore what unique questions need to be answered regarding the role of stereotypes in the context of a 360 feedback system. Therefore, the purpose of this research is to explore the theories which help to explain gender differences in performance evaluation and generalize and test those theories in the context of a 360 leadership development tool. The next few sections review Gender Stereotype theories more specifically as well as the major findings in the field of gender and leadership. Hypotheses are then offered relevant to the theories and findings presented.

LITERATURE REVIEW

Gender Stereotype Theories

There are basically two forms of thought surrounding the origin of stereotypes. There is the biological view and the social view. Those who support the biological view claim that the differences between men and women are a byproduct of evolution and are rooted in physiological characteristics that have evolved over time. Alternatively, those supporting the social perspective support the claim that differences between men and women stem from societal values and subsequent manufactured gender roles. With regard to examining stereotypes in the workplace, the social perspective has received the most attention and is most relevant for this study.

Social Role Theory. Some claim that the workplace is an extension of our social environment, and therefore spillover effects of male/female social stereotypes affect how men and women are treated at work (Gutek & Morasch, 1982). Eagly and colleagues have made the strongest arguments for this viewpoint using Social Role Theory. Social Role Theory suggests that gender stereotypes are similar to other social stereotypes in that they reflect "perceivers' observations of what people do in daily life" (Eagly & Steffen, 1984; p. 735). In other words, because we observe particular groups of people performing specific behaviors, we are prone to believing that the abilities and personal qualities required for performing those behaviors are typical of those groups. A key proposition of social role theory is that most of the behaviors we ascribe to men and women fall into two categories: agentic and communal (Bakan, 1966; Eagly, 1987). Men are stereotypically seen as displaying agentic behaviors, and women are seen as displaying communal behaviors.

Communal characteristics generally describe a concern with the welfare of other people or as social or service-oriented. Examples include being helpful, affectionate, kind, sympathetic, sensitive, gentle, and nurturing. Agentic characteristics generally describe an achievement-oriented, assertive, controlling, and confident orientation. Examples include being ambitious, aggressive, dominant, forceful, independent or self-sufficient, self-confident, decisive, and prone to act as a leader. (Eagly & Karau, 2002; Heilman, 2001). Not only are these agentic and communal conceptions of women and men different, but they are also often oppositional. Members of one sex are seen as lacking what is considered to be prevalent in members of the other sex (Heilman, 2001).

Heilman (2001) has noted that these agentic and communal social stereotypes about the attributes of men and women are pervasive, widely shared, and resistant to change (see Dodge, Gilroy, & Fenzel, 1995; Leuptow, Garovich, & Leuptow, 1995). Some researchers have found that gender provides the strongest basis for categorizing people, even when compared with other personal characteristics such as race, age, and occupation (Fiske, Haslam, & Fiske, 1991; Eagly & Karau, 2002; Stangor, Lynch, Duan, & Glass, 1992; van Kippenberg, van Twuyver, & Pepels, 1994). In addition, some have noted that stereotypes about women and men are easily and automatically activated (Banaji & Hardin, 1996; Banaji, Hardin, & Rothman, 1993; Blair & Banaji, 1996; Eagly & Karau, 2002).

Research also supports the notion that traditional stereotypes of women and men predominate in work settings as well as in non-work settings. Heilman, Block, and Martell (1995), for example, showed that even when women are depicted as managers, they are characterized as less agentic than men. In their study, working managers from a

range of industries described women managers as more competent, active, and potent than women in general. However, they described women managers as decidedly more deficient in these same attributes than male managers. Heilman et al. (1995) concluded that the increased presence of women in the workplace and their assumption of new roles do not appear to preclude gender-stereotypic perceptions. This led Heilman to propose the Lack of Fit model (Heilman, 1983, 1995, 2001).

Lack of Fit Model. Heilman (1983, 1995, 2001) developed the Lack of Fit Model to help specify the circumstances in which gender discrimination occurs. Her model suggests that the way individuals perceive job roles may be a crucial component in determining the extent to which gender influences work-related outcomes (Cleveland et al., 2000).

In Heilman's (2001) view, some jobs are more closely associated with gender stereotypes than others. This may be due to the typical number of males or females employed in these roles, or to the types of activities performed on the job (Cleveland et al., 2000). For instance, common stereotypes exist for nurses and police. When most people think of a nurse, they think of a woman. This may be because most nurses are, in fact, female. This may also be because nurses care for people, and many think that women are more likely than men to have care-giving roles. Similarly, when people think of a police officer, they usually picture a man. Again, this stereotype may arise because most police officers are, in fact, men. It may also arise because police officers enforce the law, and many think that men are more likely than women to hold an authority-type role.

Heilman (1983, 1995, 2001) also suggests that certain jobs may carry gender stereotypes because of personal qualities considered to be necessary for success. With regard to management and leadership, she noted that success is often associated with ambitiousness, competitiveness, analytical skills, and interest in power (Cleveland, et al., 2000). These are all agentically-oriented qualities. Alternatively, women are typically considered to have an entirely different set of skills including interpersonal skills and interest in consensus (Cleveland et al., 2000). Although these are also work-related, they are not skills usually associated with the stereotypic leader.

The lack of fit model implies that the role of a leader may be problematic for women because the schemas, or preconceived ideas, that people have regarding leaders/managers are different from those they have regarding women. There is much evidence to suggest that the prevailing image of a "leader" is more similar to that of a man than a woman. This has been found to be true for respondents who are young and old, male and female, and is generalizable across culture and time (Becker et al., 2002; Schein, 1973, Schein, Mueller, & Jacobson, 1989; Schein, Mueller, Lituchy, & Liu, 1996). A woman in a leadership role likely activates two conflicting schemas: a feminine, or communal schema associated with her gender and a masculine, or agentic schema associated with leader role (Becker et al., 2002).

Beginning in the 70's, researchers have repeatedly shown that many traits associated with leadership (e.g. competitiveness, self-confidence, and desire for responsibility) are seen as male-oriented or agentic traits (Brenner, Tomkiewicz, & Schein, 1989; Heilman, Block, Martell, & Simon, 1989; Powell & Butterfield, 1979, 1989, 1994; Schein, 1973, 1975). For example, in early studies by Schein (1973, 1975),

male and female managers gave their impression of either women, men, or successful middle managers. The participants perceived successful middle managers as being significantly more similar to men than women. This was particularly true for a number of agentically-oriented characteristics such as competitiveness, self-confidence, objectivity, aggressiveness, and ambitiousness. Other researchers have replicated Schein's (1973, 1975) findings in the United States (e.g. Heilman et al., 1989; Massengill & di Marco, 1979) as well as in the United Kingdom, Germany, Japan, China, and Singapore (Lee & Hoon, 1993; Schein, 2001). These studies show that the typical perceptions of the leadership role overlaps with the male stereotype (Cleveland et al., 2000).

Based on the incongruence between the stereotype of a leader and the stereotype of a woman, Heilman (1983, 1995) suggested that gender discrimination is most likely to occur when the qualities of the person do not fit with the stereotype of the role (Cleveland et al., 2000). Heilman suggested that this effect should hold true for both women and men. In other words, both women in masculine roles and men in feminine roles are more likely to experience gender discrimination than women in feminine roles and men in masculine roles, respectively. However, wage surveys suggest that women are more likely than men to be penalized in sex-role incongruent situations (Cleveland et al., 2000; Renner, Rives, & Bowlin, 2002).

Role Congruity Theory. Recently, Eagly and Karau (2002) presented their Role Congruity Theory which builds on the ideas in Social Role Theory and the Lack of Fit Model. The theory itself integrates the social psychology perspectives regarding stereotyping with the industrial/organizational psychology perspectives on perceptions of leadership and management roles.

There is an emphasis in Role Congruity Theory on the "gender" role as opposed to the "social" role discussed in Social Role Theory. Where social roles are socially shared expectations applied to persons belonging to a certain social position or social category (Biddle, 1979; Sarbin & Allen, 1968), gender roles are consensual beliefs about the attributes of men and women (Eagly & Karau, 2002). Role Congruity Theory contrasts with classic stereotype theories which portray prejudice as arising from unfavorable stereotypes toward a social group (e.g., Allport, 1954; Esses, Haddock, & Zanna, 1993). From that perspective, the lack of women in top leadership positions would be attributed to a negative stereotype toward women in general. However, such a context-free theory could not explain why there is bias against women in some roles and not others (Eagly & Karau, 2002).

The basic philosophy behind Role Congruity Theory, then, is that women in organizational leadership roles are perceived as being role incongruent because the role of a leader is inherently agentic. Eagly and Karau (2002) claim that "the potential for prejudice exists when social perceivers hold a stereotype about a social group that is incongruent with the attributes that are thought to be required for success in certain classes of social roles. When a stereotyped group member and an incongruent social role become joined in the mind of the perceiver, this inconsistency lowers the evaluation of the group member as an actual or potential occupant of the role" (Eagly & Karau, 2002; p. 574). Therefore, Role Congruity Theory suggests that bias toward female leaders generally follows from the incongruity that individuals perceive between the characteristics of women and the requirements of a leadership role.

According to Role Congruity Theory, the perceived incongruity between leadership roles and the female gender role leads to two types of bias against female leaders. First, individuals may perceive women less favorably than men as potential occupants of leadership roles. Second, individuals may evaluate behavior that fulfills the prescriptions of a leader role less favorably when these behaviors are performed by a woman. Consequences of these forms of bias include less favorable attitudes toward female as opposed to male leaders, and increased difficulty for women trying to advance into leadership positions. In presenting their Role Congruity Theory, Eagly and Karau (2002) reviewed evidence from various studies that substantiates these proposed consequences.

One of the main findings Eagly and Karau (2002) reported is that context variables of the work environment which make gender more salient are likely to enhance biases against female leaders. Specifically, women working in male-dominated settings will experience more bias. There are some contexts, at multiple levels of analysis, that have been shown to be important in affecting female leader biases. Three of these contexts include industry, job functional area, and gender composition of evaluators. These context variables are being described in the current study for a few reasons. First, they each represent different levels of interest. Industry can be investigated across organizations, functional area can be investigated both across and within organizations, and the gender composition of evaluators is specific to each individual. The argument can be made that even if women are working in agentic industries, they still may be working in a communal job role mainly surrounded by other women. Second, no single

study has ever included multiple context variables at varying levels of analysis to determine which context variable has the most impact on female rating bias.

With regard to industry, Scandura (1991) performed a study including 176 female executives from service industries, manufacturing, government, retail industries, and small businesses. Eighty percent of the women in this sample came from small companies and most were in service and retail industries. Thus, Scandura (1991) concluded that the glass-ceiling breakers were unlikely to come from the largest, most powerful organizations and were more likely to come from female-dominated industries. This study was limited in the number of industries explored and provided no information about the possible mechanisms preventing women in large companies and/or male dominated industries from reaching the top leadership positions.

With regard to functional area, research has found that males are more likely to become leaders when the tasks being performed are male-oriented (e.g. assembling mechanical components) (Butterfield & Grinnell, 1999; Carbonell, 1984; Megargee, 1969; Nyquist & Spence, 1986). Studies from the Goldberg research paradigm also tested whether women suffer disadvantages relative to men. The Goldberg studies typically consisted of experiments in which participants evaluate men and women whose characteristics have been set equal. This method of examining potential bias against women has been labeled the Goldberg Paradigm in honor of Goldberg's (1968) initial experiment of this nature (Eagly & Karau, 2002). A typical Goldberg study would involve presenting a resume or application for evaluation, with half of the participants receiving it with a female name and half receiving it with a male name (e.g. Rosen & Jerdee, 1974). In a thorough meta-analysis of a subset of Goldberg studies, Davison and

Burke (2000) showed that men were preferred over women for male gender-typed jobs and women were preferred over men for female gender-typed jobs. Given that leadership roles are usually gender-typed as masculine, this line of research supports Role Congruity Theory's prediction of bias against female candidates for leadership positions (Eagly & Karau, 2002).

The final context variable explored in this study is the gender composition of evaluators. Many of the researchers who have reported rater gender composition effects describe the composition of raters as a context variable. For instance, Bowen, Swim, and Jacobs (2000) found little evidence of overall gender bias in performance appraisals in field studies, however found significant pro-male biases when only men served as raters. Eagly and colleagues also found that men were perceived as being more effective than women to the extent that leader and subordinate roles were male-dominated numerically (Eagly, Karau, & Makhijani, 1995). In addition, other research has shown that female leaders are seen as being less effective relative to male leaders as the proportion of male subordinates increased (Eagly & Karau, 2002).

In presenting the importance of context in their Role Congruity Theory, Eagly and Karau (2002) reviewed findings from both field and lab studies. However, they chose to focus on lab studies because lab studies provide more control over contextual factors such as job tenure, interpersonal relationships between leaders and raters, and job performance (Eagly & Karau, 2002). However, although not discounting the importance of lab studies in gender-work research, other researchers have emphasized the importance of examining field data, especially when exploring issues of organizational context (Bartol, 1999).

Researchers have found that particularly in the study of gender, findings differ between lab studies and field studies such that gender effects are more likely to be found in laboratory settings (Dobbins & Platz, 1986; Osborn & Vicars, 1976). Lab studies almost always present participants with paper people or a confederate team leader with whom they have had little experience. Some researchers have criticized the lack of context, and therefore external validity, of such studies and have pointed to the need for more field research in this area (Bartol, 1999; Eagly & Karau, 2002).

Although there have been some field studies in the gender-work arena, many are limited by the samples they use. Many include middle managers only from a single functional area or industry. Martell, Parker, Emrich, and Crawford (1998) also noted that almost every study of sex stereotyping in the workplace has looked at characterizations of men and women for entry and middle levels of management, with little attention to whether those in middle management positions are perceived differently on the attributes associated with successful executives.

Bartol has also commented on the need for field research in this area. She noted that "the limited data that exist on the link between gender of rater or gender of ratee and performance evaluations in real and simulated workplaces suggest that simply using gender as a predictor does not capture the complexities of the area of inquiry" (Bartol, 1999; p. 173). She also noted that the number of studies conducted in field settings has been small and has largely concentrated on positions near the bottom of the hierarchy where women have experienced more success in being promoted. In addition, she notes that it has usually been necessary to include a broad mix of positions to attain adequate sample sizes (Bartol, 1999).

Finally, Ragins (1991) supported the view that the research setting is one artifact that contributes to inconsistent findings regarding gender effects, particularly with regard to subordinate evaluations of leaders. She notes that field studies may lead to more realistic leadership appraisals given that raters have multiple opportunities to observe leaders' behaviors and compare leaders with others in the organization.

In sum, Role Congruity Theory is derived from Social Role Theory and the Lack of Fit Model. It speaks to gender stereotypes as they pertain to leadership roles in organizational settings. Role Congruity Theory emphasizes the effect of context variables on the relationship between leader gender and performance evaluation. Although there is some support for Role Congruity Theory, there are few field studies which have successfully shown the effects of context variables such as industry, functional area, and gender composition of raters. Also, few studies have examined leader ratings in a 360 paradigm which includes evaluation of development needs. The next few sections discuss the Changing Nature of Leadership and the Use of 360 Leadership Assessments. These sections will help us to understand how current leadership theory and past findings regarding 360 assessments can help us to further determine the circumstances in which male and female leadership ratings may differ.

Changing Nature of Leadership

Although Role Congruity Theory addresses the research pertaining to context effects on the relationship between leader gender and evaluation ratings, it does not specifically address the nature of the skills being evaluated (i.e. the content of the evaluation tool). In order to understand what particular skills are valued by raters, it is important to explore current leadership theory as it pertains to gender.

Unfortunately, examining leadership in the organizational domain as it relates to gender presents a unique problem. Not only does the gender-laden nature of the organizational context need to be considered, but so does the gender-laden nature of leadership. In organizational, political, and religious domains, individuals are inundated with examples of male leaders. The typical male leader is often characterized as being agentic and therefore ideas about what constitutes good leadership tend to emphasize stereotypically masculine traits and skills.

Leaders have traditionally been viewed as strong, assertive, take-charge individuals who more or less dictate followers' actions. However, more recent trends in the leadership field highlight a mix of desired leadership skills ranging from communal to agentic. The transformational theory of leadership, particularly, incorporates a blend of both the agentic and communal leadership skills necessary for effective leadership. Investigating gender through a transformational leadership lens might allow us to better understand the types of skills and behaviors emphasized in the leadership field. In addition, the transformational leadership model has become increasingly popular in recent years and is one of the few models of leadership to be studied in relation to gender (Butterfield & Grinnell, 1999; Howell & Avolio, 1993; Howell & Higgins, 1990).

Transformational leadership is "the process of influencing major changes in the attitudes and assumptions of organizational members...and building commitment for major changes in the organization's objectives and strategies" (Yukl & Van Fleet, 1992, p. 174). Transformational leadership transcends the daily reward/punishment-based styles of leadership, and whereas more traditional styles such as transactional leadership produce incremental changes in the way followers behave, transformational leadership

produces fundamental changes in others' beliefs and attitudes. Many scholars have argued that transformational leadership is crucial for organizational effectiveness and adaptability in an age of increasing turbulence and uncertainty because it requires leaders to provide vision, share leadership (empower others), earn credibility, communicate, and promote integration and strategizing (Cleveland et al., 2000).

A few researchers have argued that some characteristics of the transformational leadership model (compared with other models such as transactional leadership) overlap with the stereotypical communal qualities of women (Maher, 1997). The transformational model emphasizes democratic relationships, participatory decision making, delegation, and team-based leadership skills (Eagly & Johnson, 1990). These characteristics of the transformational leadership model are expected to lead to positive organizational outcomes. For example, proponents of learning organizations (e.g., Garvin, 1993; Senge, 1990) emphasize supportiveness, participation, and team-based learning as key elements of organizational effectiveness. Similarly, continuous quality improvement scientists and practitioners (e.g., Deming, 1986; Juran, 1988) emphasize building cooperation, developing teamwork skills, building relationships, and empowering employees in order to improve organizational effectiveness (Eagly & Karau, 2002; Hackman & Wageman, 1995).

There is some evidence that women are more skilled in these aspects of transformational leadership than men (Denmark, 1993), although the number of studies examining gender differences in transformational styles of leadership is relatively low (Cleveland et al., 2000). However, some have studied specific leadership styles characteristic of the transformational model such as participative leadership and

consideration behaviors. For instance, Eagly and Johnson's (1990) meta-analysis of 162 studies compared the leadership styles of women and men found a tendency for women to lead in a more participative and interpersonally-oriented style (e.g. concern with maintaining interpersonal relationships by tending to others' morale and welfare...i.e. helping, explaining, friendly, available) than men. They concluded that female leaders may be more interpersonally-oriented and democratic whereas male leaders may be more task-oriented and autocratic (concern with accomplishing assigned tasks by organizing task-relevant activities...i.e. rules, procedures, high performance, definite roles).

Despite this evidence that women actually perform some leadership skills better than men, proficiency at communal-type skills does not guarantee others' perceptions of overall effectiveness as a leader. Recent studies in the area of gender and leadership style have found that "women in performance settings face a catch-22" (Rudman & Glick, 2001; p. 743). Female leaders who display communal qualities are oftentimes liked, but can be perceived as incompetent. Alternatively, women who display agentic qualities in order to be perceived as qualified for leadership roles may be perceived as competent, but may also suffer social repercussions (Glick, Zion, & Nelson, 1988; Rudman, 1998; Rudman & Glick, 2001). Research has shown that agentic women are not liked as much as their agentic male counterparts and are viewed as socially deficient and insufficiently feminine (Rudman & Glick, 2001). Therefore, Rudman and Glick (2001) noted that women vying for leadership positions face a double bind. They can either display communal qualities and be seen as nice but incompetent, or they can display agentic qualities and be seen as competent but socially inept. In both cases, women risk being disqualified for leadership roles.

An actual example where such a situation occurred was in the Supreme Court case, Price Waterhouse vs. Hopkins. Hopkins (a woman) was denied partnership in the firm despite better qualifications (e.g. more billable hours, higher customer satisfaction evaluations, more business sold) than her male counterparts who received the promotion. Members of her firm informed her that she was denied a promotion due to interpersonal problems. When pressed for details, her evaluators let her know that she was too "macho," "overcompensated for being a woman," "needed a course at charm school," and might benefit from walking, talking, and dressing more femininely (Fiske, Bersoff, Borgida, Deaux, & Heilman, 1991; p. 339). Hopkins sued her firm, and the final decision by the Supreme Court supported the notion that Hopkins was denied a leadership role in the male-dominated firm expressly based on unfair sex-stereotype expectations. The judge ruled that she was expected to be more congenial than her male counterparts solely on the basis that she was a woman (Fiske et a., 1991).

This is an important example because although by all performance accounts, Hopkins was fit for a leadership position, subjective opinions about her personal style and subsequent development needs played a role in Price Waterhouse's decision to deny her promotion. Given the discrepancies between men and women with regard to their ascension into leadership roles, it is important to understand the differences in the developmental nature of the feedback they are each given. This area has not been studied.

Many practitioners and researchers alike have pointed out the discrepancies between the types of developmental opportunities offered to male and female leadership candidates. Lyness (2002) pointed out that women are less likely to have mentors within

their organizations as well as informal social networks. Although not specifically addressed in this literature, feedback is a fundamental component of most types of development opportunities. Other studies have also pointed out the differences in female and male mentoring relationships, feedback, job assignments, etc. (see Morrison & Von Glinow, 1990).

The importance of feedback is not unique to the leadership literature, of course. However, it is not well studied in this area. Given that male and female leadership candidates are known to be evaluated differently in some cases, and in other cases they are evaluated similarly but not necessary socially accepted in similar ways, it is not unreasonable to suspect that men and women may be receiving different forms of feedback for their development needs as a function of their gender.

In sum, the nature of leadership is changing as organizations move toward models of continuous learning and improvement. The Transformational Leadership model is gaining increasing popularity, and has been hypothesized to lead to positive organizational outcomes. A few studies have shown that some characteristics of Transformational Leadership are communal in nature, and therefore it is suspected that women will be perceived as being more proficient in some transformational skills than in other, more traditional leadership skills. Research and case study findings suggest that, in fact, women can sometimes be seen as effective as men when their leadership behaviors are agentic, however they may also tend to be seen as lacking the communal skills congruent with their gender role. This raises the question: Do women need to act more agentically in order to be perceived as effective leaders? If women are given the appropriate effectiveness evaluations for their performance, is the dislike for agentic

women going to influence the types of developmental feedback they receive, primarily from male followers?

360 Degree Leadership Development Assessments

Most 360 degree feedback tools are subjective in nature, meaning that the rater is not provided with a frame of reference such as behavioral anchors by which to compare the leader's performance (Biernat & Fuegen, 2001; Biernat, Manis, & Nelson, 1991; Biernat & Vescio, 2002). Although procedures vary, typically the leader is rated by people who interact frequently with him or her, who are knowledgeable about the leader's performance, and whose opinions are valued by the leader (Mount, Judge, Scullen, Sytsma, Hezlett, 1998).

Aside from the use of multiple raters, 360 feedback systems differ from traditional evaluation methods in a few ways. First, 360 systems are used most frequently to enhance personal development and growth. Second, ratings provided in 360 systems are made anonymously or confidentially (often with the exception of the immediate supervisor) and are not accompanied by face to face discussion. The confidential nature of 360 systems is believed to promote honesty on the part of the raters and, therefore, be more beneficial to the ratee (Mount et al., 1998).

Multirater systems are believed to have some advantages over traditional evaluation systems (Hazucha, Hezlett, & Schneider, 1993; London & Beatty, 1993; London & Smither, 1995; Tornow, 1993). First, because job performance is multidimensional, raters other than the supervisor may be better suited to evaluate some aspects of performance. Second, even if raters have the same opportunities to observe a leader's performance, they may perceive it differently and therefore evaluate it

differently. Generally speaking, 360 feedback systems are assumed to provide ratees with job relevant information that would otherwise not be available (Mount et al., 1998).

Research in 360 degree assessments has primarily investigated the level of agreement between self and other ratings (e.g. Atwater, Ostroff, Yammarino, & Fleenor, 1998; Johnson & Ferstl, 1999) and, to a lesser degree, method variance (e.g., Mount et al., 1998), measurement equivalence (e.g., Maurer, Raju, & Collins, 1998), and interrater reliability (e.g., Greguras & Robie, 1998). However, few studies have explored differences between various types of raters on different types of skills. In addition, few studies have examined the differences in self-ratings on various types of skills.

There are four possible sources of ratings in most 360-degree assessments: subordinate, peer, boss, and self. Many studies in the 360-assessment field focus on the level of rater-ratee agreement (primarily in studies assessing rater accuracy). However, few studies in the 360-assessment literature have examined the circumstances in which differences between types of raters arise. In addition, there are less than a handful of studies which have examined whether or not raters differ in the specific leadership behaviors they associate most with effectiveness. For instance, a few researchers have found male preferences in ratings provided by bosses (Hooijberg & Choi, 2000; Tsui & O'Reilly, 1989) while one study found a positive female bias from female peer raters (Martell & Desmet, 2001).

Researchers have also asked managers at differing levels to rate the importance of various skills for their own jobs (e.g. Alexander, 1979; Gomez-Mejia, McCann, & Page, 1985; Paolillo, 1981; Pavett & Lau, 1983). At the first line levels, managers favored abilities involved in supervision such as monitoring potential problems and managing

conflict. At the middle management levels, respondents reported that their jobs involved greater human relations skills that involved encouraging cooperation, motivating others, and developing subordinates. At the executive levels, managers reported that their jobs required a broad range of skills including monitoring information, serving as a liaison, and engaging in long-term planning. Finally, in a study of male managers' perceptions of the successful executive skills, respondents reported that successful executives need the ability to act as change agents, have managerial courage, display a results orientation. These are all highly agentic in nature (Martell et al., 1998).

In sum, 360 feedback assessments are usually intended to be developmental in nature and are also believed to have many benefits over more traditional performance assessment systems. Given the nature of multirater assessment tools, the relationship between the leader and rater, as well as the organizational level of the leader are expected to influence the leader ratings. However, few studies have explored differences between various types of raters on different types of skills. In addition, few studies have examined the differences in self ratings on various types of skills.

Summary

Recent research and U.S. census information tells us that women are still having a difficult time advancing to leadership positions in organizations, a phenomenon commonly labeled the "glass ceiling." Many researchers have explored possible causes for the glass ceiling, and some have concluded that gender stereotypes play a large role. More specifically, researchers have proposed that stereotypes affect the quality of the performance evaluations and development opportunities provided to women.

Although researchers have examined gender differences in both the performance evaluation and development fields, most empirical studies have dealt with gender differences in evaluation. However, there are many mixed findings regarding the main effects of leader gender (as well as rater gender) on performance evaluations. Some research has found that women are generally rated lower than men, and other research has found that men are rated lower or that there is no difference at all between genders with regard to performance ratings. This batch of mixed findings has led some researchers to begin exploring moderators of gender differences and performance evaluations. Recently, Eagly and Karau (2002) proposed the Role Congruity Theory to explain the more complex relationship between gender and performance evaluation.

Role Congruity Theory generally states that women will be rated lower when there is a mismatch between their gender role and their job role. In the context of leadership positions, this implies that when a woman's stereotypically communal (e.g. helpful, affectionate, kind, sympathetic, sensitive, gentle, etc.) gender role contrasts with the stereotypically agentic (e.g. ambitious, aggressive, dominant, forceful, independent, self-confident, etc.) leadership role, individuals will see her as a less effective leader. Role Congruity Theory suggests that work context plays a large role in moderating the relationship between gender and performance evaluation to the extent that cues in the environment make either a person's gender or leadership role more salient.

There are many lab studies which generally support the notions proposed in Role Congruity Theory. These studies are typically limited to looking at bottom up ratings of paper leaders or confederate peers. These studies are also limited in the types of measures used. Oftentimes the performance evaluation instruments are developed for the

study and are not based on leadership theories about the types of skills good leaders should possess. Also, many of these studies do not assess leadership performance specifically. Rather, they address general performance on a lab task.

There are few field studies to date exploring the tenets of Role Congruity Theory, and the field studies that do exist are limited in scope. Many are limited by the industry in which the participants work, or only include women in middle management positions. Given the importance of the work context in this field, many researchers have called for the need for more field research in this area.

In addition, there are no studies to date that have taken an integrative approach to studying the glass ceiling with regard to performance evaluation and leadership development. In other words, few have examined those areas in which leadership development is intermixed with assessment methodologies. Many organizations are beginning to merge evaluation assessments with development programs, and more and more companies are using multi-rater development assessments such as 360-degree feedback surveys (Peiperl, 2001). Unfortunately, little is known about the effects of gender stereotypes on such tools. This recent trend toward using more assessment-based leadership development programs raises some interesting questions about how stereotypes may manifest themselves in these types of development paradigms. The following hypotheses section specifies some of these questions.

HYPOTHESES

As outlined in the introduction, there are three main areas of interest in this study. First, the context of the work environment is expected to impact the relationship between leader gender and leader ratings. Second, differences are expected in ratings for communal and agentic leadership skills. Third, the impacts of rater and leader level on leadership ratings are going to be explored.

The current study attempts to address these areas of inquiry through the use of a 360-degree feedback, field data set. The proposed data set contains self and others' (bosses, peers, and subordinates) ratings of male and female leaders working in various contexts. In addition, the measure used to collect the data contains both communal and agentic skill development items, as well as a set of overall leader effectiveness items.

The context variables in this data set include industry, functional area, and proportion of male raters. The degree of male or female domination in each of these contexts will be measured via the percentage of men and women operating in each context. In other words, the percentage of males and females working in each industry and job function will be used in the analyses. For proportion of male raters, the percentage of male raters (combined bosses, peers, and subordinates) for each leader will be calculated.

In addition to various context variables being used in this analysis, there are also various types of ratings. There are communal development ratings, which assess the leaders' development needs with regard to communal leadership skills. There are also agentic development ratings, which assess the leaders' development needs with regard to agentic leadership skills. These two types of development ratings are not expected to

differ in every circumstance. In cases where they are not expected to differ, they will be combined to form a rating for overall development. Finally, there are effectiveness ratings. The content of the effectiveness items focuses on leaders' general leadership ability, leadership potential, and expected success. The effectiveness items are not communally or agentically loaded, and represent the raters' overall perception of each leader's effectiveness. In some of the following hypotheses, there is reason to believe that the independent variables will affect development and effectiveness ratings differently, and in others there is reason to believe that they will be affected in the same manner. The types of ratings being used as the dependent variables in each hypothesis will be specified. Specific research findings and hypotheses are provided below.

Work Context

Eagly and Karau's (2002) Role Congruity Theory suggests that the role of leader and the stereotypical female gender role are incongruent. The theory supports the notion that women are perceived as being relatively ineffective when compared to men to the extent that leadership roles are defined in predominately masculine terms. Although Eagly and Karau (2002) discuss many findings that support their theory, there are still gaps in the research in this area with regard to the specific circumstances or environments in which the incongruence between gender roles and leadership roles is manifest.

Some researchers have noted that a major factor in enhancing the saliency of the gender of female leaders may be organizational context (Becker et al., 2002; Eagly & Karau, 2002; Ridgeway, 1992). Specifically, a woman's gender is more likely to be noticed when she is working in a male-dominated, or non-traditional context. Research in both laboratory and organizational settings support this notion and have indicated that

when phys

prominent

al., 1995, H

Th

make more

less effecti

in male-dor

compared t

contexts

Ther

females, ho

Some studie

individual-l

leadership p

al., 2000), p

the military

Foti, & de V

functional a

Heilman, 20

Few

specifically

gender diffe

study. In ac

when physical characteristics such as gender and color are novel, they become more prominent and can influence observers' perceptions in a stereotypical manner (Eagly et al., 1995; Kanter, 1977).

The current research, therefore, hypothesizes that in contexts which emphasize or make more salient the difference between gender and leader roles, women will be seen as less effective and needing more development than men. In other words, women working in male-dominated contexts may have a harder time being accepted as a leader when compared to men in the same context or when compared to women in female-dominated contexts.

There is some support for the proposition that context affects the evaluation of females; however the way in which researchers define context varies from study to study. Some studies have examined organization-level variables while others have examined individual-level variables. Contexts that have been explored include male domination of leadership positions (Eagly & Karau, 2002); proportion of male subordinates (Bowen et al., 2000); proportion of men among the raters (Bowen et al., 2000); industries including the military, education, government, and social services (Eagly & Karau, 2002; Lord, Foti, & de Vader, 1984); organizational level (Hunt, Boal, & Sorenson, 1990); and functional areas including line jobs, staff jobs, carpentry, and secretarial work (Lyness & Heilman, 2002; Robbins & DeNisi, 1993).

Few of the studies examining context variables have focused on female leaders specifically (most have not had a large enough sample of female leaders to examine gender differences), and many were limited to exploring only one context area within a study. In addition, many of the industries and functional areas included in existing

studies

etc)

explore

chosen

across

rater ge

functio

section

provide

manufac

due to th

were per

but mode

organiza

Becker e

leadershi

in bankin

females i

that wom

these indu

industries,

leads to th

studies do not translate well into the business domain (i.e. nursing, military, education, etc.). With regard to the current study, three types of organizational context will be explored: industry, functional area, and rater gender composition. These contexts were chosen because they represent a variety in levels of analysis. Industry can be assessed across organizations, functional area can be assessed across and within organizations, and rater gender composition is specific to each leader. In addition, the industries and functional areas explored in this study will be business-oriented. The following three sections describe each type of context in detail.

Industry. Industry refers to the type of product or service an organization provides. Examples of industries include healthcare, education, utilities, automotive, and manufacturing. The types of industries explored in past studies have been limited mainly due to the lack of sufficient samples. However, some research has shown that women were perceived as being substantially less effective than men in military organizations, but modestly more effective than men in education, government, and social service organizations (Eagly & Karau, 2002). Looking at more business-oriented industries, Becker et al., 2002) analyzed discrepancies in self-subordinates' perceptions of leadership behavior. They found that there were greater discrepancies for female leaders in banking, accounting, and manufacturing than for males in those settings and for females in education. These studies support the predictions of Role Congruity Theory that women in male-dominated industries (i.e. a much higher percentage of men work in these industries) are less likely to be rated to be as effective as men in male-dominated industries, but when in female-dominated industries, are seen as effective as men. This leads to the first hypothesis in this study:

Hypothesis 1: Industry will moderate the relationship between leader gender and others'(combined boss, peer, and subordinate) ratings. In male-dominated industries, female leaders will receive lower overall development ratings (i.e. needing more agentic and communal development) and lower effectiveness ratings (i.e. performing less effectively) than male leaders. In female-dominated industries, no differences are expected between male and female leaders with regard to others' overall development and effectiveness ratings.

Functional area. For purposes of this study, functional area is defined as the leader's primary work area or general job role. Examples of functional areas include engineering, production, human resources, and support staff. Functional areas can be gender typed in the sense that they tend to be viewed as consisting primarily of "man's work" or "woman's work" (Dipboye, 1985). For instance, men are typically associated with functional areas such as engineering and manufacturing whereas women are often associated with functional areas such as human resources and support staff. Many have argued that such gender types associated with certain jobs may interact with the ratee's gender and can lead to gender bias in performance evaluations (Bartol, 1999; Freedman & Phillips, 1988; Knight & Saal, 1984; Nieva & Gutek, 1980). Thus, a critical factor in evaluating context effects on leader evaluations may be whether the gender of the leader is congruent with the gender role that is typically associated with the particular job he or she is performing.

There is some empirical support for the gender/functional area congruence notion. In their meta-analysis of some of the Goldberg-paradigm experiments, Eagly et al. (1992) found strong bias against female leaders in the subgroup of experiments that portrayed men and women as basketball coaches. Robbins and DeNisi (1993) also found that raters provided lower performance ratings in gender-incongruent situations (such as a female in

a carpent

secretaria

N

(2002) fo

to as "lin

evaluatio

external a

jobs are c

male stere

consistent

stereotype

Ly

others hav

Slocum (2

organizatio

managers i

legal exper

line positio

staff positio

Unf

organization

further subst

to the second

a carpenter job) as compared to gender-congruent situations (such as a female in a secretarial job).

More recently, in a study of 485 upper level managers, Lyness and Heilman (2002) found that women received lower performance ratings than men in what they refer to as "line" jobs (business management, operations management, sales) and higher evaluations than men in what they refer to as "staff" jobs (human resources, legal, and external affairs). Lyness and Heilman (2002) concluded that the characteristics of line jobs are consistent with the forceful and achievement-oriented attributes associated with male stereotypes. Alternatively, they concluded that the characteristics of staff jobs are consistent with the caring and relationship-oriented attributes associated with female stereotypes.

Lyness and Heilman's (2002) findings are consistent with the ways in which others have defined the duties for line and staff functions. Hellriegal, Jackson, and Slocum (2002) described managers in line jobs as directing and controlling essential organizational activities such as selling or producing products, whereas they described managers in staff jobs as providing support, such as human resources, public relations, or legal expertise. Ragins and Sundstrom (1989) also supported the idea that managers in line positions tend to have greater organizational power and influence than managers in staff positions.

Unfortunately, the Lyness and Heilman (2002) study was limited to one organization in which ratings were provided by supervisors only. The current study can further substantiate their findings by looking across organizations and raters. This leads to the second hypothesis in this study:

gender
function
needing
perform
differen
develop

R

leadershi

somew ha

to no dire

1974, Le

Stumpf &

finding fe

issue (Bar

in this area

Thompson

(Yammari

few field st

but rather e

employees

In o

ratings of A

two of three

interactions

studies, how

Hypothesis 2: Functional area will moderate the relationship between leader gender and others'(combined boss, peer, and subordinate) ratings. In male-dominated functional areas, female leaders will receive lower overall development ratings (i.e. needing more agentic and communal development) and lower effectiveness ratings (i.e. performing less effectively) than male leaders. In female-dominated functional areas, no differences are expected between male and female leaders with regard to others' overall development and effectiveness ratings.

Rater Gender Composition. The direct effect of rater gender on male and female leadership ratings has been explored in numerous studies. The findings from this area are somewhat mixed. Results from a variety of lab studies suggest that rater gender has little to no direct impact on performance evaluations (Hall & Hall, 1976; Jacobson & Effertz, 1974; Lee & Alvares, 1977; London & Poplawski, 1976; London & Stumpf, 1983; Stumpf & London, 1981). Although such lab studies have been generally consistent in finding few rater gender direct effects, relatively few field studies have investigated this issue (Bartol, 1999). Some have noted that a primary reason for the lack of field studies in this area is that raters are often predominately male (Griffeth & Bedeian, 1989; Thompson & Thompson, 1985; Tsui & Gutek, 1984; Tsui & O'Reilly, 1989) or female (Yammarino & Dubinsky, 1988), precluding such tests in a single study. In addition, the few field studies that have been conducted do not specifically address leadership ratings, but rather examine job specific performance ratings provided to non-management employees.

In one such field study, Pulakos et al. (1989) analyzed the official performance ratings of Army personnel. They found that females gave higher ratings than males on two of three performance dimensions. They also found significant, but minimal, interactions with ratee gender and rating source (peer or supervisor). Two other field studies, however, did not produce findings to support rater gender direct effects. Mobley

(1982) fo

manager

Shore an

for men a

A

colleague

ratings (I

tendency

in cases v

D

gender or

gender co

Bowen et

in field st

raters. Ea

associated

effectiver

female lea

of male su

percentag

compositi

(1982) found no direct effect in a study of performance appraisal ratings for non-managerial and non-professional employees in a large supply distribution center, and Shore and Thornton (1986) found no direct effect in an analysis of performance ratings for men and women doing assembly work.

Alternatively, there have been some studies, primarily from the work of Eagly and colleagues, which have shown that rater gender does have a direct effect on leadership ratings (Eagly & Karau, 2002; Eagly et al., 1992). Their research showed a small tendency for participants to evaluate female leaders less favorably than males, especially in cases where the evaluators were men.

Despite the variation in findings regarding the direct effect of individual rater gender on leadership ratings, one relatively consistent finding has been reported when the gender composition of the group of raters is used as the dependent variable. For example Bowen et al. (2000) found little evidence of overall gender bias in performance appraisals in field studies, however they found significant pro-male biases when only men served as raters. Eagly and Karau (2002) also found that greater proportions of male raters were associated with less effective ratings for women (relative to men) on measures of leader effectiveness (see also Bowen et al., 2000). In addition, other research has shown that female leaders are seen as being less effective relative to male leaders as the proportion of male subordinates increased (Eagly & Karau, 2002). In this study, we will use the percentage of all (bosses, peers, and subordinates) male raters to represent rater composition for each leader. This leads us to the third hypothesis in this study:

Hypothesis 3: The proportion of male raters (i.e. percentage of male boss, peer, and subordinate raters) will moderate the relationship between leader gender and others'(combined boss, peer, and subordinate) ratings. When the proportion of male raters is high, female leaders will receive lower overall development ratings (i.e. needing more agentic and communal development) and lower effectiveness ratings (i.e. performing less effectively) than male leaders. When the proportion of male raters is low, no differences are expected between male and female leaders with regard to others' overall development and effectiveness ratings.

Types of Ratings

In addition to the context effects on leadership ratings, it is important to consider the content of the evaluation instrument for possible gender differences. With regard to 360 assessments specifically, researchers have been primarily concerned with investigating the psychometric viability and practical value of the tool. Researchers have focused little effort on studying the content of the evaluation items. However, given that 360 evaluations are unique from other types of assessments in that they are usually intended for development purposes, it is important to consider the content of these assessments particularly with regard to how they represent what leadership skills are desired. With regard to this study, we will be exploring three evaluation content areas: overall leadership effectiveness, development of agentically-oriented skills, and development of communally-oriented skills.

Development vs. Effectiveness Items. The Price Waterhouse vs. Hopkins case discussed in the introduction highlights an important distinction in the types of information used in promotion decisions. In that case, a female leader was denied promotion despite having performance ratings comparable to her male counterparts. The judge ruled that she was denied promotion because she was expected to be more congenial than her male counterparts solely on the basis that she was a woman (Fiske et al., 1991). Essentially, she did not receive a promotion because her superiors thought she

was lack

hypothe

like for

to or hig

role may

her fulfi

evidence

found th

regarded

less ratio

successf

at a larg

on one c

two othe

overall

highligh

differen

may not

the poss

ratings p

was lacking the communal-oriented skills expected of her as a woman. If we were to hypothesize what performance effectiveness and development ratings might have looked like for Hopkins, we might suggest that her performance effectiveness ratings were equal to or higher than her male peers while her development ratings were left wanting.

The Price Waterhouse vs. Hopkins case shows that a woman who fulfills a leader role may elicit negative reactions, even while she may also receive positive evaluation for her fulfillment of this role (Eagly & Karau, 2002). Heilman et al. (1995) found some evidence of this combination of positive and negative evaluations in a lab study. They found that even when researchers described female managers as successful, participants regarded these women as more hostile (i.e. devious, argumentative, selfish, bitter) and less rational (i.e. illogical, subjective, unable to separate feelings from ideas) than successful male managers.

Shore (1992) found similar results in a field study of female and male employees at a large petroleum company. In this study, women scored significantly higher than men on one of the performance dimensions (performance style) and equivalently to men on two other dimensions (cognitive measures and interpersonal style). However, women's overall ratings were not significantly higher than the men's ratings. Shore's (1992) study highlights the importance of looking beyond an overall rating and rather investigating different types of ratings. As Bartol (1999) put it, "Equal ratings for males and females may not always be equitable" (p. 169). This leads us to Hypothesis 4 which deals with the possible discrepancy between overall effectiveness and individual development ratings provided to men and women.

Hypothesis 4: Leader gender will moderate the relationship between others' (combined boss, peer, and subordinate) overall development (agentic and communal) and effectiveness ratings such that the two sets of ratings will be more strongly, positively related for male leaders than for female leaders. In other words, ratings indicating a lesser need for development will be more strongly related to ratings indicating greater effectiveness for men than for women.

Communal vs. Agentic Development Items. In addition to looking at development items as a whole, they can be separated into those which describe more agentically-oriented skills and those that describe more communally-oriented skills. There are many different types of skills that a good leader needs to have, and various 360 tools differentially emphasize some types of skills more than others.

Although not addressed in any current research, it seems reasonable to hypothesize that there is a difference in the types of development feedback raters provide to male and female leaders. Much of the logic behind such a hypothesis stems from the research in gender differences in leadership style. For instance, Eagly and Karau (2002) reported that women would receive less positive reactions when fulfilling their roles in a particularly dominant, assertive, directive, or self-promoting style.

Across various studies, leadership style has been categorized as either attention to task accomplishment (or initiating structure) versus attention to relationships (or consideration behavior) (Bass, 1990). Initiating structure has been shown to be associated with stereotypic masculine attributes, whereas consideration has been shown to be associated with stereotypic feminine attributes (Korabik, 1990).

Some researchers have examined the degree to which male and female leader ratings differ as a function of the general leadership styles of initiating structure and consideration. In a lab study, Bartol and Butterfield (1976) found that female leaders, relative to male leaders, received more positive ratings if they adopted a consideration

style In a

structure st

unable to r

Izraeli & Iz

Fiel

category of

(involving

(Vroom &

gender diff

initiating str

democratica

given to wo

Vroom (198

both men an

Their study s

leadership th

(Cleveland e

shown that a

than male lea

Baril, & Wats

leaders will b

skills rather th

style. In another lab study, Bass (1981) found that raters considered an initiating structure style to be more suitable for men than for women. Field research has been unable to replicate these lab findings (Butterfield & Powell, 1981; Cleveland et al., 2000; Izraeli & Izraeli, 1985; Rice, Instone, & Adams, 1984).

Field studies have been more successful when investigating another general category of leadership style: the degree to which leaders use a democratic or participative (involving others) versus autocratic (directing others) approach to decision making (Vroom & Yetton, 1973). Powell (1993) reported that meta-analyses revealed larger gender differences on participative and autocratic dimensions than on consideration and initiating structure dimensions, with women tending to be seen as leading more democratically than men. This difference may be partly due to the negative evaluations given to women who lead autocratically (Cleveland et al., 2000). In addition, Jago and Vroom (1982) showed that a participative leadership style was rated as appropriate for both men and women, but an autocratic style was deemed appropriate only for men. Their study suggests that women may have a more restricted choice in the style of leadership they can use in order to be perceived in a positive light by their followers (Cleveland et al., 2000). These findings are consistent with other studies which have shown that autocratic or dominating leadership behavior is less well received from female than male leaders (Copeland, Driskell, & Salas, 1995; Eagly & Karau, 2002; Korabik, Baril, & Watson, 1993; but see Luthar, 1996). This leads us to hypothesize that female leaders will be provided with feedback that suggests they should work on their communal skills rather than build their agentic skills.

Hypothesis 5a: Female leaders will receive lower communal development ratings (i.e. needing more development in communal skill areas) from others (combined bosses, peers, and subordinates) as compared to their male counterparts. The discrepancy between female and male ratings is expected to be more pronounced in male-dominated industries.

Hypothesis 5b: Female leaders will receive lower communal development ratings (i.e. needing more development in communal skill areas) from others (combined bosses, peers, and subordinates) as compared to their male counterparts. The discrepancy between female and male ratings is expected to be more pronounced in male-dominated functional areas.

Hypothesis 5c: Female leaders will receive lower communal development ratings (i.e. needing more development in communal skill areas) from others (combined bosses, peers, and subordinates) as compared to their male counterparts. The discrepancy between female and male ratings is expected to be more pronounced as the proportion of male raters increases.

Organizational Level

Organizational level is the third major category of interest in the current study. Here, organizational level refers to two things: the level of the rater and the level of the leader. The organizational level of the rater is important in a relational sense. For example, we are not interested in whether the boss providing ratings is an executive leader or a middle manager. Rather, we are interested in the fact that this individual is the ratee's superior.

The organizational level of the leader, however, is a different matter. In the case of the ratee, we are interested in where he or she is positioned within the organization. We need to know if the leader being evaluated is a first line supervisor, middle manager, or executive. In a sense, once the organizational level of the leader is known, inferences about the positions of the raters can be made. The following review, therefore, describes the research and hypotheses relevant to two sets of variables: leader/rater relationship and leader level.

Leader/Rater Relationship. There are four possible sources of ratings in most 360-degree assessments: subordinate, peer, boss, and self. Many studies in the 360-assessment field focus on the level of rater-ratee agreement (primarily in studies assessing rater accuracy). However, few studies in the 360-assessment literature have examined the circumstances in which differences between types of raters arise. In addition, there are less than a handful of studies which have examined whether or not raters differ in the leadership roles they associate most with effectiveness. Due to the lack of research in these areas in the 360 literature, the hypotheses proposed in this section of the current study are largely based on the theory of Relational Demography.

The theory of Relational Demography is generally based on the similarity-attraction paradigm and suggests that individuals are more attracted to other individuals who share their same demographic characteristics (Tsui & O'Reilly, 1989). These characteristics may include age, tenure, gender, and race among other things. According to the research in this area, the attraction between individuals sharing the same demographic characteristics can manifest itself in many different ways. With regard to work settings, it has been found that demographic similarity can affect individuals' attitudes toward their work groups, perceptions of advancement opportunities (Riordan & Shore, 1997), and extra-role behaviors (Tsui, Porter, & Egan, 2002). In addition, demographic similarity has also been shown to be related to supervisors' perceptions of individual helping behaviors (Tsui et al., 2002).

In the current study, the theory of Relational Demography would suggest that raters are more likely to rate leaders positively if they share the same demographic characteristics, in this case gender. A few researchers have found differences in ratings

provided to males and females by their bosses. For instance, Tsui and O'Reilly (1989) showed that demographic differences between a subordinate ratee and a superior significantly affected performance evaluations of the ratee. Gender differences were a significant predictor, with specific results indicating that ratees in opposite-gender dyads were rated lower and were liked less well than ratees in same-gender dyads. Specific interactions in this study showed that female superiors gave female ratees higher evaluations than male ratees and indicated greater liking; however they found no differences in ratings from male superiors.

Hooijberg and Choi (2000), alternatively, found that superiors (primarily male) tended to see male managers as being more effective than female managers. Unfortunately, the results of their study are limited due to the small number of women in their study. Ninety percent (90%) of their sample were white males. Although relatively inconclusive, the results of these two studies suggest that bosses might rate male and female leaders differently depending on their own gender. Specifically, it is proposed that male bosses will rate male leaders higher than female leaders, and that female bosses will rate female leaders higher than male leaders. It is also expected that both of these relationships will be stronger in male-dominated contexts. In other words, male bosses are more likely to favor other men when women are the minority, and female bosses are more likely to favor other women when they are the minority.

Hypothesis 6a: Male bosses will provide higher communal and agentic development ratings (i.e. needing less development) and higher effectiveness ratings (i.e. performing more effectively) to male leaders as compared to female leaders. The discrepancy between male and female ratings is expected to be more pronounced in male-dominated industries.

Hypothesis 6b: Male bosses will provide higher communal and agentic development ratings (i.e. needing less development) and higher effectiveness ratings (i.e. performing more effectively) to male leaders as compared to female leaders. The discrepancy between male and female ratings is expected to be more pronounced in male-dominated functional areas.

Hypothesis 6c: Male bosses will provide higher communal and agentic development ratings (i.e. needing less development) and higher effectiveness ratings (i.e. performing more effectively) to male leaders as compared to female leaders. The discrepancy between male and female ratings is expected to be more pronounced as the proportion of male raters increases.

Hypothesis 7a: Female bosses will provide higher communal and agentic development ratings (i.e. needing less development) and higher effectiveness ratings (i.e. performing more effectively) to female leaders as compared to male leaders. The discrepancy between female and male ratings is expected to be more pronounced in male-dominated industries.

Hypothesis 7b: Female bosses will provide higher communal and agentic development ratings (i.e. needing less development) and higher effectiveness ratings (i.e. performing more effectively) to female leaders as compared to male leaders. The discrepancy between female and male ratings is expected to be more pronounced in male-dominated functional areas.

Hypothesis 7c: Female bosses will provide higher communal and agentic development ratings (i.e. needing less development) and higher effectiveness ratings (i.e. performing more effectively) to female leaders as compared to male leaders. The discrepancy between female and male ratings is expected to be more pronounced as the proportion of male raters increases.

Although there are no specific findings with regard to peer ratings in the Relational Demography literature, we can generalize the theory to predict that female peers will rate female leaders higher than male leaders, and that this effect will be stronger in female contexts. Martell and Desmet (2001), for instance, reported finding a female bias in their study such that female managers rated other female managers more favorably than male managers on certain skills (e.g. inspiring, mentoring, problem solving, rewarding, and supporting), whereas male respondents displayed no bias. They claimed this pattern of results indicates a same gender bias on the part of women raters.

Hypothesis 8a: Female peers will provide higher communal and agentic development ratings (i.e. needing less development) and higher effectiveness ratings (i.e. performing more effectively) to female leaders as compared to male leaders. The discrepancy between female and male ratings is expected to be more pronounced in female-dominated industries.

Hypothesis 8b: Female peers will provide higher communal and agentic development ratings (i.e. needing less development) and higher effectiveness ratings (i.e. performing more effectively) to female leaders as compared to male leaders. The discrepancy between female and male ratings is expected to be more pronounced in female-dominated functional areas.

Hypothesis 8c: Female peers will provide higher communal and agentic development ratings (i.e. needing less development) and higher effectiveness ratings (i.e. performing more effectively) to female leaders as compared to male leaders. The discrepancy between female and male ratings is expected to be more pronounced as the proportion of male raters decreases.

There are only a few researchers (e.g. Pfeffer & Salancik, 1975; Tsui, 1984; Tsui & Ohlott, 1988) who have empirically investigated systematic differences in the specific types of leadership behaviors subordinates, peers, and bosses associate with effectiveness. Furthermore, these studies are limited to a very narrow range of industries and have not specifically examined the issue in relation to leader gender. In a field study, Maher (1997) found that there were no differences in the way subordinates evaluated their actual male and female managers. However, more recently, Hooijberg and Choi (2000) performed a field study focusing on 360-feedback for 252 middle managers from a public utility industry. They predicted that subordinates would care more about leadership skills such as involvement, that peers would care more about skills such as providing the organization with a mission, and that superiors would care more about the external side of business such as customer relations. They found mixed results including partial support for their subordinate and supervisor predictions and little support for their peer predictions. Hooijberg and Denison (2002) have recently revised Hooijberg and

Choi's (2000) predictions by suggesting that bosses would give higher ratings to leaders who show agentically-oriented leadership skills (such as providing mission), while direct reports would give higher ratings to leaders with communally-oriented skills (such as involvement).

Although these latter propositions have yet to be tested, they raise an interesting issue regarding the relationship between rater type and the types of skills most related to effectiveness. In addition, one can hypothesize how leader gender impacts the relationship between rater type and the types of skills associated with effectiveness. According to Role Congruity theory, women have a more narrow range of acceptable behaviors than men. For instance, female leaders must be seen as agentic by their bosses, but not too agentic as to appear masculine. Likewise, female leaders must appear communal to their subordinates, but not too communal as to appear soft. Men, alternatively, may have a more broad set of behaviors they can acceptably engage in. A male leader who obviously cares about his subordinates is more likely than a female to be considered as being an effective leader as opposed to being considered soft and incapable. In addition, men are expected by their bosses to appear strong and are less likely than women to be considered too assertive or controlling. Based on this logic, Hypothesis 9 proposes that both rater type (boss or subordinate) and leader gender will impact the relationships between agentic and effectiveness ratings, as well as the relationship between communal and effectiveness ratings.

Hypothesis 9: Rater type (boss or subordinate) and leader gender will interact in their influence on the relationship between agentic and effectiveness ratings, and on the relationship between communal and effectiveness ratings. It is expected that the relationship between agentic and effectiveness ratings will be higher for bosses than for subordinates, more so for men than for women. In addition, the relationship between communal and effectiveness ratings will be higher for subordinates than for bosses, more so for men than for women.

In addition to subordinate, peer, and boss ratings in 360 assessments, leaders are often asked to provide self-ratings of their performance. Self-ratings are being used more and more frequently within performance appraisal systems in organizations (Atwater, 1998), and Bartol (1999) suggests that gender differences in self-ratings may have implications for directly or indirectly affecting the ratings that individuals receive from their supervisors. There is some evidence to support the idea that there are gender-related differences in self-ratings of leadership effectiveness, however findings are conflicting.

Deaux (1979), for example, found that male heads of a retail department store chain viewed themselves as performing significantly better and having greater ability than did their female peers. In addition, the males rated themselves has having higher intelligence, more difficult jobs, more approval for their work, and better relationships with the supervisor than did their female counterparts. However, a more recent study by Tsui and Gutek (1984) indicated a trend toward equivalent or even higher self-ratings by women rather than by men. Results from their study of middle managers in a multi-company corporation showed that female manager self-ratings of performance effectiveness were significantly higher than those of males.

Examples of studies showing no differences between male and female self-ratings include Shore and Thornton (1986), who found no gender differences in ratings on familiar tasks in a real work setting in which performance feedback was available. Also,

in a study focusing on managers, Wohlers, Hall, and London (1993) concluded that male and female managers from a variety of organizations did not differ in self-ratings.

As these studies show, there is little consistent evidence for a direct relationship between leader gender and self-rating bias. Therefore, at this point little is known about the contextual circumstances which affect the likelihood of women producing higher self-ratings than do men (Bartol, 1999). However, just as Role Congruity Theory helped us hypothesize about the factors affecting the ratings provided to female leaders from subordinates, peers, and bosses, it can also help guide predictions about the possible variables affecting female self-ratings.

Role Congruity Theory claims that context plays a large role in predicting when gender impacts leadership evaluation. The three types of contexts proposed for study in the current research are industry type, functional area, and rater gender composition. If we were to generalize our predictions concerning others' ratings within these contexts to self-ratings, then we can propose that self-ratings are going to be impacted by the context in which the leader is performing.

Role Congruity Theory also suggests that women will respond to their own performance similarly to how others respond such that female leaders will rate themselves lower within contexts inconsistent with their gender role. Eagly and colleagues have performed research in this area, and attribute lower self-ratings to stereotype expectancies. They found that expectancies associated with gender roles foster behaviors consistent with these roles (Eagly, 1987; Eagly & Karau, 2002; Eagly, Wood, & Diekman, 2000). Others have also reported that gender stereotype expectancies are especially powerful in yielding demonstrations of behavioral confirmation (Geis,

1993; Skrypnek & Snyder, 1982; Snyder, 1981), including studies of stereotype threat (Steele, 1997). For instance, Spencer, Steele, and Quinn (1999) showed that it is possible to undercut women's math performance through raising anxiety about confirming people's low expectations about women's inferior math ability. Based on this research, it is reasonable to assume that women are more likely confronted with stereotype expectancies in male-dominated work contexts as opposed to environments where they are surrounded by other females.

This is also consistent with Relational Demography theory which states that individuals are likely to compare their own demographic characteristics with those around them and then determine whether or not their own demographics are similar. Relational Demography theory suggests that the level of an individual's demographic similarity to the demographic composition of those around them drives the individual's work-related attitudes and behaviors (Riordan & Shore, 1997). Although Relational Demography research has not specifically addressed the issue of self-ratings, the theory suggests that individuals will be more confident about their own performance when performing in situations where they share demographic characteristics with those around them.

Hypothesis 10a: Industry will moderate the relationship between leader gender and self-ratings. In male-dominated industries, female leaders will give themselves lower agentic and communal development ratings (i.e. needing more agentic and communal development) than will male leaders. In female-dominated industries, no differences are expected between male and female leaders with regard to self-ratings of agentic and communal development needs.

Hypothesis 10b: Functional area will moderate the relationship between leader gender and self-ratings. In male-dominated functional areas, female leaders will give themselves lower agentic and communal development ratings (i.e. needing more agentic and communal development) than will male leaders. In female-dominated functional areas, no differences are expected between male and female leaders with regard to self-ratings of agentic and communal development needs.

Hypothesis 10c: The proportion of male raters will moderate the relationship between leader gender and self-ratings. When the proportion of male raters is high, female leaders will give themselves lower agentic and communal development ratings (i.e. needing more agentic and communal development) than will male leaders. When the proportion of male raters is low, no differences are expected between male and female leaders with regard to self-ratings of agentic and communal development needs.

In addition to context variables, the current study has proposed that leader development ratings provided to males and females will vary partly as a function of the types of items (i.e. communally or agentically-oriented) contained in the evaluation instrument. Again, in generalizing these hypotheses to self-ratings, we propose that females will provide higher self-ratings on communal-type skills than will men.

There is some research to support the claim that women see themselves as having a higher level of communal leadership skills than do men. For example, in a study of middle managers' self-reports of how they influence their bosses, females reported that they acted more out of organizational interest than self-interest, considered others' viewpoints, and focused on the interpersonal aspects as well as the task aspects of the influence episode (Lauterbach & Weiner, 1996). In another study, both male and female leaders reported behaving more agentically toward their subordinates relative to their bosses. However, women reported using a more communal style overall, regardless of their own organizational level in relation to those they interacted with (Moskowitz, Suh, & Desaulniers, 1994).

Hypothesis 11: Female leader self-ratings will be higher for communal skills (i.e. needing less development in communal skill areas) than male leader self-ratings of communal skills.

We also propose that in situations where females rate themselves highly in agentic skills, others will see these leaders as being less effective. This proposition is consistent with Role Congruity Theory in that women who see themselves as displaying agentic skills are likely to be perceived by others as using a leadership style inconsistent with the stereotypical female gender role.

Hypothesis 12: Leader gender will moderate the relationship between others' (combined boss, peer, and subordinate) effectiveness ratings and agentic development self-ratings such that the two sets of ratings will be more strongly, positively related for male than for female leaders. In other words, self ratings indicating a lesser need for development of agentic skills will be more strongly related to others' ratings indicating greater effectiveness for men than for women.

Finally, there is another hypothesis that can be made about the relationship between self and others' ratings. London and Wohlers (1991) found that the self-perceptions of female managers were more in agreement with their subordinates' perceptions than were those of male managers. This is consistent with the notion in hypotheses 9 (i.e. effectiveness ratings for female leaders will be positively related to subordinate communal development scores) and 11 (i.e. female leader self-ratings will be higher for communal skills than male leader self-ratings of communal skills) such that female leaders and subordinates will have more commonality regarding the skills they rate highly than will male leaders and subordinates.

Hypothesis 13: Self-subordinate rater agreement on communal and agentic development ratings will be higher for female leaders than for male leaders.

Leader Level. For purposes of this study, leader level refers to the ratee's position in the organization. Although researchers are often interested in those holding the highest

positions in organizations, they often define leadership more broadly. As mentioned in the introduction, most research studies examine middle managers because they are greater in number and more accessible. Although many studies only examine those in the middle ranks of organizations, many have speculated about the generalizability of findings to high-level leaders. Eagly and Karau (2002), for instance, proposed that the incongruity between the female gender role and the leader role is likely to be more extreme at higher levels of leadership. The findings from a study by Lyness and Judiesch (1999) are consistent with Eagly and Karau's proposition. Their study included managerial employees in a large financial services organization in which women at higher levels of management were less likely to be promoted when compared to men.

Adding to the complexity of the leader level issue, Eagly and Karau (2002) describe middle managers as having a particularly multifaceted role relative to first-line managers or top leaders. They, and others, claim that the role of middle managers is socially complex and requires stereotypically feminine interpersonal skills (Paolillo, 1981). Given the more socially complex characteristics of the middle management role, Eagly and Karau (2002) suggest that the gender-leader role incongruity might be lower for middle managers than for first level or executive level leaders. Therefore, they conclude that bias against female leaders is especially likely above and below the middle management level and that women should fare well, relative to men, in middle-level leadership positions as opposed to first line or executive positions. In this study, we are mainly concerned with the differences between leaders at the middle management and executive levels. For that reason, the following hypothesis addresses differences between these two groups.

Hypothesis 14: Leader level will moderate the relationship between leader gender and others'(combined boss, peer, and subordinate) ratings. At the executive level, female leaders will receive lower agentic and communal development ratings (i.e. needing more development in agentic and communal leadership skills) and lower effectiveness ratings (i.e. performing less effectively) than male leaders. At the middle-management level, no differences are expected between male and female leaders with regard to others' ratings.

METHODS

Design

The hypotheses presented in the last section were studied using data from an archival set of responses to a 360 leadership evaluation: the Denison Leadership Development Survey (DLDS; Denison & Neale, 1996). The responses on this survey were provided in a single time period. The decision was made to use field data as opposed to lab data in this study for many reasons.

First, multiple researchers have highlighted the need for further field research in the area of gender and leadership evaluation (Bartol, 1999; Eagly & Karau, 2002). Given that the use of 360 assessments is becoming more and more prevalent in practice and research, this seems an appropriate forum for investigating the relationship between gender and leader evaluation.

Second, the ratings provided in this data set were of real leaders by actual peers, subordinates, and bosses. Although data gathered in a lab would allow for more control, it would be difficult to simulate a 360 assessment in the lab while maintaining external validity. This data set gave us an opportunity to use data from a real evaluation/development instrument currently being used in many organizations rather than creating a contrived one for a lab study. In addition, field 360 ratings allowed us to assess self, top-down, and bottom-up ratings as opposed to just bottom-up ratings that could be provided in a lab study.

Third, in addition to variation in types of raters, this data set contained variation in the contexts within which the data were collected, as well as variation in the types of ratings provided to individuals. The contexts in this data set included the industries and

functional areas in which the leaders worked, as well as the proportion of male raters for each leader. The types of ratings included development and leadership effectiveness. Further, the development items in the DLDS spanned both traditional and transformational leadership skills.

Finally, this particular data set allowed us to assess ratings provided to female executive leaders as opposed to strictly middle management, as most studies have done. Lab studies are usually restricted to examining bottom-up, "subordinate" ratings of paper leaders or team peers. In addition, as Bartol (1999) noted, the number of studies conducted in field settings has been small and have been largely concentrated on positions near the bottom of the hierarchy where women have experienced more success in being promoted.

Participants

The data were collected during the years 1996-2002 from 206 different American companies[1]. The sample included organizational members from 51 various industries including automotive, manufacturing, telecommunications, pharmaceutical, and others. Subjects in this data set were all members of companies that paid Denison Consulting Company for their services. Although it is impossible to say if all individual participants were required to fill out the survey or if they volunteered, it is the case that all surveys were completed for developmental purposes only. Survey results were provided to the leaders only. Therefore, survey results were not provided to the leaders' bosses, peers, or subordinates and they were not used for performance evaluation purposes.

The data set includes ratings of 3,938 leaders who were rated by 29,566 subordinates, peers, and bosses. It should be noted, however, that not all of these leaders

and raters were included in every analysis. For example, some leaders did not provide self-ratings and some leaders and raters did not report their gender, therefore, these leaders and raters were not included in analyses where these factors needed to be known. Of those leaders who reported their gender, 988 were women and 2702 were men. This means that there were approximately three times more male than female leaders in this data set.

Table 1 shows the breakdown of the numbers of raters for male and female leaders. As the table shows, female subordinate and peer raters made up 52% of female leaders' raters whereas female subordinate and peer raters accounted for only 24% of male leaders' raters. Alternatively, male subordinates and peers made up 63% of male leaders' raters whereas male subordinates and peers accounted for only 34% of female leaders raters. In addition, female leaders were four times more likely to have a female boss rater than were male leaders. It is important to restate here that leaders chose their own raters. These numbers suggest that women leaders were more likely to chose women raters than were male leaders.

Despite differences between the types of raters, both male and female leaders had a similar average number of raters. As Table 2 shows, male and female leaders both had an average of approximately 8-9 total raters, 4 subordinate raters, 3-4 peers, and 1-2 boss raters.

In addition to gender, leaders and raters filled out a variety of demographic information including their age, education, years with organization, annual salary, and ethnic background (see Appendix A). The demographic characteristics of the participants who provided demographic information are presented in Table 3. The results are

[1] Eight companies were excluded from the study because they were not American organizations.

categorized into the following groups: all leaders, male leaders, female leaders, all bosses, male bosses, female bosses, all peers, male peers, female peers, subordinates, male subordinates, and female subordinates.

There were little to no differences between groups or genders in the area of ethnicity. Most of the participants were white/Caucasian (84-90%). Males reported being slightly older than females within each group (leaders, bosses, peers, and subordinates). This may be due in part to the relatively large percentage of male compared to female leaders, bosses, and peers in the 50-59 age range. Across groups, most leaders, peers, and subordinates were between the ages of 30 and 49 whereas most bosses were between the ages of 40 and 59.

Overall, a higher percentage of males reported having higher education levels than females within each group (leaders, bosses, peers, and subordinates). Across groups, most leaders, bosses, peers, and male subordinates reported having either a bachelor's or master's degree. Most female subordinates reported having some college or a bachelor's degree. Similarly, a higher percentage of males reported a higher number of years spent in their organizations than females within each group. Across groups, subordinates (male mean=5.48, 4-10 years; female mean=5.15, 4-10 years) reported slightly less years in the organization than peers (male mean=6.03, 6-10 years; female mean=5.68, 4-10 years) and leaders (male mean=5.97, 4-10 years; female mean=5.56, 4-10 years) who reported slightly less years than bosses (male mean=6.40, 6-15 years; female mean=5.94, 4-10 years).

The demographic with the largest differences between groups was salary. Salary was measured on an eight-point scale ranging from 1=$25,000 or less to, 8=$200,001

plus (see Appendix A). Within each group (leaders, bosses, peers, subordinates), males reported higher salaries than females. The salary discrepancies between men and women were highest for bosses (male mean=6.42, or approximately $100,000 to $200,000; female mean=5.38, or approximately $75,000 to $150,000) and subordinates (male mean=4.58, or approximately $50,000 to $100,000; female mean=3.42, or approximately $35,000 to $75,000). Across groups, female subordinates report being paid the least (mean=3.42, or approximately $35,000-$75,000) and male bosses reported being paid the most (mean=6.42, or approximately $100,000 to $200,000). Female leaders and peers reported being paid about the same as male subordinates (female leader mean=4.78, female peer mean=4.55, male subordinate mean=4.58; or approximately $50,000 to $100,000), and male leaders and peers reported being paid about the same as female bosses (male leader mean=5.36, male peer mean=5.41, female boss mean=5.38; or approximately $75,000 to $150,000).

Looking strictly at the leaders who were rated in this data set, male leaders had significantly ($p<0.05$) higher salaries than female leaders even after controlling for age, education, tenure, ethnicity, functional area, and organizational level (male leader: N=1277, mean=5.48, SD=1.40; female leader: N=536, mean=4.86, SD=1.40; R^2=0.42). Female leaders reported an average salary of $50,000 to $100,000 whereas male leaders reported an average salary of $75,000 to $150,000.

Measures

Denison Leadership Development Survey. The instrument used to gather the data used in this study is a 360 degree leadership assessment called the Denison Leadership Development Survey (DLDS; Denison & Neale, 1996). It allowed each leader to make a

self assessment of both their developmental areas and overall effectiveness as a leader. In addition, each leader chose up to 25 bosses, peers, and direct reports to rate them in the same areas. Responses were confidential in the sense that no names were attached to the feedback reports that were provided to the leaders. In fact, raters had to agree that their ratings could be shown in the case that they were the only rater within a boss, subordinate, or peer category. Surveys were completed on-line and took approximately 20-30 minutes to complete. For more details about the administration and reporting of the DLDS, see the excerpt taken from the DLDS facilitator guide (Denison & Neale, 2001) in Appendix B.

The DLDS was developed based on Quinn's (1988) Competing Values Framework (CVF). Others have used this framework effectively to explore leadership and organizational culture issues (Buenger, Daft, Conlon, & Austin, 1996; Denison, Hooijberg, & Quinn, 1995; Hart & Quinn, 1993; Hooijberg, 1996; Hooijberg & Choi, 2000). Quinn's (1988) CVF provided a useful model of leadership for the current study because it addresses both the traditional and transformational roles of a leader. The tenets of CVF are that organizations are complex and dynamic, therefore managers and leaders have to fulfill many competing expectations; they have to express both flexibility and control, and maintain both an internal and external focus.

Essentially, then, both the CVF and the DLDS are based on a quadrant type model where the skills reflecting the leaders' flexibility and control are crossed with the skills reflecting their focus on internal and external issues (see Figure 1). The first quadrant includes skills that represent both an external focus and flexible leadership style. In the DLDS, this quadrant is called the *Adaptability* dimension. The second quadrant includes

skills that represent an external focus and stable leadership style. This is called the *Mission* dimension in the DLDS. The third quadrant includes skills that represent an internal focus and flexible leadership style. This is called the *Involvement* dimension in the DLDS. Finally, the fourth quadrant includes skills that represent both an internal focus and stable leadership style. This dimension is called *Consistency* in the DLDS. In addition, within each of these four main dimensions (Adaptability, Mission, Involvement, and Consistency), the DLDS specifies three subdimensions as shown in Table 4.

As Table 4 shows, the Adaptability dimension contains skills dealing with organizational learning and change, and customer focus. The Mission dimension includes skills dealing with providing followers with the strategies, goals, and vision for the organization. The Involvement dimension contains skills dealing with the people side of leadership: empowering others, building teamwork, and developing others. Finally, the Consistency dimension includes skills dealing with coordinating and integrating organizational systems, resolving conflicts, and defining and enforcing organizational core values. The excerpt from the DLDS facilitator guide (Denison & Neale, 2001) in Appendix C provides a more detailed description of each dimension and sub-dimension. Also, the specific items within each sub-dimension of the DLDS are listed in Table 5[2]. The items in this survey are all stated positively such that high ratings indicate proficiency in a skill (i.e. does not need further development). The rating responses are in a Likert-type format and range from 1=strongly disagree to 7=strongly agree.

For purposes of this study, we were primarily concerned with which set of items dealt most with agentic-type and communal-type leadership skills. Conceptually, there

[2] For a report on the psychometric properties of the DLDS, see Cho (2000) and Hooijberg and Denison (2002).

are some items that appeared to represent either agentic or communal skills. As described in the introduction, agentic leadership styles are usually associated with characteristics such as decisiveness, assertiveness, achievement orientation, confidence, ambitiousness, and an autocratic style. This implies that leaders with an agentic style of leadership are concerned with providing their followers with direction, establishing high standards of performance, and independently assessing the needs of the organization. Alternatively, a communal style of leadership is generally associated with characteristics such as nurturing others, caring for the welfare of others, being sensitive to individuals' needs, and taking a service orientation.

In order to determine which items in the DLDS had the best conceptual fit with the notions of agentic and communal leadership skills, subject matter experts were given definitions of agentic and communal skills and asked to rate each item of the DLDS on a scale of 1 (agentic) to 5 (communal). The subject matter experts were Industrial/Organizational Psychology graduate students who were all trained in item analysis. Table 6 shows the item means, standard deviations, percent of subject matter experts who rated the items above 3 and below 3, and if the item was chosen for inclusion in the study as a communally or agentically-oriented item.

In all, 22 subject matter experts rated the items. One person's ratings were dropped due to missing ratings and outliers leaving 21 usable ratings. The items were then chosen based on two criteria. First, the item had to have a majority of people agreeing that it fit on one side of the scale or the other. In other words, 50% or more of subject matter experts had to rate the item as a 1 or 2 in order for it to be considered agentic, and 50% or more had to rate it as a 4 or 5 for the item to be considered

communal. Second, the mean subject matter expert rating had to be equal to or above 3.50 or equal to or below 2.50 in order for the item to be considered communal or agentic, respectively. Using these criteria, 22 agentic and 32 communal items were chosen for use in this study (see Appendix D for a list of the communal and agentic items). Reliabilities of the communal and agentic scales were then computed across raters in the actual data set. Both scales produced high reliability coefficients (communal α=0.85, agentic α=0.90).

The agentic and communal scales were significantly, highly correlated (r=0.83, corrected r=0.95). The nature of the DLDS instrument was such that if a leader was rated highly (needing less development) on one set of items, he or she was also likely to be rated highly on most other items. Although every effort was made to create conceptually distinct scales, the strong statistical relationship between the two sets of development items made it more difficult to find results for those hypotheses suggesting different effects for agentic and communal skills types. Normally, the strong relationship between these scales would have helped to make a case for combining them, thereby dismissing some of the hypotheses presented in this study. However, for purposes of this study, the development scales were treated as conceptually distinct where hypothesized to be so[3].

Leadership Effectiveness Items. There were seven leadership effectiveness items. These items were asked at the end of the DLDS and addressed the overall effectiveness of the ratee as a leader. These items had the same Likert-type response scale as the development items (1=strongly disagree to 7=strongly agree; i.e. higher ratings indicated higher leadership effectiveness). The effectiveness items were not tied to any

[3] Analyses for the set of combined agentic and communal items were also run for each hypothesis, but did not produce results different from those reported in the Results section.

performance-related incentives such as pay raises or promotions. Therefore, the effectiveness items used in this survey were not used in a traditional performance appraisal sense. The leadership effectiveness items can be seen in Appendix E. The reliability of the effectiveness items was assessed across groups (α=0.95). The effectiveness items correlated highly with the agentic and communal scales (r=0.81 and r=0.82 respectively). These correlations did not differ between the sets of ratings provided to male and female leaders. The correlation between the development and effectiveness items put limitations on the likelihood of finding evidence for hypotheses suggesting a difference between these types of items. However, as with the development scales, the set of effectiveness items was treated as a conceptually distinct scale based on item content.

Organizational Level. The organizational level (or leader position) categories on the Denison Leadership Develop Survey can be seen in Appendix F. Each leader was responsible for indicating his or her own level. The seven levels used in the survey were compiled into three levels: line, middle, and executive. Those who reported 2 (line-management, N=695) were grouped into a line management category given that they supervise non-management personnel. Those who reported 3 (middle management, N=1286) were grouped into the middle management category given that they supervise line managers. Finally, those who reported 4 (senior management, N=753), 5 (executive/senior vice president, N=322), and 6 (CEO/President, N=128) were grouped into the senior/executive management category given that they supervise middle managers. Those ratees who reported an organizational level of 1 (non-management, N=175) or 7 (owner, N=37) were not coded given that we can not be sure these

individuals are in leadership roles in their organizations. Table 7 shows the number of male and female leaders in each major category. As can be seen in the table, 27% of female leaders were at the line management level and 32% were at the executive level. Alternatively, 20% of male leaders were at the line management level and 40% were at the executive level.

Functional area. Each leader was responsible for indicating his or her own functional area (see Appendix G). Each of the functional areas used in the DLDS were coded according to how male (or female) dominated it is. The functional areas used in the DLDS were matched up with corresponding occupational areas in the Standard Occupational Classification (SOC) system used by the U.S. Bureau of Labor Statistics (2001b; http://www.bls.gov/cps/cpsaat11.pdf). The SOC system is used by all Federal statistical agencies to classify workers into occupational categories for the purpose of collecting, calculating, or disseminating data. In this system, all workers are classified into one of over 820 occupations according to their occupational definition. To facilitate classification, occupations are combined to form 23 major groups, 96 minor groups, and 449 broad occupations.

Although the occupations used by the U.S. Bureau of Labor Statistics are not perfectly representative of the functional areas used in the DLDS, this was the closest classification system available which also provided statistics on the gender make up of the workforce within each occupation. Therefore, each DLDS functional area was located in the SOC system and the corresponding percentage of females working within that occupation was coded in the current data set. This was performed by two subject matter experts at Denison Consulting. These subject matter experts were chosen based

on their familiarity with the DLDS functional area options. In cases where the experts disagreed (for 3 of the 10 functional areas, or 30% of cases), they discussed the categorization until they reached consensus. Consensus was reached for all functional areas. Each DLDS functional area and its corresponding DOC classification are also provided in Appendix G.

Table 8 shows the number of leaders in each functional area. It is important to note here that individuals could only be included when they provided the functional area in which they worked. As the table shows, most leaders worked within the sales and marketing (% female leaders=20, % male leaders=22), administration (% female leaders=20, % male leaders=17), and professional (% female leaders=21, % male leaders=15) functional areas. This is true for both male and female leaders. A higher percentage of female leaders worked in the human resources (% female leaders=11, % male leaders=3) and professional (% female leaders=21, % male leaders=15) functional areas whereas a higher percentage of male leaders worked in the manufacturing (% male leaders=10, % female leaders=3) and engineering (% male leaders=8, % female leaders=3) functional areas.

Table 9 shows the number of leaders that worked within male and female dominated functional areas. Not surprisingly, more female leaders in this data set worked in functional areas typically consisting of a high percentage of women. Similarly, more male leaders from this data set worked in functional areas typically consisting of a high percentage of men.

Industry Type. The industry type of each company was not provided in this data set, but was coded by Denison Consulting subject matter experts. The industry categories

used were taken from the Standard Industrial Classification (SIC) system that is used throughout the federal government to group establishments into industries (U.S. Bureau of Labor Statistics, 2001c). Two subject matter experts from Denison Consulting were interviewed and asked what type of business each company in the data set was engaged in. The subject matter experts were chosen based on their knowledge of the companies in the data set. These individuals were asked to place each company from the data set into the appropriate SIC category. In cases where the experts disagreed (for 21 of the 206 companies included in the study, or 11% of cases), they discussed the categorization until they reached consensus. Consensus was reached for all companies. The 2001 U.S. Bureau of Labor Statistics information regarding the percentage of females working in each industry can be found at http://www.bls.gov/cps/cpsaat18.pdf.

There were 51 total industries represented in this data set. Table 10 shows the number of leaders in each major industry area. It is important to note here that one subset of companies (20 total) in this data was not coded for industry and therefore not included in any analyses which examined industry. Those companies were consulting firms which used the DLDS with their own clients. Although the personal information provided by the individuals within these consulting companies was still useful, we had no way of tracking what industries these individuals worked in. In other words, we did not have the company names from which these individuals actually worked, rather we had the company name of the consulting firm that delivered them the survey. Therefore, rather than grouping all of these individuals into a consulting service industry which would have misrepresented their true industry affiliations, these individuals were not included in the industry analyses (878 leaders total: 275 female leaders and 603 male leaders). Of the

leaders who were included, there were male leaders in 48 of the 51 industries and female leaders in 41 of the 51 industries.

As can be seen in Table 10, 30% of male leaders worked in the manufacturing industry while 29% of female leaders work in the services industry. A higher percentage of female leaders worked in the transportation, communications, and other public utilities (% female leaders=9, % male leaders=5); finance, insurance, and real estate (% female leaders=17, % male leaders=10); and services (% female leaders=29, % male leaders=17) industries whereas a higher percentage of male leaders worked in the manufacturing (% male leaders=30, % female leaders=17), wholesale and retail trade (% male leaders=15, % female leaders=11), and construction (% male leaders=11, % female leaders=6) industries.

Table 11 shows the number of leaders that worked within male and female dominated industries. As expected, more female leaders in this data set worked in industries typically consisting of a high percentage of women. Similarly, more male leaders from this data set worked in industries typically consisting of a high percentage of men. The industry percent of females and functional area percent of females were correlated; however, the correlation coefficients were low: r=0.17 for male leaders and r=0.01 for female leaders.

Proportion of Male Raters. This variable represents the percentage of boss, peer, and subordinate raters (not including self) who were male. This variable was computed simply by taking the number of male raters and dividing by the total number of raters for each leader. Table 12 shows that about half of female leaders (48%) had less than 40%

male raters (mean % male raters=41%). Alternatively, most of the male leaders (64%) had over 60% male raters (mean % male raters=66%).

The proportion of male raters and the percentage of women in industries were negatively correlated (r=-0.31 for male leaders, r=-0.39 for female leaders) indicating that as the percent of women working within an industry increased, the proportions of male raters decreased. The proportion of male raters was also correlated, albeit to a small degree, with the functional area percent of females (r=-0.14 for male leaders, r=-0.13 for female leaders.

Analyses

With the exception of Hypotheses 10c, 11, and 13 which were tested using Analysis of Variance techniques, most of the hypotheses in this study required Hierarchical Linear Modeling (HLM) for testing.

HLM. This data set was composed of numerous levels of nested data which made it suitable for Hierarchical Linear Modeling (HLM) analysis (Raudenbush & Bryk, 2002; Snijders & Bosker, 1999). HLM is a statistical method that accounts for the dependencies that occur in the hierarchical nature of nested data (Hofmann, 1997). In the current study, individual level ratings were nested within each leader, and each leader was nested within various industries and functional areas. While HLM is similar to ordinary least squares (OLS) analysis, it has a few advantages in the analysis of cross-level data. First, because HLM techniques model both the individual and group level variances, it can assess within and between group differences independently from one another (Hofmann, 1997). Second, HLM can generally handle missing data and unbalanced designs (Hofmann, 1997).

Multi-rater feedback studies typically aggregate individual level ratings to the leader level. Those that involve higher level context variables typically disaggregate those variables to the leader level. Most studies then use some variant of analysis of variance (multiple ANOVAs or MANOVAs) to study group differences in ratings (see Lyness & Heilman, 2002; Shore & Thornton, 1986). Many have commented on the problems with aggregating individual level data to a higher group variable (see Klein & Kozlowski, 2000). In fact, in a study exploring gender and race differences in interviewer ratings, Sacco, Scheu, Ryan, and Schmitt (in press) showed that using analysis of variance techniques can increase the Type I error rate (finding an effect when it does not really exist) in nested data structures. They found more significant findings when using MANOVA procedures as compared to accounting for the clustering in the data through using HLM analyses. Therefore, they suggest using HLM for clustered data sets such as the one being used in this study to reduce the risk of Type I error.

HLM also has a few limitations that will need to be addressed in this study. First, HLM has limited use as a multivariate procedure and at this point, is almost always used in a univariate paradigm. Therefore, only one dependent variable can be examined at a time. The hypotheses in this study were designed with this in mind. Also, HLM is limited in the number of levels and interactions it can reasonably assess. Technically, this data set was composed of four levels: Individual ratings (level-1) were clustered within each leader (level-2), each leader was clustered within companies (level-3), and companies were clustered within industries (level-4). Unfortunately, anything over a three-level HLM model is extremely rare and beyond the capabilities of most analysis programs. In addition, this study was conceptually concerned with the contextual aspects

of each company that could explain differences between ratings provided to male and female leaders. This data set did not contain any information about each company other than the industry it belonged to. Therefore, even if differences were found at the company level, little theoretical rationale could be provided for the result given the data available. This study, therefore, disaggregated the level-4 variables (industry and functional area) to level-3 (company).

The hypotheses offered in this study which required HLM analyses were tested via four general models (see Figures 2-5). First, Hypotheses 1, 2, 5a-b, 6a-b, 7a-b, and 8a-b all proposed a moderation effect of a level-3 context variable (industry or functional area) on the relationship between a level-2 variable (leader gender) and a level-1 set of ratings (see Figure 2). Second, Hypotheses 3, 5c, 6c, 7c, 8c, 12, and 14 all proposed a moderation effect of a level-2 variable (usually proportion of male raters) on the relationship between a level-2 variable (usually leader gender) and a level-1 set of ratings (see Figure 3). Third, Hypotheses 4 and 10a-b proposed a moderation effect of a level-2 variable (leader gender for example) on the relationship between two level-1 variables (for example others' development ratings and others' effectiveness ratings; see Figure 4). Finally, Hypothesis 9 proposed a 2-level, 3-way interaction. Rater type (level-1) was expected to moderate the relationship between agentic ratings (level-1) and effectiveness ratings (level-1) and also the relationship between communal ratings (level-1) and effectiveness ratings (level-1). Leader gender (level-2) was then expected to moderate the effect of rater type. The formulas for each of these four models are also presented in Figures 2-5.

Hypothesis testing in HLM involves the same procedure regardless of the number of levels or interactions in each model. It involves testing successively more complex models, with the statistical significance of parameters in initial models as a requirement for finding significance in later models (Sacco, Scheu, Ryan, & Schmitt, in press). Therefore, the first step in HLM hypothesis testing is evaluating the null model (see Figure 6). The null model contains no predictors and simply shows whether or not there is a significant amount of variance in the outcome variables that can be explained by the grouping variables. The variance components of the null model error terms can be used to calculate the intra-class correlation coefficients or ICCs. The ICC represents the proportion of between group differences due to each grouping variable.

Based on whether or not the between group variances are significant, the next step in HLM hypothesis testing involves adding more predictors to the model, one by one. Each model is them compared to the subsequent model to determine if the predictor helped to explain more variance in the outcome variable. Although R^2 is typically used to represent the variance explained in these models, some have argued that R^2 is not an appropriate measure for use in multi-level models (Raudenbush & Bryk, 2002). However, to date, no other measure has been offered and therefore R^2 is still used.

HLM tests cross-level relationships via a regression model, and provides information about both intercepts and slopes. Although the moderation effects in which we were most interested in the current study were represented in the slope coefficients (intercept and slope variance coefficients are labeled in Figures 2-5 for each of the HLM models being tested in this study), an issue regarding the interpretation of the intercept coefficients should be pointed out. The interpretation of intercept coefficients can be

affected by centering procedures. Although significance tests in HLM are not affected by the method of centering used, centering is a procedure that impacts the interpretability of the intercept results by altering the scaling of the variables.

There are two forms of centering (in addition to leaving the data in its raw form): group mean and grand mean. Hofmann and Gavin (1998) note that grand mean centering produces equivalent models to not centering at all (raw metric), but affects the interpretation of the intercept coefficients. Group mean centering, alternatively, is not equivalent to either grand mean centered or raw metric models. Although Hofmann and Gavin (1998) point out that there is no "right or wrong" method of centering, some methods are more suited to answering different types of questions.

Kreft, DeLeeuw, and Aiken (1995) and Raudenbush and Bryk (2002) recommend using grand mean centering for two reasons. First, they note that grand mean centering usually results in the reduction of covariance between the intercepts and slopes, which can reduce potential problems with multicollinearity (see Hofman & Gavin, 1998). In addition, Raudenbush and Bryk (2002) note that grand mean centering is appropriate for dichotomous variables (i.e. gender and leader level as used in this study) in cases where there are a discrepant number of individuals in each group (which was the case in this study given that there were more male than female leaders, and more middle managers than executives). Therefore, all variables in this study were grand mean centered. Although the intercept values represented different things in the different HLM models being used in this study, when grand mean centered, it is important to remember that the intercept coefficients for each grouping variable represented an adjusted group mean as opposed to a raw group mean.

RESULTS

Overview

The specific analyses for each hypothesis proposed in this study are summarized in Table 13. First, some basic descriptive information is presented. Following that, some tests of main effects are described. Although not hypothesized, these main effects tests were performed to assess some of the trends in the descriptives data most relevant to the hypotheses. Then, the results of the 3-level and 2-level moderation hypotheses tests are reported. The relevant hypotheses within each set of results are mentioned where appropriate. The final outcome of each hypothesis test is also summarized in Table 13.

General Descriptives

The means and standard deviations for the agentic and communal scales are provided in Table 14. The data in this table are broken down by rater group, rater gender, and leader gender. For female leaders, the agentic scale means ranged from 5.42 (SD=0.82) to 5.67 (SD=0.86) and the communal scale means ranged from 5.53 (SD=0.79) to 5.83 (SD=0.52). Female leader self ratings were higher than the ratings from the other rater groups for the communal scale (self mean=5.83, SD=0.52), followed by the male subordinate mean rating (mean=5.73, SD=0.81). For the agentic scale, male subordinates provided female leaders with the highest mean rating (mean=5.67, SD=0.86).

For male leaders, the agentic scale means ranged from 5.43 (SD=0.84) to 5.66 (SD=0.64) and the communal scale means ranged from 5.45 (SD=0.80) to 5.78 (SD=0.54). Male leader rated themselves higher than did any other rater group on both scales (agentic mean=5.66, SD=0.64; communal mean=5.78, SD=0.54) followed by

female subordinates (agentic mean=5.57, SD=0.95; communal mean=5.56, SD=0.95). For both male and female leaders, male peers provided the lowest ratings than any other rater group for both the communal (female leader mean=5.53, SD=0.79; male leader mean=5.45, SD=0.80) and agentic scales (female leader mean=5.42, SD=0.82; male leader mean=5.43, SD=0.84).

Across leader groups, female leaders rated themselves slightly higher than male leaders on the communal scale (female mean=5.83, SD=0.52; male mean=5.78, SD=0.54) whereas male leaders rated themselves slightly higher on the agentic scale (male mean=5.66, SD=0.64; female mean=5.63, SD=0.68). All other rater groups rated female leaders higher than male leaders on both scales with the exception of male peers who rated male and female leaders similarly on the agentic scale (female leader mean=5.42, SD=0.82; male leader mean=5.43, SD=0.84).

The most apparent differences between male and female leader ratings came from male subordinates and female peers. These two groups rated female leaders more highly than male leaders on both the agentic and communal scales. For the agentic scale, the mean rating from male subordinates was 5.53 (SD=0.91) for male leaders and 5.67 (SD=0.86) for female leaders. Again for the agentic scale, the mean rating from female peers was 5.49 (SD=0.88) for male leaders and 5.63 (SD=0.80) for female leaders. For the communal scale, the mean rating from male subordinates was 5.53 (SD=0.86) for male leaders and 5.73 (SD=0.81) for female leaders. Again for the communal scale, the mean rating from female peers was 5.50 (SD=0.88) for male leaders and 5.69 (SD=0.78) for female leaders. The lowest apparent differences between male and female leader ratings came from male bosses. Male bosses rated male and female leaders similarly on

both the agentic (male leader mean=5.56, SD=0.77; female leader mean=5.57, SD=0.77) and communal (male leader mean=5.60, SD=0.71; female leader mean=5.63, SD=0.71) scales.

The standard deviations for the agentic ratings ranged from 0.68 to 0.92 for female leaders and from 0.64 to 0.95 for male leaders. The standard deviations for the communal ratings ranged from 0.52 to 0.89 for female leaders and from 0.54 to 0.95 for male leaders. Across male and female leaders, the standard deviations are fairly consistent for each rater group with the female subordinates showing the most amount of variance (female leader agentic SD=0.92, female leader communal SD=0.89, male leader agentic SD=0.95, male leader communal SD=0.95) and leader self ratings showing the least variance (female leader agentic SD=0.68, female leader communal SD=0.52, male leader agentic SD=0.64, male leader communal SD=0.54).

Overall, both male and female leaders rated themselves higher on the communal scale than the agentic scale. In addition, the other rater groups tended to rate women higher on the communal scale as opposed to the agentic scale. Although this was also true in a few cases for male leaders, the agentic and communal scale means were more similar for male leaders than for female leaders across most of the rater groups.

Table 14 also shows the means and standard deviations of the effectiveness ratings provided to male and female leaders by each rater group. The results shown on the table are consistent with the pattern of results for the communal and agentic scales. For female leaders, subordinates provided the highest mean ratings (female subordinate mean=5.90, SD=1.14; male subordinate mean=5.90, SD=1.06) and male peers provided the lowest mean rating (mean=5.41, SD=1.09). For male leaders, female subordinates

provided the highest mean rating (mean=5.78, SD=1.22) and male peers provided the lowest mean rating (mean=5.37, SD=1.13).

Across leader groups, every rater group produced a higher mean rating for female leaders than male leaders except for the male bosses. The male bosses rated male and female leaders quite similarly, though the male leader mean rating was slightly higher (female leader mean=5.62, SD=0.97; male leader mean=5.64, SD=0.98). The biggest differences between male and female leader ratings came from male subordinates and female peers. These two groups rated female leaders (male subordinate mean=5.90, SD=1.06; female peer mean=5.63, SD=1.09) more highly than male leaders (male subordinate mean=5.72, SD=1.13; female peer mean=5.47, SD=1.17).

The standard deviations for the effectiveness ratings were higher than those for the communal and agentic ratings, ranging from 0.97 to 1.14 for female leaders and 0.98 to 1.22 for male leaders. However, across male and female leaders, the standard deviations were fairly consistent for each rater group with the female subordinates showing the most amount of variance (female leader SD=1.14, male leader SD=1.22) and the male bosses showing the least (female leader SD=0.97, male leader SD=0.98).

Main Effects

Prior to testing the hypotheses, a few tests of main effects were conducted to determine the significance of some of the relationships noted in the descriptives section that were relevant to the relationships proposed in the study. One-way multivariate analysis of variance (MANOVA) and 2-level, main effect HLM analyses were conducted to examine the main effects of leader gender, leader level, and rater sex on agentic, communal, and effectiveness ratings.

Leader Gender Main Effects. Leader gender was significantly related to others' (subordinates, peers, and bosses) communal, agentic, and effectiveness ratings such that female leaders were rated more highly (i.e. more effective and needing less development) than male leaders on every dimension (see Table 15). In addition, female leaders' self ratings were significantly higher than male leaders' self ratings of communal skills (i.e. needing less communal skill development). There were no significant differences in self ratings of agentic skills (see Table 16).

As Tables 15 and 16 show, the R^2 values for the main effects of leader gender range from 0.001 to 0.004. These values are quite low, suggesting that although significant, leader gender described a very small amount of the variance in each set of ratings.

Leader Level Main Effects. Organizational level was significantly related to others' (subordinates, peers, and bosses) ratings for the agentic scale only (see Table 15). Others' rated executives' as needing less agentic development than middle managers'. In addition, although not statistically significant, others' rated executives as needing more communal development than middle managers. As for self ratings, executives rated themselves as needing significantly less communal and agentic development than did middle managers (see Table 16).

As with the leader gender main effects, the variance in ratings explained by leader level was very low (R^2 values ranged from 0.0001 to 0.016). This indicates that leader gender described a very small amount of variance in each set of ratings.

Rater Gender Main Effects. Rater gender was significantly related to others' (subordinates, peers, and bosses) communal, agentic, and effectiveness ratings. Female

raters rated leaders higher than male raters did on each scale. This indicates that female raters considered leaders as being more effective and needing less development than did male raters. The variance in ratings explained by rater gender was low (R^2 = 0.004 to 0.005).

HLM Null Models

For all hypotheses requiring the use of HLM procedures, the null models were assessed in order to determine whether or not there was a significant amount of variance in the dependent variables that could be explained by the grouping variables. In the 2-level models, each ICC represented the proportion of variance in the outcome that was between groups (in most cases, leader gender). In the level-3 models, the ICC for the level-2 variance component represented the proportion of variance among the level-2 variables within the level-3 variables (for example, the proportion of variance among leader genders within industries), while the ICC for the level-3 variance component represented the proportion of variance among the level-3 variables (industry or functional area). Table 17 shows the null models for each outcome variable for each hypothesis. As can be seen in Table 17, of the 70 between group variance components calculated, all were significant except for four level-2 components and one level-3 component. For Hypothesis 7a, the between gender variance within industry was not significant for female bosses' communal ratings. For Hypothesis 7b, the between gender variance within functional area was not significant for either female bosses' communal rating or for female bosses agentic ratings. In addition, in Hypothesis 7b the variance among functional areas was not significant. Finally, for Hypothesis 7c, the variance among leader genders was not significant for female bosses' communal ratings.

Normally, if the between group variance coefficients are not significant, it is a sign that the grouping variables are unlikely to explain variance in the outcome variables and therefore the model does not require further testing. However, in this study, all steps for all hypotheses were completed up to the final models for sake of completion. The final models (in their complete form) for all hypotheses requiring HLM analyses are presented in Tables 18-21.

HLM Model 1: 3-Level, 2-Way Interactions

This first set of hypotheses dealt with the moderation effect of context variables (industry and functional area) on the relationship between leader gender and others' ratings (see Figure 2). In these hypotheses, others' individual ratings were level-1 variables. These ratings were all clustered within the second-level variable: leader. The leaders, in turn, were clustered within the third-level context variables. Therefore, 3-level HLM analysis was used for each of these hypotheses. Table 18 shows the results for all of the 3-level, 2-way interaction hypotheses.

Hypotheses 1, 5a, 6a, 7a, and 8a. Hypotheses 1, 5a, 6a, 7a, and 8a all proposed that industry would moderate the relationship between leader gender and others' ratings. There were four significant interactions resulting from this set of analyses, but three of the four relationships were not in the direction expected. First, for Hypothesis 1, industry moderated the relationship between leader gender and others' development ratings (R^2=0.0081) such that others' rated female leaders more positively than male leaders (i.e. needing less development), especially in industries composed of higher percentages of women. Although both men and women were rated higher in female dominated industries, women were rated significantly higher than men. Second, the same industry

moderation effect was found for others' effectiveness ratings in Hypothesis 1: others' rated female leaders as more effective than male leaders, especially in industries composed of higher percentages of women (R^2=0.0069). Third, when looking at communal development items specifically (Hypothesis 5a), the same result was found (R^2=0.0067). Others' rated female leaders as needing less communal development than male leaders, especially in female dominated industries. All three of these results were opposite of what was expected. Males were hypothesized to receive higher ratings, especially in male-dominated settings.

The fourth and final significant relationship found for this set of hypotheses came from the tests for Hypothesis 8a. Here again, it was found that women were rated higher, especially in female-dominated industries. However, in the case of Hypothesis 8a, this was hypothesized to be true. Female peers rated female leaders more highly than male leaders on agentic skills (i.e. needing less agentic development), especially in female-dominated industries (R^2=0.0139). No effect was found for communal development ratings in Hypothesis 8a.

Hypotheses 1, 5b, 6b, 7b, and 8b. Hypotheses 1, 5b, 6b, 7b, and 8b all proposed that functional area would moderate the relationship between leader gender and others' ratings. There were no significant moderation effects found for functional area.

HLM Model 2: 2-Level, 2-Way Interactions

This set of hypotheses all proposed that a level-2 variable (usually proportion of male raters) would moderate the relationship between a level-2 variable (usually leader gender) and a level-1 outcome variable (usually others' ratings). The model for this set

of hypotheses can be seen in Figure 3. The results for this set of hypotheses are reported in Table 19.

Hypotheses 3, 5c, 6c, 7c, and 8c. Hypotheses 3, 5c, 6c, 7c, and 8c all predicted that the proportion of male raters would moderate the relationship between leader gender and others' ratings. The percentage of male raters for each leader was used to represent the proportion of male raters in this study. There was one significant finding in this set of hypotheses. As predicted in Hypothesis 8c, female peer effectiveness ratings were higher for women than for men, and were especially high when the proportion of male raters was low (R^2=0.0059).

Hypothesis 12. Hypothesis 12 predicted that the relationship between self agentic ratings and others' effectiveness ratings would be moderated by leader gender. In this model, both leader gender and self-ratings were level-2 variables. Support was not found for this hypothesis.

Hypothesis 14. Hypothesis 14 predicted that the relationship between leader gender and others' ratings would be moderated by leaders' organizational level. Both leader gender and leader level in this case were level-2 variables. Support was not found for this model. Although not included in the hypotheses, an analysis was also run to test for a leader gender by leader level interaction on self-ratings. As in the case with others' ratings, support was not found for this interaction with regard to development or effectiveness ratings.

HLM Model 3: 2-Level, 2-Way Interactions

These hypotheses all proposed that a level-2 variable (for example, leader gender) would moderate the relationship between two level-1 variables (such as two sets of ratings). The general model for this set of hypotheses can be seen in Figure 4. The results for this set of hypotheses are reported in Table 20.

Hypothesis 4. This hypothesis predicted that leader gender would moderate the relationship between two sets of ratings (others' effectiveness and development ratings). Here, each set of others' ratings were treated as level-1 variables and leader gender was treated as a level-2 moderator. The moderation effect of leader gender was not significant.

Hypotheses 10a-b. Hypothesis 10 predicted that context (industry and functional area) would have a direct effect on female leaders' self-ratings. Given that there was only one self-rating for each leader, self-ratings were modeled at the leader level. In previous hypotheses, the leader level was described as a level-2 variable because the dependent variables in previous hypotheses were at the individual rater level (level-1). In Hypotheses 10a-b, however, the leader level was the lowest level of analysis and therefore became level-1. Similarly, the context variables that were modeled as level-3 variables in previous hypotheses were modeled as level-2 variables in the analyses for Hypotheses 10a-b. In short, because leader self-ratings and not individual-level others' ratings were modeled in Hypothesis 10, self-ratings were considered a level-1 variable and context variables were considered to be level-2.

There was a significant moderation effect of industry on the relationship between leader gender and self agentic ratings. As hypothesized, women rated themselves lower

than men (i.e. needing more development) on agentic skills in male dominated industries. In addition, although no differences were expected between male and female self-ratings in female-dominated industries, the data showed that females rated themselves higher (i.e. needing less agentic development) than men in female dominated industries (R^2=0.0292).

HLM Model 4: 2-Level, 3-Way Interaction

There was only one hypothesis which used this model: Hypothesis 9. Hypothesis 9 was a 2-level, 3-way interaction and can be seen in Figure 5.

Hypothesis 9. Hypothesis 9 stated that rater type (boss or subordinate) and leader gender would interact in their influence on the relationship between agentic and effectiveness ratings, and on the relationship between communal and effectiveness ratings. It was expected that the relationship between agentic and effectiveness ratings would be higher for bosses than for subordinates, more so for men than for women. In addition, the relationship between communal and effectiveness ratings was expected to be higher for subordinates than for bosses, more so for men than for women. This hypothesis was not supported (see Table 21).

Analysis of Variance Tests

Three hypotheses (10c, 11, and 13) were tested using Analysis of Variance analyses. The results for these hypotheses are reported in Table 22.

Hypothesis 10c. Hypothesis 10 predicted that the proportion of male raters would moderate the relationship between leader gender and self ratings. Given that there was only one self-rating for each leader, self-ratings were modeled at the leader level. In previous hypotheses, the leader level was described as a level-2 variable because the

dependent variables in previous hypotheses were at the individual rater level (level-1). In Hypothesis 10c, however, the leader level was the lowest level of analysis and therefore became level-1. Similarly, the proportion of male raters was a leader level variable. Given that all of the variables in this hypothesis were at the same level, HLM was not required. Rather, a 2-Way MANOVA was performed. No significant interactions were found for this hypothesis.

Hypothesis 11. Hypothesis 11 predicted that there would be a direct effect of leader gender on leader communal development self-ratings. In this instance, both leader gender and leader self-ratings were modeled at the leader level. Therefore, a one-way ANOVA was sufficient to test this hypothesis. The main effect of leader gender on communal self-ratings was significant in the direction hypothesized. Women had higher self ratings of communal skills (needing less communal development) than did men ($R2=0.002$). However, analyses of main effects showed that women also rated themselves higher on agentic skills (needing less agentic development).

Hypothesis 13. Hypothesis 13 predicted that there would be a difference between male and female leaders with respect to self-subordinate rater agreement on communal and agentic development scores. Contrary to the other hypotheses, Hypothesis 13 required that individual level data be aggregated. Specifically, a self-subordinate rater agreement index was created for each leader. Although HLM could have been used here to examine the correlational relationship between self and subordinate ratings, it could not be used to examine the level of agreement between these sets of ratings. Therefore, in order to create a measure of self-subordinate agreement, subordinate scores were averaged for each leader. Then, a difference score was created by subtracting the

subordinate average score from the leader self-rating. The absolute value of the difference score was used as the dependent variable for this hypothesis. Therefore, this hypothesis test assessed the direct effect of leader gender on self-subordinate agreement. Both of these variables were modeled at the same level given that they both existed at the leader level. Since there were no cross-level relationships hypothesized here, a one-way MANOVA was sufficient to test Hypothesis 13. There was a significant main effect of leader gender on self-subordinate rater agreement on agentic ratings. The level of agreement for agentic ratings was higher (i.e. difference score was lower) between male leaders and subordinates than between female leaders and subordinates (R^2= 0.002). This was contrary to the hypothesis that the level of self-subordinate agreement would be higher for female leaders than for male leaders.

DISCUSSION

There are few empirical explanations for the disparity in the numbers of men and women in leadership positions. Some have pointed to the pipeline explanation, which states that there are simply not enough women with the appropriate education and background to fill top leadership roles. However, given that women are receiving equivalent degrees at the same rate as men and are equally represented at lower levels in organizations, the pipeline explanation is losing steam (Carli & Eagly, 2001; Wellington, Kropf, & Gerkovich, 2003). In addition, some researchers have noted that even when women enter management positions at the same time as men, they progress more slowly and fail to move as far up the ladder as their male colleagues (Heilman, 1995; Morrison, White, & Van Velsor, 1992).

The continuing lack of female representation in the executive suite has led some researchers to examine the role of gender bias and stereotypes in organizational leadership. Eagly and colleagues (see Eagly & Karau, 2002) have performed a sizeable amount of work in this area, and have proposed Role Congruity Theory to explain how gender stereotypes affect women striving for leadership roles in organizations. Generally, Role Congruity Theory states that stereotypes about women are at odds with stereotypes about leaders. For instance, women are typically associated with feminine or communal characteristics such as sensitivity and nurturing, whereas leaders are typically associated with a more masculine or agentic set of characteristics including assertiveness and ambition.

Role Congruity Theory also asserts that the organizational context within which individuals perform may serve to either weaken or enforce the incongruence of gender

and leader stereotypes. Specifically, Eagly and colleagues (see Eagly & Karau, 2002) have proposed that variables in the work context that make a female's gender more salient will result in more bias concerning that woman's ability to lead. For example, they suggest that when women work in male-dominated settings, their gender becomes a novel and distinguishing characteristic that increases the likelihood that they will experience difficulty in attaining leadership roles.

Given the possibility that stereotypes affect the advancement of women in organizations, it is important to understand how and when these stereotypes are manifested. Some have proposed that stereotypes reveal themselves through gender bias in performance evaluations. Unfortunately, the findings are mixed with regard to the relationship between gender and performance evaluations. One potential reason for this is that studies involving this relationship are often performed in lab settings, few of which include measures of leadership skills. In addition, many of the field studies that do exist were performed in a single organization thereby precluding tests of the possible between-organization context effects suggested in Role Congruity Theory.

Others have suggested examining the differential leadership development opportunities provided to men and women in organizations. There are fewer empirical studies of gender differences in development opportunities than there are of gender differences in performance evaluation, however, evidence suggests that in some cases women receive fewer (or lesser quality) development experiences than their male counterparts.

Although gender differences in performance evaluation and development opportunities have both been cited as barriers for women hoping for leadership roles, past

studies have not examined possible gender differences in ratings provided for developmental purposes. Given the increasing prevalence of multi-rater development tools, it is important to understand how what we know about how possible gender differences in performance evaluations translate into a multi-rater appraisal of leadership development needs. It would also be helpful to know if there are differences in the types of leadership skills that leaders are perceived as needing as a function of their gender. In other words, are women rated as having adequate communal skills and inadequate agentic type skills? Finally, within the framework of Role Congruity Theory, we can ask how context variables (such as industry, functional area, and proportion of male raters) affect the relationship between gender and ratings of leader development needs. The purpose of this study was to investigate the role of gender in the assessment of leadership development needs using Role Congruity Theory as a foundation. The broader goal of this research was to contribute to the understanding of how gender plays a role in the advancement of women in organizations.

Findings

A field data set was used in this study to test possible gender differences in development needs and effectiveness ratings. The data set was large, containing leader ratings from multiple rater sources in over 200 companies, over 50 different industries, and nine functional areas. There were many benefits to using field data for testing the relationships proposed in this study. First, the ratings provided in this data set were of real leaders by actual peers, subordinates, and bosses. This allowed us to assess self, top-down, and bottom-up ratings as opposed to just bottom-up ratings that could be provided in a lab study. Second, this data set contained variation in the contexts within which the

leaders performed as well as variation in the types of ratings (e.g. communal and agentic development, and effectiveness) provided to individuals. Third, the data set allowed us to assess ratings provided to leaders at different organizational levels which was an advantage over past studies which have tended to examine effects at the middle management level only.

However, despite the size and depth of this particular data set, few significant results were found. In addition, it is important to note that even for the few relationships in this study that proved to be statistically significant, the amounts of variance explained by the predictor variables were very low. In fact, there was a small amount of between group (i.e. gender) variance to be explained prior to adding predictors.

It is not uncommon to find statistically significant relationships that account for a small amount of variance in a study with large sample sizes. Eagly and Karau (2002) note that small effect sizes are a relatively common phenomenon in this field of research, and they have claimed that small biases can produce large consequences when repeated over individuals and occasions. In fact, Martell, Lane, and Emrich (1996) performed computer simulations which showed that over time, a small bias against women of 1% of the variance in initial performance ratings produced senior management levels with only 35% women. In addition, they showed that a 5% initial bias produced only 29% senior women. Martell & Robison-Cox (2002) subsequently found that introducing a small bias against women at each yearly review readily creates the disproportionately small numbers of female senior executives that are typical of corporations. Although these simulation studies do not address the effect of first introducing bias at higher levels, (i.e. after women have already progressed through the middle management ranks), they help

to justify how the chances of women rising to high-level positions in organizations are greatly reduced even when slight prejudice is consistently acted on.

Despite the propensity of small effect sizes in this type of research, the effects found in the current study were so small as to be negligible. The largest effects found in this study explain approximately 0.01% of the variance in ratings whereas the amounts found in many studies have ranged between 1% to 5% variance at the lowest (Eagly & Karau, 2002). Therefore, even when comparing the findings of this study relative to others which have found small effects, the effects in this study are slight. For this reason, it is difficult to argue that the few significant effects found in this study provide evidence of gender bias in ratings. However, some of the trends in the data are notable and are described below.

Main Effects. Although there were no hypotheses predicting main effects of gender on ratings, the data showed that women were rated higher than men by others on both development dimensions as well as on overall effectiveness. In addition, women rated themselves significantly higher on communal skills than did men. Therefore, the data suggested a slight propensity for female leaders to be perceived by themselves and by others as needing less development and being more effective than male leaders. This was an unexpected finding which served to negate many of the hypotheses offered in this study given that most of the hypotheses predicted that men would receive higher ratings than would women.

In addition, there was also a significant effect of rater gender on communal, agentic, and effectiveness ratings such that females provided higher ratings than males. This finding did not affect the hypotheses to a great extent given that rater groups were

usually not compared with one another. Because HLM can only handle univariate analyses, the ratings from various groups were either grouped together or ratings from a specific group (for instance, "female peers") were assessed in independent analyses. Once again, it is important to note that the variances explained by leader gender and rater gender were small.

Descriptive Data. Although significance tests were not performed for all of the descriptive data presented, there were some interesting trends in the data. One example is the type of raters who provided the higher ratings to leaders within the sets of male and female leader ratings. Perhaps not surprisingly, both male and female leader self ratings were higher than ratings from any other set of raters. However more surprisingly, male subordinates provided the highest ratings than any other group to female leaders whereas female subordinates provided the highest ratings to male leaders, albeit the effects were trivial. This pattern of comparisons is not consistent with the theory of Relational Demography. Relational Demography Theory would suggest that raters are likely to provide higher ratings to demographically similar as opposed to demographically dissimilar ratees.

Another interesting trend in the descriptive data was that female leaders reported being paid significantly less than male leaders even after controlling for age, tenure, education, organizational level, functional area, and ethnicity. This was the one finding in the study that had a notable effect (R^2=0.42). This finding is consistent with those from other studies showing that even when women are in positions comparable to men, and have equivalent experience and education, they are paid less (Renner, Rives, & Bowlin, 2002).

Context Effects. A large number of hypotheses in this study predicted that context would play a moderating role in the relationship between gender and some given set of ratings such that when the leaders' gender was salient (i.e. leader gender was the minority), rating differences would be greatest. The types of context variables examined were industry type, functional area, and proportion of male raters. The information of interest regarding these context variables was the degree to which they were female or male dominated.

The set of hypothesis which tests for an industry moderation effect produced four significant interactions that all said the same thing: individuals rated female leaders more positively than male leaders, especially in female-dominated industries. This was opposite of what was expected. It was hypothesized that men would receive higher ratings, especially in male-dominated contexts. Although the effect of this relationship was very small, and more information is needed to draw any conclusions about rating bias, the trend in the data is more consistent with a pro-female bias in ratings as opposed to the pro-male bias that was hypothesized.

Alternatively, the slightly higher ratings of female leaders could be due to true ability differences. Women may be more competent given that they became leaders despite working in male dominated contexts, or perhaps because women need to be more competent in general to reach the same leadership positions as less competent men. Unfortunately, there are many alternative explanations that could not be controlled for in this study. It could be that individuals responsible for rating women were overcompensating in their rating so as not to appear biased. It could also be that

individuals rated women based on a different standard than that on which they rated men. In short, too little is known to draw any strong conclusions about this set of findings.

The type of effect found to be significant with industry was not replicated with functional area. In fact, this variable failed to produce any effects on the relationship between leader gender and ratings. Part of the issue here could have been that the functional area options on the DLDS are not completely parallel to the occupations they were paired with from the U.S. Bureau of Labor Statistics occupational classification system. Although subject matter experts reached consensus on where the functional area options fit into the classification system, these classifications were very broad. Therefore, the extent to which the percentages of women working within each functional area are accurately represented is unclear.

In addition, it is possible that industry and functional area interact such that the male domination of a functional area only has an effect to the extent which that function is being performed within a male-dominated industry. In other words, a woman working in human resources in a female-dominated industry may receive different types of ratings than a woman working in human resources in a male-dominated industry. Although not tested in this study, this might help to explain the lack of effect of functional area.

Finally, there was one significant finding with regard to the proportion of male raters context variable. This context variable affected the relationship between leader gender and ratings in the same way industry did, which is not surprising given that the percentage of women working in each industry was moderately correlated with the proportion of male raters for each leader. Findings showed that females were rated higher than males, especially when the proportion of male raters was low (i.e. the

proportion of female raters was high). Again, it is difficult to draw any conclusions about why this occurred, but it is clear that this did not provide support for the gender bias in ratings that was hypothesized.

Kanter (1977) theorized that once a critical mass of the proportion of women in the workplace is reached, they will no longer be seen as the minority and therefore their status would be better appreciated. This suggests that perhaps the context effects in this study would only occur in the most male-dominated of settings. In other words, only in highly male-dominated settings would females receive lower ratings than men. The full range of data were analyzed in this study, however, further analyses of only the most male and female dominated settings may have shown more gender differences in ratings. (See Lortie-Lussier & Rinfret, 2002 for conflicting finding).

Rater Effects. The only significant rater effect found in this study was with regard to self ratings. As hypothesized, female leaders rated themselves lower on agentic skills (i.e. needing more agentic development) than did males in male-dominated industries. However, whereas no gender differences were expected in female-dominated industries, findings showed that females rated themselves higher (i.e. needing less agentic development) than men in female-dominated industries. This might suggest that women are more confident in their level of agentic skills in female-dominated industries, and interestingly, that men are less confident about their agentic skills in female dominated industries. Alternatively, women may not believe that agentic skills are as necessary in female dominated industries, therefore they did not focus on their development needs in that skill area. The finding with regard to men in female-dominated industries is not so straight forward.

One possibility is that the roles of male leaders in female dominated industries differ from those of women. Men may be more likely to be responsible for traditional leadership duties (e.g. providing strategic direction) whereas women may be more responsible for communal duties (e.g. mentoring and developing others). If this is the case, then perhaps focusing on agentic development is more adaptive for men given the types of leadership tasks they perform.

Another possibility is that male leaders' traditional approaches to leadership have failed in female dominated contexts. Yet, rather than considering a shift in leadership style congruent with the context within which they are performing, they instead focus on improving traditionally masculine leadership skills. This implies that men are less likely to recognize the situations in which communal leadership skills are needed, instead believing that agentic leadership will be effective across situations if only developed more fully.

Limitations

Although the data used in this study provided a unique opportunity to examine possible gender differences in development and effectiveness ratings in light of Role Congruity Theory, it had some limitations that precluded drawing conclusions about the role of rating bias in the glass ceiling effect.

Nature of Rating Dimensions. Although every effort was made to create two, conceptually distinct development scales, those used in this study were highly correlated with one another. In addition, the development and effectiveness items were also highly correlated. This essentially created a unidimensional measure of leadership. Unfortunately, this prevented us from making conclusions about how perceptions of

development needs related to overall leadership effectiveness. In this particular data set, good leaders received high ratings across development and effectiveness items which raises questions about whether or not responses from this type of instrument would be useful in future studies attempting to compare development and performance ratings. Raters need to be able to make a distinction in the types of skills being rated. Ideally, these data would have included responses to a second, distinct measure of effectiveness in order to better compare perceptions of performance with perceptions of development needs.

Although the effectiveness items required raters to evaluate the overall effectiveness of each leader, these items were collected in a manner consistent with the development items. In other words, the items assessing overall leadership effectiveness used the same rating scale as the development items, were given on the same survey at the same time as the development items, were reported back to the leaders in the same manner as the development ratings (confidentially), and were not tied to any performance-related incentives such as pay raises or promotions. In addition, the same respondents provided information about both leadership development needs and effectiveness. These factors suggest a common method bias in the data set.

Some meta-analyses indicate that the concern about inflated correlations due to common method bias may not be as large a problem as Campbell (1982) noted. For instance, Crampton and Wagner (1994) found that the "percept-percept inflation may be more the exception than the rule in micro-research on organizations" (p.72). Doty and Glick (1998) also state that the biasing effect of common method variance to distort true correlations among constructs usually does not invalidate our theoretical interpretations

and research conclusions. Yet, Crampton & Wagner (1994) also found that correlations among leadership behaviors are more susceptible to common methods bias than other areas of research. Therefore, using caution in interpreting results is warranted. Given the strong relationship between scales in this study, conclusions can really only be drawn about individuals' perceptions of leadership performance in general. We cannot make conclusions about differences in distinct skill and effectiveness ratings.

Leader Characteristics. Another limitation of this study was the lack of information regarding how leaders came to be in their industries and functional areas. The women working in the male dominated industries or functional areas may have different characteristics from women working in the female dominated industries or functional areas. The possibility exists that women working in male dominated contexts self-selected into those jobs because they are more similar to the men in those contexts. It is also possible that the women who were hired into these male-dominated contexts necessarily had to be high performers to be selected in the first place, which would help explain why these women were rated as highly as men in male-dominated industries. For instance, if the women working in male dominated industries have the agentic skills required within that setting, and are also perceived as having good communal skills as a function of being women, then perhaps they receive higher ratings than men because they not only possess traditional leadership characteristics, but also display interpersonal and supportive leadership.

Standards of Performance. The tool used in this data set was subjective, and therefore, the data did not contain information regarding what or who raters were comparing each leader against. It is also important to emphasize that leaders were

responsible for choosing their own raters. It is also clear from the descriptive data that women had more female raters than did men, suggesting that women were more likely to choose female raters and men were more likely to choose male raters. If leaders chose raters who they thought would rate them more positively, then the variance in ratings was likely reduced as compared to a situation in which raters were chosen at random.

Shifting standards theory (Biernat & Fuegen, 2001; Biernat, Manis, & Nelson, 1991; Biernat & Vescio, 2002) states that in the case of subjective measures as used in this study, raters compare each gender to others within that same gender group. Therefore, a good female leader is not the same as a good male leader, rather, they are both good only relative to others in their own gender groups. Those adhering to this theory would say that gender stereotypes will only show through when you can be sure that individuals are rating leaders of both genders against the same standard of performance. This suggests that the types of rating forms used for research should contain relative comparison information such as a behaviorally anchored response scale.

We also had no information regarding the specific types of leadership tasks that leaders were performing. It is possible that if the women in these organizations are performing less complex, less challenging, or lower-profile leadership tasks than their male counterparts, then despite comparable ratings, they will still be seen as less qualified for higher level roles during the promotion process.

Future Research

Although this study did not provide evidence of gender bias in ratings of development needs or effectiveness, inferences could not be made concerning whether or not bias exists in some other aspect of leader development or promotion processes. In

other words, gender stereotyping may play a role in the glass ceiling phenomenon; however, results of this study suggest that those stereotypes are not manifested in the types of 360 degree leadership development ratings used here. In addition, little support was found for the principles of Role Congruity Theory. The context variables used in this study did not have the anticipated moderation effects on the relationship between leader gender and ratings. However, this is not to imply that Role Congruity Theory would not hold in a study investigating different types of context variables. The idea that organizational context influences the development of female leaders, thereby affecting women's progression up the organizational ladder, is still a worthy subject of investigation. Future research should investigate organizational context from a different perspective. For example, the gender composition of work groups or departments could be examined. Also, it would be helpful to measure the level of perceived gender stereotypes within organizations. Although this might be difficult in a field setting, a lab study could be designed in which the degree of agentic and communal contextual characteristics are manipulated, allowing researchers to have more control over context effects.

Promotion Criteria. Although investigating performance ratings and development needs can tell us part of the glass ceiling story, it is important to better understand the whole story. In other words, what is the full set of both formal and informal criteria used when making promotion decisions? Although ratings are typically used, and are an important part of most promotion systems, ratings alone do not predict who will be promoted. If the various components of promotion processes were better understood, then perhaps there would be more avenues through which to investigate the role of

stereotypes as they relate to leadership advancement. In the current study, the ratings provided were not directly linked to promotion decisions. Rather, they were linked to the feedback individuals received about where they need to improve their skills. Therefore, we are unable to make any direct inferences about how rating bias (or lack thereof) helps us to understand the glass ceiling effect.

Although we have female leaders representing all organizational levels, we had no progression information. In other words, we had no information about how long it took leaders to advance into their current roles or how long they have been in those roles. We also could not assume that the raters who provided information in this study would have provided the same types of information had this tool been used in making promotion decisions.

The case of Price Waterhouse vs. Hopkins is relevant here. Essentially, Hopkins was denied promotion despite having ratings that were equivalent to her male peers who did get promoted. She was partly denied promotion because her bosses met informally and decided that she lacked the appropriate communal skills. Although part of the promotion process, it was likely unknown to employees that meetings such as this played such as large role in promotion decisions. This case provides anecdotal support that a woman who fills a leader role may elicit negative reactions even while receiving positive performance evaluations (Eagly & Karau, 2002), and that these negative reactions can manifest in other ways (possibly through informal mechanisms) that effect the advancement of women.

Development Opportunities. Given that no evidence of gender bias in ratings resulted from this study, more questions can be asked about where and when bias may

play a part in the progression of women into leadership roles. In this study, men and women were given similar development feedback. However, the types of learning and performance opportunities men and women are given in order to improve their skills may still differ.

The results of a recent survey given to 825 Fortune 1000 CEO's and women executives (vice president level and above) showed that both men and women think that the largest barrier to the advancement of women is lack of line experience (defined in this case as those with profit and loss experience) (Wellington, Kropf, & Gerkovich, 2003). The second largest barrier cited by women in the survey is exclusion from informal networks. Interestingly, exclusion from informal networks was not rated highly by men as a barrier for women. This suggests that men have a difficult time recognizing when women are not provided with the same informal opportunities that they are. Men may also be underestimating, or taking for granted the importance of informal opportunities. Third, stereotypes about women's roles and abilities were cited by both men and women to be a likely barrier to the advancement of women. Despite the lack of findings in the current study, the fact that individuals are aware of leadership stereotypes and admit that that they may present a barrier for women lends credence to the idea that these biases do exist and warrant further research.

This anecdotal evidence supports other research suggesting that the developmental opportunities afforded women are different, and often detrimentally so, from those provided to men. For instance Lyness (2002) and others (see Morrison & Von Glinow, 1990; Ragins & Cotton, 1991) have noted that women are less likely than men to

receive mentoring opportunities, challenging assignments, performance feedback, and access to informal social networks within their organizations.

McCauley (2001) notes that the leadership development process has three components: developmental experiences, the ability to learn, and an organizational context that supports development. Opportunities to learn are provided via developmental experiences, and the context of the organization can affect the quality and quantity of those opportunities. McCauley also states that people learn the most under high challenge situations. If the assertions of research are true, and women are not receiving these opportunities or challenges, then perhaps they will be viewed as less qualified than men for leadership roles.

In addition to opportunities to learn, Ford and colleagues (Ford, Quinones, Sego, & Sorra, 1992; Quinones, Ford, Sego, & Smith, 1995) have pointed out the importance of opportunities to perform to enable individuals to transfer the skills they have learned. They specify three dimensions of opportunity to perform: breadth, activity level, and type of tasks performed. They performed a study using an Air Force sample, and found that airmen received differential opportunities to perform. They also found that context variables affected whether or not opportunities were provided. Specifically, a supervisor's perceptions of the airman's capability, skills, and likability affected the likelihood of that airman receiving opportunities to perform trained tasks on the job (Ford, et al, 1992). In the context of gender bias, this may mean that regardless of ability, if the stereotype that females are incapable of performing important leadership roles is believed, then females are unlikely to receive opportunities to controvert the stereotype.

In future studies, it would be helpful to know what types of organizational context variables make development opportunities more or less similar for men and women. For instance, it would be helpful to have information about the organizational culture with regard to leader development and promotion choice points. This may help us to assess how the opportunities offered to men and women may differ in such a way as to create differential qualifications for higher level roles.

Changing Nature of Leadership. Although many have commented that the number of female executives is disproportionately low, the number of female middle managers is on the rise. Some have attributed this to the changing nature of leadership roles. Diekman and Eagly (2000), for instance, introduce the idea of dynamic stereotypes. The Theory of Dynamic Stereotypes suggests that as the social roles of groups change, so do stereotypes about those groups. Given the notion provided by Social Role Theory that the role behavior of group members shapes the group stereotype, groups should have dynamic stereotypes to the extent that their typical social roles are perceived to change over time.

Applied to gender and leadership, this theory makes two predictions about perceived change. First, perceivers should think that gender differences are diminishing because of increasing similarity in the roles of men and women (both in society and at work). Second, the female stereotype should be particularly prone to change given the large scale changes that have occurred in women's social roles over the past 50 years. Based on the theory of dynamic stereotypes, one could make the argument that roles are converging and as higher level leadership roles begin to take on a more communal characterization, women will be more likely to be promoted into these roles.

It would be interesting to study the ways in which leader roles may already be changing toward more androgynous definitions (Schein, 2001). Given the landscape of today's organizations (e.g. rapidly changing, globally competitive, high-technology), it would be worthy of future research to understand how leadership roles are adjusting. Whereas the traditional autocratic and directive style of management may have worked well under more hierarchical organizational structures, it may be the case that a new style (perhaps one more consistent with a participatory, delegatory, and open-communication structure) is better suited to modern organizations. The implications of a changing ideal for leadership may have direct implications on the advancement of women into such roles Eagly & Karau, 2002).

Furthermore, as globalization continues, organizations are likely to be faced with the cultural impact of gender stereotypes with regard to leadership. From a research perspective, it will be important to investigate how gender stereotypes affect the advancement of women in cultures where stereotypes are more or less prevalent than they are in the United States. Additionally, with regard to practical implications, female leaders working overseas could benefit from a better understanding of how cultural contexts interact with stereotypes about their leadership abilities. This may be most important in cases where women in important leadership roles are placed within countries where female leaders and communal approaches to leadership are not well accepted.

Conclusion

In sum, this study did not provide evidence of gender bias in leadership development and effectiveness ratings. In addition, the results of this study showed little support for the moderating effect of context variables on the relationship between leader

gender and ratings as suggested by Role Congruity Theory. There were some limitations to this study that could possibly have impacted our ability to find evidence of rating bias, and have precluded our ability to speak directly to the glass ceiling phenomenon. However, it is important to note that the lack of findings with regard to gender differences in a data set of this size and quality tells us that perhaps bias is not manifesting itself in ratings, but perhaps in some other stage of the development/advancement process. The results of this study might also suggest that gender bias simply does not exist in this context.

It will be important for future studies to begin where this study left off, and examine points at which gender stereotypes may play a role after development needs are assessed. Future research should focus on furthering our understanding of the criteria used for promotion decisions and the potential impact of gender bias on those decision processes. In addition, future research should address the possible gender differences in employee development, specifically in the areas of opportunities to learn and perform. Finally, it is important to expand our investigation of various contextual characteristics that serve to either weaken or reinforce the incongruity between gender and leader stereotypes.

REFERENCES

Alexander, L. D. (1979). The effect level in the hierarchy and functional area have on the extent Mintzberg's roles are required by managerial jobs. Academy of Management Proceedings, 39, 186-189.

Allport, G. W. (1954). The nature of prejudice. Reading, MA: Addison-Wesley.

Antonioni, D. (1996). Designing an effective 360-degree appraisal feedback process. Organizational Dynamics, 25, 24-38.

Atwater, L. (1998). The advantages and pitfalls of self-assessment in organizations. In J. W. Smither (Ed.), Performance appraisal: State of the art in practice (pp. 331-369). San Francisco, CA: Jossey-Bass.

Atwater, L., Ostroff, C., Yammarino, F. J., & Fleenor, J. W. (1998). Self-other agreement: Does it really matter? Personnel Psychology, 51, 577-598.

Atwater, L., Roush, P., & Fischthal, A. (1995). The influence of upward feedback on self and follower ratings of leadership. Personnel Psychology, 48, 35-59.

Atwater, L. E., & Waldman, D. (1998). Accountability in 360-degree feedback: Is it time to take the 360-degree feedback method to its next step? HRMagazine, 43, 96-102.

Bakan, D. (1966). The duality of human existence: An essay on psychology and religion. Chicago: Rand McNally.

Banaji, M. R., & Hardin, C. D. (1996). Automatic stereotyping. Psychological Science, 7, 136-141.

Banaji, M. R., Hardin, C. D., & Rothman, A. J. (1993). Implicit stereotyping in person judgment. Journal of Personality and Social Psychology, 65, 272-281.

Bartol, K. M. (1999). Gender influences on performance evaluations. In G. N. Powell (Ed.), Handbook of gender & work (pp. 165-178). Thousand Oaks, CA: Sage.

Bartol, K. M., & Butterfield, D. A. (1976). Sex effects in evaluating leaders. Journal of Applied Psychology, 61, 446-454.

Bass, B. M. (1981). Individual capability, team response, and productivity. In E. A. Fleishman & M. D. Dunnette (Eds.), Human performance and productivity. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bass, B. M. (1990). Bass & Stogdill's handbook of leadership: Theory, research, and managerial applications (3rd ed.). New York: Free Press.

Becker, J., Ayman, R. A., & Korabik, K. (2002). Discrepencies in self/subordinates' perceptions of leadership behavior: leader's gender, organizational context, and leader's self-monitoring. Group & Organizational Management, 27, 226-244.

Bennett, S. K. (1982). Student perceptions of and expectations for male and female instructors: Evidence relating to the question of gender bias in teaching evaluation. Journal of Educational Psychology, 74, 170-179.

Biddle, B. J. (1979). Role theory: Expectations, identities, and behaviors. New York: Academic Press.

Biernat, M., & Fuegen, K. (2001). Shifting standards and the evaluation of competence: Complexity of gender-based judgment and decision making. Journal of Social Issues, 57, 707-724.

Biernat, M., Manis, M., & Nelson, T. E. (1991). Stereotypes and standards of judgment. Journal of Personality and Social Psychology, 60, 485-499.

Biernat, M., & Vescio, T. K. (2002). She swings, she hits, she's great, she's benched: Implications of gender-based shifting standards for judgment and behavior. Personality and Social Psychology Bulletin, 28, 66-77.

Bigoness, W. J. (1976). Effect of applicant's sex, race, and performance on employers' performance ratings: Some additional findings. Journal of Applied Psychology, 61, 80-84.

Blair, I. V., & Banaji, M. R. (1996). Automatic and controlled processes in stereotype priming. Journal of Personality and Social Psychology, 70, 1142-1163.

Brenner, O. C., Tomkiewicz, J., & Schein, V. E. (1989). The relationship between sex role stereotypes and requisite management characteristics revisited. Academy of Management Journal, 32, 662-669.

Bowen, C., Swim, J. K., & Jacobs, R. R. (2000). Evaluating gender biases on actual job performance of real people: A meta-analysis. Journal of Applied Social Psychology, 30, 2194-2215.

Buenger, V., Daft, R. L., Conlon, E. J., & Austin, J. (1996). Competing values in organizations: Contextual influences and structural consequences. Organization Science, 7, 557-576.

Butterfield, D. A., & Grinnell, J. P. (1999). "Re-viewing" gender, leadership, and managerial behavior: Do three decades of research tell us anything? In G. N. Powell (Ed.) Handbook of gender & work. (pp. 223-238). Thousand Oaks, CA: Sage.

Butterfield, D. A., & Powell, G. N. (1981). Effects of group performance, leader sex, and rater sex on ratings of leader behavior. Organizational Behavior and Human Performance, 28, 129-141.

Campbell, J. P. (1982). Editorial: Some remarks from the outgoing editor. Journal of Applied Psychology, 67, 691-700.

Carbonell, J. L. (1984). Sex roles and leadership revisited. Journal of Applied Psychology, 69, 44-49.

Carli, L. L., & Eagly, A. H. (2001). Gender, hierarchy, and leadership: An introduction. Journal of Social Issues, 57, 629-636.

Catalyst. (2000). Census of women corporate officers and top earners. New York: Catalyst.

Cho, H. J. (2000). Diagnosing leadership development: Validating a model and method. Ann Arbor, MI: Denison Consulting.

Cleveland, J. N., Stockdale, M., & Murphy, K. R. (2000). Women and men in organizations: Sex and gender issues at work. Mahwah, NJ: Lawrence Erlbaum.

Copeland, C. L., Driskell, J. E., & Salas, E. (1995). Gender and reactions to dominance. Journal of Social Behavior and Personality, 10, 53-68.

Crampton, S. M., & Wagner III, J. A. (1994). Percept-percept inflation in microorganizational research: An investigation of prevalence and effect. Journal of Applied Psychology, 79, 67-76.

Davison, H. K., & Burke, M. J. (2000). Sex discrimination in simulated employment contexts: A meta-analytic investigation. Journal of Vocational Behavior, 56, 225-248.

Deaux, K. (1979). Self-evaluations of male and female managers. Sex Roles, 5, 571-580.

Deming, W. E. (1986). Out of the crisis. Cambridge, MA: Massachusetts Institute of Technology, Center for Advanced Engineering Study.

Denison, D. R., Hooijberg, R., & Quinn, R. E. (1995). Paradox and performance: Toward a theory of behavioral complexity in managerial leadership. Organization Science, 6, 524-540.

Denison, D. R. & Neale, W. S. (1996). The Denison leadership development survey. Ann Arbor, MI: Aviat.

Denison, D. R., & Neale, W. S. (2001). Denison leadership development survey: Facilitator guide. Denison Consulting, LLC: Ann Arbor, MI.

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

Diekman, A. B., & Eagly, A. H. (2000). Stereotypes as dynamic constructs: Women and men of the past, present, and future. Personality and Social Psychology Bulletin, 26, 1171-1188.

Dipboye, R. L. (1985). Some neglected variables in research on discrimination in appraisals. Academy of Management Review, 10, 116-127.

Dobbins, G. H., & Platz, S. J. (1986). Sex differences in leadership: How real are they? Academy of Management Review, 11, 118-127.

Dodge, K. A., Gilroy, F. D., & Fenzel, L. M. (1995). Requisite management characteristics revisited: Two decades later. Journal of Social Behavior and Personality, 10, 253-264.

Doty, D. H., & Glick, W. H. (1998). Common methods bias: Does common methods variance really bias results? Organizational Research Methods, 1, 374-406.

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

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

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. Psychological Review, 109, 573-598.

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

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

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.

Eagly, A. H., Wood, W., & Diekman, A. B. (2000). Social role theory of sex differences and similarities: A current appraisal. In T. Eckes & H. M. Trautner (Eds.), The developmental social psychology of gender (pp. 123-174). Mahwah, NJ: Erlbaum.

Esses, V. M., Haddock, G., & Zanna, M. P. (1993). Values, stereotypes, and emotions as determinants of intergroup attitudes. In D. M. Mackie & D. L. Hamilton (Eds.), Affect, cognition, and stereotyping: Interactive processes in group perception (pp. 137-166). San Diego, CA: Academic Press.

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.

Fiske, A. P., Haslam, N., & Fiske, S. T. (1991). Confusing one person with another: What errors reveal about the elementary forms of social relations. Journal of Personality and Social Psychology, 60, 656-674.

Ford, J. K., Quinones, M. A., Sego, D. J., & Sorra, J. S. (1992). Factors affecting the opportunity to perform trained tasks on the job. Personnel Psychology, 45, 511-527.

Frank, F. D., & Drucker, J. (1977). The influence of evaluatee's sex on evaluations of a response on a managerial selection instrument. Sex Roles, 3, 59-64.

Freedman, S. M., & Phillips, J. S. (1988). The changing nature of research on women at work. Journal of Management, 14, 231-251.

Garvin, D. A. (1993). Building a learning organization. Harvard Business Review, 71, 78-91.

Geis, F. L. (1993). Self-fulfilling prophecies: A social psychological view of gender. In A. E. Beall & R. J. Sternberg (Eds.), The psychology of gender (pp. 9-54). New York: Guilford Press.

Giannantonio, C. M., Olian, J. D., & Carroll, S. J. (1995). An experimental study of gender and situational effects in a performance evaluation of a manager. Psychological Reports, 76, 1004-1006.

Glick, P., Zion, C., & Nelson, C. (1988). What mediates sex discrimination in hiring decisions? Journal of Personality and Social Psychology, 55, 178-186.

Goldberg, P. (1968). Are women prejudiced against women? Transaction, 5, 316-322.

Gomez-Mejia, L. R., McCann, J. E., & Page, R. C. (1985). The structure of managerial behaviors and rewards. Industrial Relations, 24, 147-154.

Greguras, G. J., & Robie, C. (1998). A new look at within-source interrater reliability of 360-degree feedback ratings. Journal of Applied Psychology, 83, 960-968.

Griffeth, R. W., & Bedeian, A. G. (1989). Employee performance evaluations: Effects of ratee age, rater age, and ratee gender. Journal of Organizational Behavior, 10, 83-90.

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

Hackman, J. R., & Wageman, R. (1995). Total quality management: Empirical, conceptual, and practical issues. Administrative Science Quarterly, 40, 309-342.

Hall, F. S., & Hall, D. T. (1976). Effects of job incumbents' race and sex on evaluations of managerial performance. Academy of Management Journal, 19, 476-481.

Hamner, W. C., Kim, J. S., Baird, L., & Bigoness, W. J. (1974). Race and sex as determinants of ratings by potential employers in a simulated work-sampling task. Journal of Applied Psychology, 59, 705-711.

Hart, S. L., & Quinn, R. E. (1993). Roles executives play: CEOs, behavioral complexity, and firm performance. Human Relations, 46, 543-574.

Hazucha, J. F., Hezlett, S. A., & Schneider, R. L. (1993). The impact of 360-degree feedback on management skill development. Human Resource Management, 32, 325-352.

Heilman, M. E. (1983). Sex bias in work settings: The lack of fit model. In B. Staw and L. Cummings (Eds), Research in organizational behavior (Vol. 5). Greenwich, CT: JAI.

Heilman, M. E. (1995). Sex stereotypes and their effects in the workplace: What we know and what we don't know. Journal of Social Behavior and Personality, 10, 3-26.

Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. Journal of Social Issues, 57, 657-674.

Heilman, M. E., Block, C., & Martell, R. (1995). Sex stereotypes: Do they influence perceptions of managers? Journal of Social Behavior and Personality, 10, 237-252.

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.

Hellriegel, D., Jackson, S. E., & Slocum, J. W., Jr. (2002). Management: A competency-based approach. Cincinnati, OH: South-Western.

Hofmann, D. A. (1997). An overview of the logic and rationale of hierarchical linear models. Journal of Management, 23, 723-744.

Hofmann, D. A., & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. Journal of Management, 24, 623-641.

Hooijberg, R. (1996). A multidirectional approach toward leadership: An extension of the concept of behavioral complexity. Human Relations, 49, 917-946.

Hooijberg, R., & Choi, J. (2000). Which leadership roles matter to whom? In examination of rater effects on perceptions of effectiveness. Leadership Quarterly, 11, 341-364.

Hooijberg, R., & Denison, D. R. (2002). What makes leaders effective? A stakeholder approach to leadership effectiveness. Unpublished manuscript.

Howell, J. M., & Avolio, B. J. (1993). Transformational leadership, transactional leadership, locus of control, and support for innovation: Key predictors of consolidated, business-unit performance. Journal of Applied Psychology, 78, 891-902.

Howell, J. M., & Higgins, C. A. (1990). Champions of technological innovation. Administrative Science Quarterly, 35, 317-341.

Hunt, J. G., Boal, K. B., & Sorenson, R. L. (1990). Top management leadership: Inside the black box. Leadership Quarterly, 1, 41-65.

Izraeli, D. N., & Izraeli, D. (1985). Sex effects in evaluating leaders: A replication study. Journal of Applied Psychology, 70, 540-546.

Jacobson, M. B., & Effertz, J. (1974). Sex roles and leadership: Perceptions of the leader and the led. Organizational Behavior and Human Performance, 12, 383-396.

Jago, A. G., & Vroom, V. H. (1982). Sex differences in the incidence and evaluation of participative leader behavior. Journal of Applied Psychology, 67, 776-783.

Johnson, J. W., & Ferstl, K. L. (1999). The effects of interrater and self-other agreement on performance improvement following upward feedback. Personnel Psychology, 52, 272-303.

Juran, J. M. (1988). Juran on planning for quality. New York: Free Press.

Kanter, R. M. (1977). Men and women of the corporation. New York: Basic Books.

Kaufmann, G., Isaksen, S. G., & Lauer, K. (1996). Testing the "glass ceiling" effect on gender differences in upper level management: The case of innovator orientation. European Journal of Work and Organizational Psychology, 5, 29-41.

Kirnan, J. P., Farley, J. A., Geisinger, K. F. (1989). The relationship between recruiting source, applicant quality, and hire performance: An analysis by sex, ethnicity, and age. Personnel Psychology, 42, 293-308.

Klein, K. J., & Kozlowski, S. W. J. (Eds.). (2000). Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions. San Francisco, CA: Jossey-Bass.

Knight, P. A., & Saal, F. E. (1984). Effects of gender differences and selection agent expertise on leader influence and performance evaluations. Organizational Behavior and Human Performance, 34, 225-243.

Korabik, K. (1990). Androgyny and leadership style. Journal of Business Ethics, 9, 9-18.

Korabik, K., Baril, G. L., & Watson, C. (1993). Managers' conflict management style and leadership effectiveness: The moderating effects of gender. Sex Roles, 29, 405-420.

Kreft, I. G. G., DeLeeuw, J., & Aiken, L. S. (1995). The effect of different forms of centering in Hierarchical Linear Models. Multivariate Behavioral Research, 30, 1-21.

Lauterbach, K. E., & Weiner, B. J. (1996). Dynamics of upward influence: How male and female managers get their way. Leadership Quarterly, 7, 87-107.

Lee, D. M., & Alvarez, K. M. (1977). Effects of sex on descriptions and evaluations of supervisory behavior in a simulated industrial setting. Journal of Applied Psychology, 62, 405-410.

Lee, J., & Hoon, T. H. (1993). Business students' perceptions of women in management-The case in Singapore. Management Education and Development, 24, 415-429.

Leuptow, L. B., Garovich, L., & Leuptow, M. B. (1995). The persistence of gender stereotypes in the face of changing sex roles: Evidence contrary to the sociocultural model. Ethology and Sociobiology, 16, 509-530.

London, M., & Beatty, R. W. (1993). 360-degree feedback as a competitive advantage. Human Resource Management, 32, 352-373.

London, M., & Poplawski, J. R. (1976). Effects of information on stereotype development in performance appraisal and interview context. Journal of Applied Psychology, 61, 199-205.

London, M., & Smither, J. W. (1995). Can multi-source feedback change perceptions of goal accomplishment, self-evaluations, and performance-related outcomes? Theory-based applications and directions for future research. Personnel Psychology, 48, 803-841.

London, M., & Stumpf, S. A. (1983). Effects of candidate characteristics on management promotion decisions: An experimental study. Personnel Psychology, 36, 241-259.

London, M., & Wohlers, A. J. (1991). Agreement between subordinate and self-ratings in upward feedback. Personnel Psychology, 44, 375-390.

Lord, R. G., Foti, R. J., & de Vader, C. L. (1984). A test of leadership categorization theory: Internal structure, information processing, and leadership perceptions. Organizational Behavior and Human Decision Processes, 34, 343-378.

Lortie-Lussier, M., & Rinfret, N. (2002). The proportion of women managers: Where is the critical mass? Journal of Applied Social Psychology, 32, 1974-1991.

Luthar, H. K. (1996). Gender differences in evaluation of performance and leadership ability: Autocratic vs. democratic managers. Sex Roles, 35, 337-361.

Lyness, K. S. (2002). Finding the key to the executive suite: Challenges for women and people of color. In R. Silzer (Ed.). The 21st century executive: Innovative practices for building leadership at the top. San Francisco: Jossey Bass. (p. 229-273).

Lyness, K. S., & Heilman, M. E. (2002). Gender fit and ratings of upper-level line and staff managers. Paper presented at the 17th Annual Conference of the Society for Industrial and Organizational Psychology, Toronto, Ontario, Canada.

Lyness, K. S., & Judiesch, M. K. (1999). Are women more likely to be hired or promoted into management positions? Journal of Vocational Behavior, 54, 158-173.

Maher, K. J. (1997). Gender-related stereotypes of transformational and transactional leadership. Sex Roles, 37, 209-225.

Martell, R. F., & Desmet, A. L. (2001). A diagnostic-ratio approach to measuring beliefs about the leadership abilities of male and female managers. Journal of Applied Psychology, 86, 1223-1231.

Martell, R. F., Lane, D. M., & Emrich, C. E. (1996). Male-female differences: A computer simulation. American Psychologist, 51, 157-158.

Martell, R. F., Parker, C., Emrich, C. G., & Crawford, M. S. (1998). Sex stereotyping in the executive suite: "Much ado about something." Journal of Social Behavior and Personality, 13, 127-138.

Martell, R. F., & Robison-Cox, J. (April, 2002). Gender bias and organizational mobility: A computer simulation. In R. F. Martell (Chair), Symposium presented at the Society of Industrial-Organizational Psychology Annual Convention, Toronto, Canada.

Massengill, D., & di Marco, N. (1979). Sex-role stereotypes and requisite management characteristics: A current replication. Sex Roles, 5, 561-570.

Maume, D. J., Jr. (1999). Glass ceilings and glass escalators: Occupational segregation and race and sex differences in managerial promotions. Work and Occupations, 1999, 483-509.

Maurer, T. J., Raju, N. S., & Collins, W. C. (1998). Peer and subordinate performance appraisal measurement equivalence. Journal of Applied Psychology, 83, 693-702.

McCauley, C. D. (2001). Leader training and development. In S.J Zaccaro and R. J. Klimosky (Eds.) The nature of organizational leadership: Understanding performance imperatives confronting today's leaders. San Francisco: Jossey-Bass. (p. 347-383).

Megargee, E. I. (1969). Influence of sex roles on the manifestation of leadership. Journal of Applied Psychology, 53, 377-382.

Mobley, W. H. (1982). Supervisor and employee race and sex effects on performance appraisals: A field study of adverse impact and generalizability. Academy of Management Journal, 25, 598-606.

Morrison, A. M., & Von Glinow, M. A. (1990). Women and minorities in management. American Psychologist, 45, 200-208.

Morrison, A. M., White, R. P., Van Velsor, E., & the Center for Creative Leadership. (1992). Breaking the glass ceiling: Can women reach the top of America's largest corporations? (updated ed.). Reading, MA: Addison-Wesley.

Moskowitz, D. S., Suh, E. J., & Desaulniers, J. (1994). Situational influences on gender differences in agency and communion. Journal of Personality and Social Psychology, 66, 753-761.

Mount, M. K., Judge, T. A., Scullen, S. E., Sytsma, M. R., & Hezlett, S. A. (1998). Trait, rater and level effects in 360-degree performance ratings. Personnel Psychology, 51, 557-576.

Newport, F. (2001, February 21). Americans see women as emotional and affectionate, men as more aggressive. Retrieved from the Gallup Poll News Service Web site: http://www.gallup.com/poll/releases/pr010221.asp.

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

Nyquist, L. V., & Spence, J. T. (1986). Effects of dispositional dominance and sex-role expectations on leadership behaviors. Journal of Personality and Social Psychology, 50, 87-93.

Osborn, R. N., & Vicars, W. M. (1976). Sex stereotypes: An artifact in leader behavior and subordinate satisfaction analysis? Academy of Management Journal, 19, 439-449.

Ott, E. M. (1989). Effects of the male-female ratio at work: Policewomen and male nurses. Psychology of Women Quarterly, 13, 41-57.

Paolillo, J. G. (1981). Manager's self assessments of managerial roles: The influence of hierarchical level. Journal of Management, 7, 43-52.

Pavett, C. M., & Lau, A. W. (1983). Managerial work: The influence of hierarchical level and functional specialty. Academy of Management Journal, 26, 170-177.

Pazy, A. (1986). The persistence of pro-male bias despite identical information regarding causes of success. Organizational Behavior and Human Decision Processes, 38, 366-377.

Peiperl, M. A. (2001). Getting 360 feedback right. Harvard Business Review, January 2001, 142-147.

Pfeffer, J., & Salancik, G. R. (1975). Determinants of supervisory behavior: A role set analysis. Human Relations, 28, 139-154.

Powell, G. N. (1993). Women and men in management. Newbury Park, CA: Sage.

Powell, G. N., & Butterfield, D. A. (1979). The "good manager": Masculine or androgynous? Academy of Management Journal, 22, 395-403.

Powell, G. N., & Butterfield, D. A. (1989). The "good manager": Did androgyny fare better in the 1980s? Group and Organization Studies, 14, 216-233.

Powell, G. N., & Butterfield, D. A. (1994). Investigating the "glass ceiling" phenomenon: An empirical study of actual promotions to top management. Academy of Management Journal, 37, 68-86.

Powell, G. N., Butterfield D. A., & Parent, J. D. (2002). Gender and managerial stereotypes: Have the times changed? Journal of Management, 28, 177-193.

Pulakos, E. D., & Wexley, K. N. (1983). The relationship among perceptual similarity, sex and performance ratings in manager-subordinate dyads. Academy of Management Journal, 26, 129-139.

Pulakos, E. D., White, L. A., Oppler, S. H., & Borman, W. C. (1989). Examination of race and sex effects on performance ratings. Journal of Applied Psychology, 74, 770-780.

Quinn, R. E. (1988). Beyond rational management: Mastering the paradoxes and competing demands of high performance. San Francisco, CA: Jossey-Bass.

Quinones, M. A., Ford, J. K., Sego, D. J., & Smith, E. M. (1995). The effects of individual and transfer environment characteristics on the opportunity to perform trained tasks. Training Research Journal, 1, 29-48.

Ragins, B. R. (1991). Gender effects in subordinate evaluations of leaders: Real or artifact? Journal of Organizational Behavior, 12, 259-268.

Ragins, B. R., & Cotton, J. (1991). Easier said than done: Gender differences in perceived barriers to gaining a mentor. Academy of Management Journal, 34, 939-951.

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

Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical linear models: Applications and data analysis methods (2nd Ed.). Thousand Oaks, CA: Sage.

Reilly, R. R., Smither, J. W., & Vasilopoulos, N. L. (1996). A longitudinal study of upward feedback. Personnel Psychology, 49, 599-612.

Renner, C., Rives, J. M., & Bowlin, W. F. (2002). The significance of gender in explaining senior executive pay variations: An exploratory study. Journal of Managerial Issues, 14, 331-345.

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

Ridgeway, C.L. (1992). Gender, interaction, and inequality. New York: Springer-Verlag.

Riordan, C. M., & Shore, L. M. (1997). Demographic diversity and employee attitudes: An empirical examination of relational demography within work units. Journal of Applied Psychology, 82, 342-358.

Robbins, T. L., & DeNisi, A. S. (1993). Moderators of sex bias in the performance appraisal process: A cognitive analysis. Journal of Management, 19, 113-126.

Rosen, B., & Jerdee, T. H. (1974). Effects of applicant's sex and difficulty of job on evaluations of candidates for managerial positions. Journal of Applied Psychology, 59, 511-512.

Rudman, L. A. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. Journal of Personality and Social Psychology, 74, 629-645.

Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. Journal of Social Issues, 57, 743-762.

Sacco, J. M., Scheu, C. R., Ryan, A. M., & Schmitt, N. (in press). An investigation of race and sex similarity effects in interviews: A multi-level approach to relational demography. Journal of Applied Psychology.

Sarbin, T. R., & Allen, V. L. (1968). Role theory. In G. Lindzey & E. Aronson (Eds.), Handbook of social psychology (2nd ed., Vol. 1, pp. 488-567). Reading, MA: Addison-Wesley.

Scandura, T. A. (1991). Breaking the glass ceiling in the 1990's (Technical Report). Washington, DC: U.S. Department of Labor, Women's Bureau.

Schein, V. E. (1973). The relationship between sex role stereotypes and requisite management characteristics. Journal of Applied Psychology, 57, 95-100.

Schein, V. E. (1975). Relationships between sex role stereotypes and requisite management characteristics among female managers. Journal of Applied Psychology, 60, 340-344.

Schein, V. E. (2001). A global look at psychological barriers to women's progress in management. Journal of Social Issues, 57, 675-688.

Schein, V.E., Mueller, R., & Jacobson, C. (1989). The relationship between sex role stereotypes and requisite management characteristics among college students. Sex Roles, 20, 103-110.

Schein, V.E., Mueller, R., Lituchy, T., & Liu, J. (1996). Think manager-think male: A global phenomenon? Journal of Organizational Behavior, 17, 33-41.

Senge, P. M. (1990). The fifth discipline: The art and practice of the learning organization. New York: Doubleday.

Shore, L. M., & Thornton, G. C. (1986). Effects of gender on self and supervisory ratings. Academy of Management Journal, 29, 115-129.

Shore, T. H. (1992). Subtle gender bias in the assessment of managerial potential. Sex Roles, 27, 499-515.

Simmons, W. W. (2001, January 11). When it comes to choosing a boss, Americans still prefer men. Retrieved from the Gallup Poll News Service Web site: http://www.gallup.com/poll/releases/pr010111.asp.

Skrypnek, B. J., & Snyder, M. (1982). On the self-perpetuating nature of stereotypes about women and men. Journal of Experimental Social Psychology, 18, 277-291.

Snijders, T. A. B., & Bosker, R. J. (1999). Multilevel analysis: An introduction to basic and advanced multilevel modeling. Thousand Oaks, CA: Sage.

Snyder, M. (1981). On the self-perpetuating nature of social stereotypes. In D. L. Hamilton (Ed.), Cognitive processes in stereotyping and intergroup behavior (pp. 183-212). Hillsdale, NJ: Erlbaum.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. Journal of Experimental Social Psychology, 35, 4-28.

Stangor, C., Lynch, L., Duan, C., & Glass, B. (1992). Categorization of individuals on the basis of multiple social features. Journal of Personality and Social Psychology, 62, 207-218.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. American Psychologist, 52, 613-629.

Stumpf, S. A., & London, M. (1981). Capturing rater policies in evaluating candidates for promotion. Academy of Management Journal, 24, 752-766.

Stroh, L. K., Brett, J. M., & Reilly, A. H. (1992). All the right stuff: A comparison of female and male managers' career progression. Journal of Applied Psychology, 77, 251-260.

Thompson, D. E., & Thompson, T. A. (1985). Task-based performance appraisal for blue-collar jobs: Evaluation of race and sex effects. Journal of Applied Psychology, 70, 747-753.

Tornow, W. W. (1993). Perceptions or reality: Is multi-perspective measurement a means or an end? Human Resource Management, 32, 221-230.

Tsui, A. S. (1984). A role set analysis of managerial reputation. Organizational Behavior and Human Performance, 34, 64-96.

Tsui, A. S., & Gutek, B. A. (1984). A role set analysis of gender differences in performance, affective relationships, and career success of industrial middle managers. Academy of Management Journal, 27, 619-635.

Tsui, A. S., & Ohlott, P. (1988). Multiple assessment of managerial effectiveness: Interrater agreement and consensus in effectiveness models. Personnel Psychology, 41, 779-803.

Tsui, A. S., & O'Reilly, C. A. (1989). Beyond simple demographic effects: The importance of relational demography in superior-subordinate dyads. Academy of Management Journal, 32, 402-423.

Tsui, A. S., Porter, L. W., & Egan, T. D. (2002). When both similarities and dissimilarities matter: Extending the concept of relational demography. Human Relations, 55, 899-929.

U.S. Bureau of the Census. (2000). Current population reports: Educational attainment in the United States: March 2000 (Table 1: Educational attainment of the population 15 years and over, by age, sex, race, and Hispanic origin). Retrieved from http://www.census.gov/population/socdemo/education/p20-536/tab01.txt.

U.S. Bureau of Labor Statistics. (2001a). News: The employment situation: July 2001 (Table A-1: Employment status of the civilian population by sex and age). Retrieved from http://www.bls.gov/news.release/pdf/empsit.pdf.

U.S. Bureau of Labor Statistics. (2001b). Demographic characteristics of the labor force (Table 11: Employed persons by detailed occupation, sex, race, and Hispanic origin). Retrieved from http://www.bls.gov/cps/cpsaat11.pdf.

U.S. Bureau of Labor Statistics. (2001c). Demographic characteristics of the labor force (Table 18: Employed persons by detailed industry, sex, race, and Hispanic origin). Retrieved from http://www.bls.gov/cps/cpsaat18.pdf.

Van Der Leeden, R., & Busing, F. M. T. A. (1994). First iteration versus igls/rigls estimates in two-level models: A monte carlo study with ML3. Preprint PRM 94-03, Psychometrics and Research Methodology, Leiden, Netherlands.

van Knippenberg, A., van Twuyver, M., & Pepels, J. (1994). Factors affecting social categorization processes in memory. British Journal of Social Psychology, 33, 419-431.

Vroom, V., & Yetton, P. (1973). Leadership and decision making. Pittsburgh, PA: University of Pittsburgh Press.

Wellington, S., Kropf, M. B., & Gerkovich, P. R. (2003). What's holding women back: As barriers shift, lack of line experience has become a chief obstacle. Harvard Business Review, 81, 18-19.

Williams, C. L. (1992). The glass escalator: Hidden advantages for men in the "female" professions. Social Problems, 39, 253-267.

Williams, C. L. (1995). Still a man's world: Men who do "women's work." Berkeley: University of California Press.

Woehr, D. J., & Roch, S. G. (1996). Context effects in performance evaluation: The impact of ratee sex and performance level on performance ratings and behavioral recall. Organizational Behavior and Human Decision Processes, 66, 31-41.

Wohlers, A. J., Hall, M. J., & London, M. (1993). Subordinates rating managers: Organizational and demographic correlates of self/subordinate agreement. Journal of Occupational and Organizational Psychology, 66, 263-275.

Yammarino, F. J., & Dubinsky, A. J. (1988). Employee responses: Gender or job related differences? Journal of Vocational Behavior, 32, 366-383.

Yoder, J. D., & Sinnett, L. M. (1985). Is it all in the numbers? A case study of tokenism. Psychology of Women Quarterly, 9, 413-418.

Yukl, G., & Van Fleet, D. D. (1992). Theory and research on leadership in organizations. In M. D. Dunnette & L. M. Hough (Eds.), Handbook of industrial and organizational psychology (2nd ed., vol. 3) (pp. 147-198). Palo Alto, CA: Consulting Psychologists Press.

Table 1. Number of Raters for Male and Female Leaders

Rater	# Raters for Female and Male Leaders	% Raters for Female and Male Leaders	# Raters for Female Leaders	% Raters for Female Leaders	# Raters for Male Leaders	% Raters for Male Leaders
Female Subordinates	4,259	19	1,799	31	2,368	15
Male Subordinates	5,630	26	768	13	4,722	30
Female Peers	2,686	12	1,218	21	1,421	9
Male Peers	6,591	30	1,239	21	5,189	33
Female Bosses	469	2	228	4	232	1
Male Bosses	2,343	11	512	9	1,768	11
TOTAL RATERS	21, 978	100	5,764	100	15,700	100

Table 2. Average Number of Raters Per Leader

Rater Group	Female Leaders	Male Leaders
Average Number of Total Raters	8.82	8.52
Average Number of Subordinate Raters	3.93	4.02
Average Number of Peer Raters	3.76	3.77
Average Number of Boss Raters	1.27	1.30

Table 3. Leader and Rater Demographics

Demographics	All Leaders		Male Leaders		Female Leaders		All Bosses		Male Bosses		Female Bosses	
Age	#	%	#	%	#	%	#	%	#	%	#	%
Under 20	0	0	0	0	0	0	1	0	1	0	0	0
20-29	147	4	100	4	47	5	41	1	20	1	19	4
30-39	1123	33	780	31	338	37	566	20	454	20	102	22
40-49	1473	43	1075	43	393	43	1221	43	973	42	222	48
50-59	659	19	535	21	123	14	875	31	752	32	112	24
60 or over	46	1	37	1	9	1	126	4	116	5	9	2
Education	#	%	#	%	#	%	#	%	#	%	#	%
High school	132	4	94	4	38	4	80	3	61	3	17	4
Some college	387	11	270	11	117	13	212	8	175	8	33	7
Associate's	166	5	119	5	47	5	113	4	83	4	30	7
Bachelor's	1052	31	769	31	275	30	794	28	654	29	128	28
Some graduate work	409	12	301	12	105	12	328	12	266	12	50	11
Master's	944	28	695	28	247	27	931	33	752	33	164	36
Doctoral	284	8	219	9	65	7	307	11	278	12	22	5
Other	50	1	32	1	18	2	30	1	23	1	7	2
Years	#	%	#	%	#	%	#	%	#	%	#	%
Less than 6 months	93	3	63	3	29	3	55	2	32	1	22	5
6 months-1 year	118	3	85	3	32	3	61	2	52	2	8	2
1-2 years	279	8	198	8	78	9	134	5	96	4	35	8
2-4 years	464	13	298	12	165	18	280	10	221	10	50	11
4-6 years	373	11	257	10	114	12	309	11	264	11	40	9
6-10 years	526	15	372	15	154	17	366	13	277	12	80	18
10-15 years	596	17	455	18	138	15	531	19	437	19	82	18
More than 15 years	994	29	785	31	206	22	1065	38	917	40	139	30
Salary	#	%	#	%	#	%	#	%	#	%	#	%
$25,000 or less	12	0	7	0	5	1	22	1	12	1	9	3
$25,001-35,000	38	1	20	1	18	2	23	1	8	1	15	5
$35,001-50,000	158	6	85	4	73	10	59	3	40	3	19	6
$50,001-75,000	763	29	511	27	251	34	131	7	87	6	40	12
$75,001-100,000	634	24	439	23	195	26	305	16	216	14	86	26
$100,001-150,000	598	22	471	24	122	17	475	25	386	25	81	24
$150,001-200,000	213	8	181	9	32	4	395	21	351	23	38	11
$200,001 or over	252	9	211	11	40	5	507	26	456	29	43	13
Ethnicity	#	%	#	%	#	%	#	%	#	%	#	%
Asian	117	4	81	4	36	5	51	2	41	2	8	2
African American	64	2	38	2	25	3	33	2	23	1	10	3
Hispanic	74	3	54	3	20	3	78	4	68	4	8	2
White/Caucasian	2281	86	1640	87	635	86	1948	89	1584	88	333	90
Other	102	4	77	4	25	3	90	4	81	5	9	2

Table 3 (continued). Leader and Rater Demographics

Demographic	All Peers		Male Peers		Female Peers		All Subord		Male Subord		Female Subord	
Age	#	%	#	%	#	%	#	%	#	%	#	%
Under 20	6	0	1	0	3	0	11	0	7	0	2	0
20-29	424	5	239	4	179	7	1211	12	525	10	655	16
30-39	2878	31	1932	30	858	33	3515	35	1934	35	1455	35
40-49	3890	42	2646	41	1141	44	3347	34	1910	35	1333	32
50-59	1937	21	1500	23	394	15	1679	17	1010	18	631	15
60 or over	178	2	143	2	34	1	170	2	109	2	57	1
Education	#	%	#	%	#	%	#	%	#	%	#	%
High school	403	4	210	3	179	7	745	8	345	6	374	9
Some college	1013	11	625	10	361	14	1756	18	804	15	911	22
Associate's	507	6	361	6	134	5	977	10	465	9	479	12
Bachelor's	2948	32	2086	33	775	30	3199	32	1873	34	1213	30
Some graduate work	930	10	648	10	247	10	857	9	474	9	338	8
Master's	2478	27	1743	27	665	26	1725	17	1092	20	580	14
Doctoral	847	9	647	10	183	7	424	4	313	6	101	2
Other	115	1	71	1	42	2	183	2	81	1	91	2
Years	#	%	#	%	#	%	#	%	#	%	#	%
Less than 6 months	116	1	72	1	36	1	333	3	185	3	142	3
6 months-1 year	270	3	187	3	74	3	565	6	283	5	265	7
1-2 years	679	7	441	7	217	8	1005	10	506	9	479	12
2-4 years	1343	15	844	13	465	18	1823	19	908	17	841	21
4-6 years	1073	12	693	11	348	14	1217	12	683	13	501	12
6-10 years	1445	16	978	15	432	17	1491	15	819	15	631	15
10-15 years	1667	18	1213	19	412	16	1334	14	773	14	523	13
More than 15 years	2566	28	1910	30	587	23	1987	20	1244	23	693	17
Salary	#	%	#	%	#	%	#	%	#	%	#	%
$25,000 or less	111	2	43	1	63	3	418	6	109	3	288	9
$25,001-35,000	218	3	90	2	112	6	766	10	197	5	549	17
$35,001-50,000	487	7	248	5	234	12	1394	19	489	12	875	27
$50,001-75,000	1478	22	870	19	579	30	2313	31	1386	34	884	28
$75,001-100,000	1531	23	1071	24	431	22	1192	16	836	20	330	10
$100,001-150,000	1592	24	1227	27	339	18	916	12	719	17	181	6
$150,001-200,000	561	9	457	10	98	5	255	3	215	5	33	1
$200,001 or over	597	9	509	11	74	4	231	3	167	4	52	2
Ethnicity	#	%	#	%	#	%	#	%	#	%	#	%
Asian	264	4	201	4	56	3	344	4	218	5	117	4
African American	167	2	91	2	73	4	245	3	90	2	154	5
Hispanic	197	3	153	3	42	2	249	3	143	3	104	3
White/Caucasian	6253	88	4340	88	1749	88	6478	85	3576	85	2709	84
Other	236	3	149	3	77	4	332	4	164	4	153	5

Table 4. Denison Leadership Development Survey Skills, Dimensions, and Subdimensions (Denison & Neale, 1996)

Type of Skill	Dimension	Subdimension
External Focus Flexible	Adaptability	Promotes Organizational Learning
		Emphasizes Customer Focus
		Creates Change
External Focus Stable	Mission	Defines Strategic Direction and Intent
		Defines Goals and Objectives
		Creates Shared Vision
Internal Focus Flexible	Involvement	Empowers People
		Builds Team Orientation
		Develops Organizational Capability
Internal Focus Stable	Consistency	Manages Coordination and Integration
		Works to Reach Agreement
		Defines Core Values

Table 5. Denison Leadership Development Survey Items (Denison & Neale, 1996)

Dimension	Sub-Dimension	Items
Involvement (Internal Focus Flexible)	Empowers People	1. Sees that decisions are made at the lowest possible level. 2. Shares information so that everyone gets the information s/he needs. 3. Creates an environment where everyone feels that his/her effort can make a difference. 4. Involves everyone in shaping the plans and decisions that affect them. 5. Ensures that the necessary resources are available to do the job. 6. Conveys confidence in people's competence to do their job. 7. Encourages others to take responsibility. 8. Delegates authority so that others can do their work more effectively.
	Builds Team Orientation	9. Builds effective teams that get the job done. 10. Encourages effective teamwork by others. 11. Knows how to use a team approach to solve problems. 12. Knows when to use a team approach to solve problems. 13. Fosters teamwork within the work unit. 14. Knows how to design work so that it can be done by a team. 15. Values the contributions of the people s/he works with. 16. Acknowledges and celebrates team accomplishments.
	Develops Organizational Capability	17. Builds the capabilities of employees into an important source for competitive advantage. 18. Knows how to utilize the diversity of the workforce. 19. Coaches others in the development of their skills. 20. Is sensitive and responsive to diversity issues when dealing with others. 21. Helps direct reports create realistic development plans and create opportunities for them. 22. Uses rewards and recognition to motivate good performance. 23. Develops his/her own people so that they are ready for promotion. 24. Builds employee skills so that the organization always has good "bench strength".
Consistency (Internal Focus Stable)	Defines Core Values	25. Does the "right thing" even when it is not popular. 26. "Practices" what s/he "preaches". 27. Has an ethical code that guides his/her behavior. 28. Helps define the organization's culture, values, and ethical standards. 29. Helps employees learn to apply the organization's values when dealing with others. 30. Lives up to promises and commitments. 31. Has earned the confidence and trust of others. 32. Clearly articulates a set of fundamental beliefs that are not negotiable.
	Works to Reach Agreement	33. Helps people to reach consensus, even on difficult issues. 34. Works to find alternatives that will benefit all when confronted with a disagreement. 35. Helps people in his/her organization be effective at reaching agreement on key issues. 36. Incorporates diverse points of view when making decisions. 37. Promotes constructive discussion among people with conflicting ideas. 38. Is willing to compromise when necessary in order to reach agreement. 39. Works toward win/win solutions when disagreements occur. 40. Reconciles differences by seeking to clarify and understand others.
	Manages Coordination and Integration	41. Works hard to foster the alignment of goals across all functional areas. 42. Builds coordination across departmental boundaries. 43. Uses informal networks to get things done. 44. Builds relationships with key people in other functions and levels. 45. Helps create an environment that facilitates coordination of projects across functional units. 46. Makes certain that things do not "fall between the cracks". 47. Builds support for ideas through contracts with other departments. 48. Establishes mechanisms that facilitate effective cross-functional communication.

Table 5 (continued). Denison Leadership Development Survey Items (Denison & Neale, 1996)

Dimension	Sub-Dimension	Items
Adaptability (External Focus Flexible)	Creates Change	49. Continuously looks for new and better ways to do work. 50. Encourages creative thinking. 51. Challenges the way that things have always been done and looks for a better way. 52. Champions change that goes beyond the scope of his/her job. 53. Challenges organizational practices that are nonproductive. 54. Foresees problems before they arise. 55. Serves as a model that creates change in other parts of the organization. 56. Generates innovative ideas and solutions to problems.
	Emphasizes Customer Focus	57. Encourages direct contact with customers. 58. Responds quickly and effectively to customer feedback. 59. Ensures that employees have a deep understanding of customer wants and needs. 60. Uses customer comments and recommendations to change organizational practices. 61. Actively seeks feedback from customers. 62. Continuously tries to improve service to customers. 63. Incorporates customer input into the planning process. 64. Recognizes the need to respond quickly to customer concerns.
	Promotes Organizational Learning	65. Deals constructively with failures and mistakes. 66. Views failures as an opportunity for learning and improvement. 67. Creates a working environment in which learning is an important objective. 68. Openly accepts criticism without being defensive. 69. Works well under conditions of ambiguity and uncertainty. 70. Knows the strengths and weaknesses of the competition. 71. Encourages others to learn about the best practices in the industry. 72. Helps others to understand "the big picture".
Mission (External Focus Stable)	Defines Strategic Direction and Intent	73. Provides employees with a clear mission that gives meaning and direction to their work. 74. Implements strategies by developing clear goals, objectives, and tactics. 75. Focuses on long-term strategies, rather than quick fix "band-aid" solutions. 76. Effectively allocates resources in line with strategic priorities. 77. Helps define strategies and tactics that keep his/her organization competitive. 78. Has a clear strategy for the future of his/her own part of the organization. 79. Is able to meet short-term demands without losing sight of the long-term strategy. 80. Communicates a clear and compelling rationale for the business strategy.
	Defines Goals and Objectives	81. Sets clear goals that are ambitious, but realistic. 82. Holds individuals and teams accountable for achieving goals and objectives. 83. Provides clear directions and priorities for employees. 84. Establishes high standards of performance. 85. Involves employees in the goal-setting process so goals and objectives are understood and shared. 86. Tracks progress against stated goals. 87. Effectively communicates the goals and objectives of the organization. 88. Aligns goals and objectives with the strategy and vision.
	Creates Shared Vision	89. Helps create a shared vision of what this organization will be like in the future. 90. Communicates the organizational vision to his/her employees. 91. Uses the vision to create excitement and motivation for employees. 92. Realizes short-term goals without compromising long-term vision. 93. Organizes work so that everyone sees the connection between the vision and daily activities. 94. Translates the vision into reality in a way that helps guide individual action. 95. Inspires others with his/her vision of the future. 96. Engages others in ways that ensure buy-in and commitment.

Table 6. Agentic and Communal Development Item Ratings

#	DLDS Item	M	SD	%Rated 1 or 2	%Rated 4 or 5	Agentic/ Communal
1	Sees that decisions are made at the lowest possible level.	4.14	0.85	5	81	C
2	Shares information so that everyone gets the information s/he needs.	4.76	0.44	0	100	C
3	Creates an environment where everyone feels that his/her effort can make a difference.	4.81	0.40	0	100	C
4	Involves everyone in shaping the plans and decisions that affect them.	4.71	0.72	5	95	C
5	Ensures that the necessary resources are available to do the job.	3.02	0.64	19	24	--
6	Conveys confidence in people's competence to do their job.	4.14	0.57	0	90	C
7	Encourages others to take responsibility.	3.67	1.32	33	62	C
8	Delegates authority so that others can do their work more effectively.	3.14	1.42	33	48	--
9	Builds effective teams that get the job done.	3.36	1.20	24	57	--
10	Encourages effective teamwork by others.	4.24	0.44	0	100	C
11	Knows how to use a team approach to solve problems.	3.86	0.85	5	67	C
12	Knows when to use a team approach to solve problems.	3.95	0.74	5	81	C
13	Fosters teamwork within the work unit.	4.38	0.50	0	100	C
14	Knows how to design work so that it can be done by a team.	3.57	0.87	14	62	C
15	Values the contributions of the people s/he works with.	4.43	0.68	0	90	C
16	Acknowledges and celebrates team accomplishments.	4.60	0.58	0	95	C
17	Builds the capabilities of employees into an important source for competitive advantage.	2.57	1.29	57	19	--
18	Knows how to utilize the diversity of the workforce.	3.40	0.83	14	48	--
19	Coaches others in the development of their skills.	4.24	0.77	5	90	C
20	Is sensitive and responsive to diversity issues when dealing with others.	4.57	0.60	0	95	C
21	Helps direct reports create realistic development plans and create opportunities for them.	3.35	0.99	25	50	--
22	Uses rewards and recognition to motivate good performance.	3.07	1.05	33	33	--
23	Develops his/her own people so that they are ready for promotion.	3.50	1.05	19	62	C
24	Builds employee skills so that the organization always has good "bench strength".	3.05	1.12	43	33	--

Table 6 (continued). Agentic and Communal Development Item Ratings

#	DLDS Item	M	SD	%Rated 1 or 2	%Rated 4 or 5	Agentic/ Communal
25	Does the "right thing" even when it is not popular.	2.55	1.24	52	24	--
26	"Practices" what s/he "preaches".	3.05	0.97	24	24	--
27	Has an ethical code that guides his/her behavior.	3.52	0.75	5	50	C
28	Helps define the organization's culture, values, and ethical standards.	3.14	0.85	14	29	--
29	Helps employees learn to apply the organization's values when dealing with others.	3.69	0.68	0	57	C
30	Lives up to promises and commitments.	3.38	0.74	10	43	--
31	Has earned the confidence and trust of others.	3.33	1.06	19	57	--
32	Clearly articulates a set of fundamental beliefs that are not negotiable.	1.33	0.58	95	0	A
33	Helps people to reach consensus, even on difficult issues.	4.07	0.93	5	86	C
34	Works to find alternatives that will benefit all when confronted with a disagreement.	4.48	0.51	0	100	C
35	Helps people in his/her organization be effective at reaching agreement on key issues.	4.02	0.72	5	86	C
36	Incorporates diverse points of view when making decisions.	4.24	0.54	0	95	C
37	Promotes constructive discussion among people with conflicting ideas.	4.10	0.70	5	90	C
38	Is willing to compromise when necessary in order to reach agreement.	4.10	0.62	5	95	C
39	Works toward win/win solutions when disagreements occur.	4.10	1.00	10	90	C
40	Reconciles differences by seeking to clarify and understand other's points of view.	4.29	0.46	0	100	C
41	Works hard to foster the alignment of goals across all functional areas.	3.07	0.84	29	33	--
42	Builds coordination across departmental boundaries.	3.50	0.81	14	62	C
43	Uses informal networks to get things done.	3.02	1.25	33	48	--
44	Builds relationships with key people in other functions and levels.	2.67	1.20	48	29	--
45	Helps create an environment that facilitates coordination of projects across functional units.	3.19	0.98	29	52	--
46	Makes certain that things do not "fall between the cracks".	2.60	0.86	48	14	--
47	Builds support for ideas through contracts with other departments.	2.71	0.96	52	29	--
48	Establishes mechanisms that facilitate effective cross-functional communication.	3.57	0.81	14	67	C

Table 6 (continued). Agentic and Communal Development Item Ratings

#	DLDS Item	M	SD	%Rated 1 or 2	%Rated 4 or 5	Agentic/ Communal
49	Continuously looks for new and better ways to do work.	2.38	0.74	57	5	A
50	Encourages creative thinking.	3.21	0.68	14	38	--
51	Challenges the way that things have always been done and looks for a better way.	2.05	0.80	76	5	A
52	Champions change that goes beyond the scope of his/her job.	2.29	0.96	51	5	A
53	Challenges organizational practices that are nonproductive.	1.64	0.57	95	0	A
54	Foresees problems before they arise.	2.60	0.54	38	0	--
55	Serves as a model that creates change in other parts of the organization.	2.36	0.65	57	0	A
56	Generates innovative ideas and solutions to problems.	2.55	0.77	38	5	--
57	Encourages direct contact with customers.	3.40	0.97	14	43	--
58	Responds quickly and effectively to customer feedback.	3.48	0.68	5	48	--
59	Ensures that employees have a deep understanding of customer wants and needs.	3.88	0.63	0	76	C
60	Uses customer comments and recommendations to change organizational practices.	3.40	0.74	14	57	--
61	Actively seeks feedback from customers.	3.79	0.60	0	71	C
62	Continuously tries to improve service to customers.	3.45	0.77	14	52	--
63	Incorporates customer input into the planning process.	3.81	0.60	5	81	C
64	Recognizes the need to respond quickly to customer concerns.	3.40	0.70	10	43	--
65	Deals constructively with failures and mistakes.	3.07	0.87	24	38	--
66	Views failures as an opportunity for learning and improvement.	3.29	0.85	14	48	--
67	Creates a working environment in which learning is an important objective.	3.43	0.98	19	62	--
68	Openly accepts criticism without being defensive.	3.19	1.17	24	52	--
69	Works well under conditions of ambiguity and uncertainty.	2.52	0.87	43	10	--
70	Knows the strengths and weaknesses of the competition.	1.95	0.74	76	0	A
71	Encourages others to learn about the best practices in the industry.	2.71	1.01	48	29	--
72	Helps others to understand "the big picture".	3.14	1.06	19	48	--

Table 6 (continued). Agentic and Communal Development Item Ratings

#	DLDS Item	M	SD	%Rated 1 or 2	%Rated 4 or 5	Agentic/ Communal
73	Provides employees with a clear mission that gives meaning and direction to their work.	2.48	1.12	62	29	A
74	Implements strategies by developing clear goals, objectives, and tactics.	1.67	0.66	90	0	A
75	Focuses on long-term strategies, rather than quick fix "band-aid" solutions.	2.36	0.82	57	10	A
76	Effectively allocates resources in line with strategic priorities.	1.88	0.71	81	0	A
77	Helps define strategies and tactics that keep his/her organization competitive.	1.79	0.93	76	5	A
78	Has a clear strategy for the future of his/her own part of the organization.	1.62	0.74	86	0	A
79	Is able to meet short-term demands without losing sight of the long-term strategy.	2.21	0.82	57	0	A
80	Communicates a clear and compelling rationale for the business strategy.	1.98	0.95	76	10	A
81	Sets clear goals that are ambitious, but realistic.	1.74	0.62	90	0	A
82	Holds individuals and teams accountable for achieving goals and objectives.	1.76	0.62	90	0	A
83	Provides clear directions and priorities for employees.	1.83	0.58	90	0	A
84	Establishes high standards of performance.	1.62	0.67	90	0	A
85	Involves employees in the goal-setting process so goals and objectives are understood and shared.	3.98	0.68	5	86	C
86	Tracks progress against stated goals.	2.26	0.80	62	5	A
87	Effectively communicates the goals and objectives of the organization.	2.62	0.80	48	14	--
88	Aligns goals and objectives with the strategy and vision.	2.14	0.79	62	0	A
89	Helps create a shared vision of what this organization will be like in the future.	3.10	1.09	33	52	--
90	Communicates the organizational vision to his/her employees.	2.98	1.01	33	38	--
91	Uses the vision to create excitement and motivation for employees.	3.12	1.12	29	43	--
92	Realizes short-term goals without compromising long-term vision.	2.55	0.80	48	10	--
93	Organizes work so that everyone sees the connection between the vision and daily activities.	3.38	0.67	10	48	--
94	Translates the vision into reality in a way that helps guide individual action.	2.48	0.75	50	5	A
95	Inspires others with his/her vision of the future.	2.62	1.02	43	14	--
96	Engages others in ways that ensure buy-in and commitment.	3.43	0.93	14	57	--

Table 7. Number of Leaders in Each Organizational Level

Organizational Level	# Female and Male Leaders	% Female and Male Leaders	# Female Leaders	% Female Leaders	# Male Leaders	% Male Leaders
Line Management	695	22	224	27	470	20
Middle Management	1286	40	344	42	938	40
Executive	1203	38	263	32	934	40
TOTAL RATERS	3184	100	821	100	2342	100

Table 8. Number of Leaders in Each Functional Area

Functional Area	# Female and Male Leaders	% Female and Male Leaders	# Female Leaders	% Female Leaders	# Male Leaders	% Male Leaders
Finance and Accounting	298	9	80	9	216	9
Engineering	222	7	22	3	199	8
Manufacturing and Production	266	8	28	3	238	10
Research and Development	177	5	49	6	128	5
Sales and Marketing	689	21	170	20	513	22
Purchasing	141	4	32	4	109	5
Human Resources	167	5	94	11	73	3
Administration	581	18	168	20	411	17
Support Staff	149	5	32	4	117	5
Professional Staff	539	17	181	21	355	15
TOTAL	3229	100	856	100	2359	100

Table 9. Number of Leaders in Functional Area Percent-Female Ranges

% Females in Each Functional Area[a]	# Female and Male Leaders	% Female and Male Leaders	# Female Leaders	% Female Leaders	# Male Leaders	% Male Leaders
Above 60%	316	10	126	15	190	8
Between 40%-60%	2248	70	631	74	1604	68
Below 40%	665	21	99	12	565	24
TOTAL	3229	100	856	100	2359	100

[a]Complete range is 10.4%-78.7%

Table 10. Number of Leaders in Each Major Industry

Industry	# Female and Male Leaders	% Female and Male Leaders	# Female Leaders	% Female Leaders	# Male Leaders	% Male Leaders
Construction	306	10	46	6	240	11
Manufacturing	798	27	121	17	640	30
Transportation, Communications, and Other Public Utilities	180	6	63	9	115	5
Wholesale and Retail Trade	409	14	80	11	311	15
Finance, Insurance, and Real Estate	336	11	125	17	206	10
Services	614	20	212	29	359	17
Public Administration	353	12	74	10	273	13
TOTAL	2996	100	721	100	2144	100

Table 11. Number of Leaders in Industry Percent-Female Ranges

% Females in Each Industry[a]	# Female and Male Leaders	% Female and Male Leaders	# Female Leaders	% Female Leaders	# Male Leaders	% Male Leaders
Above 60%	751	25	325	45	408	19
Between 40%-60%	763	25	191	26	510	24
Below 40%	1482	49	205	28	1226	57
TOTAL	2996	100	721	100	2144	100

[a]Complete range is 9.7%-92.0%

Table 12. Number of Leaders in Proportion of Male Raters Ranges

% of Male Raters[a]	# Female and Male Leaders	% Female and Male Leaders	# Female Leaders	% Female Leaders	# Male Leaders	% Male Leaders
Above 60%	2067	54	269	27	1727	64
Between 40%-60%	861	23	246	25	586	22
Below 40%	882	23	473	48	389	14
TOTAL	3810	100	988	100	2702	100

[a]Complete range is 0.0%-100.0%

Table 13. Summary of Study Hypotheses, Analyses, and Results

Hypotheses	Independent Variables[a]	Moderators	Dependent Variables	Analyses	Results
H1: Industry will moderate the relationship between leader gender and others' (combined boss, peer, and subordinate) ratings. In male-dominated industries, female leaders will receive lower overall development ratings (i.e. needing more agentic and communal development) and lower effectiveness ratings (i.e. performing less effectively) than male leaders. In female-dominated industries, no differences are expected between male and female leaders with regard to others' overall development and effectiveness ratings.	Leader gender^{L2}	Industry^{L3}	Others' development ratings^{L1} Others' effectiveness ratings^{L1}	HLM: 3 levels 2-way interaction	Not Supported
H2: Functional area will moderate the relationship between leader gender and others' (combined boss, peer, and subordinate) ratings. In male-dominated functional areas, female leaders will receive lower overall development ratings (i.e. needing more agentic and communal development) and lower effectiveness ratings (i.e. performing less effectively) than male leaders. In female-dominated functional areas, no differences are expected between male and female leaders with regard to others' overall development and effectiveness ratings.	Leader gender^{L2}	Functional area^{L3}	Others' development ratings^{L1} Others' effectiveness ratings^{L1}	HLM: 3 levels 2-way interaction	Not Supported

[a] The level of each variable is indicated in superscript for each hypothesis. Note that the level of each variable may vary depending on what the lowest level of analysis is for each hierarchical model. For example, leader gender may be a level-2 variable when included in an analysis modeling its effects on lower, individual-level variables such as peer ratings. However, it may also be a level-1 variable when modeling its effects on another leader-level variable such as self-ratings. In the case of leader self ratings, there is only one rating per leader and therefore is a leader level variable.

Table 13 (continued). Summary of Study Hypotheses, Analyses, and Results

Hypotheses	Independent Variables[a]	Moderators	Dependent Variables	Analyses	Results
H3: The proportion of male raters (i.e. percentage of male boss, peer, and subordinate raters) will moderate the relationship between leader gender and others' (combined boss, peer, and subordinate) ratings. When the proportion of male raters is high, female leaders will receive lower overall development ratings (i.e. needing more agentic and communal development) and lower effectiveness ratings (i.e. performing less effectively) than male leaders. When the proportion of male raters is low, no differences are expected between male and female leaders with regard to others' overall development and effectiveness ratings.	Leader gender^{L2}	Proportion of male raters^{L2}	Others' development ratings^{L1} Others' effectiveness ratings^{L1}	HLM: 2 levels 2-way interaction	Not Supported
H4: Leader gender will moderate the relationship between others' (combined boss, peer, and subordinate) overall development (agentic and communal) and effectiveness ratings such that the two sets of ratings will be more strongly, positively related for male leaders than for female leaders. In other words, ratings indicating a lesser need for development will be more strongly related to ratings indicating greater effectiveness for men than for women.	Others' development ratings^{L1}	Leader gender^{L2}	Others' effectiveness ratings^{L1}	HLM: 2 levels 2-way interaction	Not Supported
H5a: Female leaders will receive lower communal development ratings (i.e. needing more development in communal skill areas) from others (combined bosses, peers, and subordinates) as compared to their male counterparts. The discrepancy between female and male ratings is expected to be more pronounced in male-dominated industries.	Leader gender^{L2}	Industry^{L3}	Others' communal ratings^{L1}	HLM: 3 levels 2-way interaction	Not Supported

Table 13 (continued). Summary of Study Hypotheses, Analyses, and Results

Hypotheses	Independent Variables[a]	Moderators	Dependent Variables	Analyses	Results
H5b: Female leaders will receive lower communal development ratings (i.e. needing more development in communal skill areas) from others (combined bosses, peers, and subordinates) as compared to their male counterparts. The discrepancy between female and male ratings is expected to be more pronounced in male-dominated functional areas.	Leader gender[L2]	Functional area[L3]	Others' communal ratings[L1]	HLM: 3 levels 2-way interaction	Not Supported
H5c: Female leaders will receive lower communal development ratings (i.e. needing more development in communal skill areas) from others (combined bosses, peers, and subordinates) as compared to their male counterparts. The discrepancy between female and male ratings is expected to be more pronounced as the proportion of male raters increases.	Leader gender[L2]	Proportion of male raters[L2]	Others' communal ratings[L1]	HLM: 2 levels 2-way interaction	Not Supported
H6a: Male bosses will provide higher communal and agentic development ratings (i.e. needing less development) and higher effectiveness ratings (i.e. performing more effectively) to male leaders as compared to female leaders. The discrepancy between male and female ratings is expected to be more pronounced in male-dominated industries.	Leader gender[L2]	Industry[L3]	Male bosses' communal ratings[L1] Male bosses' agentic ratings[L1] Male bosses' effectiveness ratings[L1]	HLM: 3 levels 2-way interaction	Not Supported

Table 13 (continued). Summary of Study Hypotheses, Analyses, and Results

Hypotheses	Independent Variables[a]	Moderators	Dependent Variables	Analyses	Results
H6b: Male bosses will provide higher communal and agentic development ratings (i.e. needing less development) and higher effectiveness ratings (i.e. performing more effectively) to male leaders as compared to female leaders. The discrepancy between male and female ratings is expected to be more pronounced in male-dominated functional areas.	Leader gender[L2]	Functional area[L3]	Male bosses' communal ratings[L1] Male bosses' agentic ratings[L1] Male bosses' effectiveness ratings[L1]	HLM: 3 levels 2-way interaction	Not Supported
H6c: Male bosses will provide higher communal and agentic development ratings (i.e. needing less development) and higher effectiveness ratings (i.e. performing more effectively) to male leaders as compared to female leaders. The discrepancy between male and female ratings is expected to be more pronounced as the proportion of male raters increases.	Leader gender[L2]	Proportion of male raters[L2]	Male bosses' communal ratings[L1] Male bosses' agentic ratings[L1] Male bosses' effectiveness ratings[L1]	HLM: 2 levels 2-way interaction	Not Supported
H7a: Female bosses will provide higher communal and agentic development ratings (i.e. needing less development) and higher effectiveness ratings (i.e. performing more effectively) to female leaders as compared to male leaders. The discrepancy between female and male ratings is expected to be more pronounced in male-dominated industries.	Leader gender[L2]	Industry[L3]	Female bosses' communal ratings[L1] Female bosses' agentic ratings[L1] Female bosses' effectiveness ratings[L1]	HLM: 3 levels 2-way interaction	Not Supported

Table 13 (continued). Summary of Study Hypotheses, Analyses, and Results

Hypotheses	Independent Variables[a]	Moderators	Dependent Variables	Analyses	Results
H7b: Female bosses will provide higher communal and agentic development ratings (i.e. needing less development) and higher effectiveness ratings (i.e. performing more effectively) to female leaders as compared to male leaders. The discrepancy between female and male ratings is expected to be more pronounced in male-dominated functional areas.	Leader gender[L2]	Functional area[L3]	Female bosses' communal ratings[L1] Female bosses' agentic ratings[L1] Female bosses' effectiveness ratings[L1]	HLM: 3 levels 2-way interaction	Not Supported
H7c: Female bosses will provide higher communal and agentic development ratings (i.e. needing less development) and higher effectiveness ratings (i.e. performing more effectively) to female leaders as compared to male leaders. The discrepancy between female and male ratings is expected to be more pronounced as the proportion of male raters increases.	Leader gender[L2]	Proportion of male raters[L2]	Female bosses' communal ratings[L1] Female bosses' agentic ratings[L1] Female bosses' effectiveness ratings[L1]	HLM: 2 levels 2-way interaction	Not Supported
H8a: Female peers will provide higher communal and agentic development ratings (i.e. needing less development) and higher effectiveness ratings (i.e. performing more effectively) to female leaders as compared to male leaders. The discrepancy between female and male ratings is expected to be more pronounced in female-dominated industries.	Leader gender[L2]	Industry[L3]	Female peers' communal ratings[L1] Female peers' agentic ratings[L1] Female peers' effectiveness ratings[L1]	HLM: 3 levels 2-way interaction	Partially Supported

Table 13 (continued). Summary of Study Hypotheses, Analyses, and Results

Hypotheses	Independent Variables[a]	Moderators	Dependent Variables	Analyses	Results
H8b: Female peers will provide higher communal and agentic development ratings (i.e. needing less development) and higher effectiveness ratings (i.e. performing more effectively) to female leaders as compared to male leaders. The discrepancy between female and male ratings is expected to be more pronounced in female-dominated functional areas.	Leader gender[L2]	Functional area[L3]	Female peers' communal ratings[L1] Female peers' agentic ratings[L1] Female peers' effectiveness ratings[L1]	HLM: 3 levels 2-way interaction	Not Supported
H8c: Female peers will provide higher communal and agentic development ratings (i.e. needing less development) and higher effectiveness ratings (i.e. performing more effectively) to female leaders as compared to male leaders. The discrepancy between female and male ratings is expected to be more pronounced as the proportion of male raters decreases.	Leader gender[L2]	Proportion of male raters[L2]	Female peers' communal ratings[L1] Female peers' agentic ratings[L1] Female peers' effectiveness ratings[L1]	HLM: 2 levels 2-way interaction	Partially Supported
H9: Rater type (boss or subordinate) and leader gender will interact in their influence on the relationship between agentic and effectiveness ratings, and on the relationship between communal and effectiveness ratings. It is expected that the relationship between agentic and effectiveness ratings will be higher for bosses than for subordinates, more so for men than for women. In addition, the relationship between communal and effectiveness ratings will be higher for subordinates than for bosses, more so for men than for women.	Bosses' and subordinates' communal ratings[L1] Bosses' and subordinates' agentic ratings[L1]	Rater type[L1] Leader gender[L2]	Bosses' and subordinates' effectiveness ratings[L1]	HLM: 2 levels 3-way interaction	Not Supported

Table 13 (continued). Summary of Study Hypotheses, Analyses, and Results

Hypotheses	Independent Variables[a]	Moderators	Dependent Variables	Analyses	Results
H10a: Industry will moderate the relationship between leader gender and self-ratings. In male-dominated industries, female leaders will give themselves lower agentic and communal development ratings (i.e. needing more agentic and communal development) than will male leaders. In female-dominated industries, no differences are expected between male and female leaders with regard to self-ratings of agentic and communal development needs.	Leader gender[L1]	Industry[L2]	Self communal ratings[L1] Self agentic ratings[L1]	HLM: 2 levels 2-way interaction	Partially Supported
H10b: Functional area will moderate the relationship between leader gender and self-ratings. In male-dominated functional areas, female leaders will give themselves lower agentic and communal development ratings (i.e. needing more agentic and communal development) than will male leaders. In female-dominated functional areas, no differences are expected between male and female leaders with regard to self-ratings of agentic and communal development needs.	Leader gender[L1]	Functional area[L2]	Self communal ratings[L1] Self agentic ratings[L1]	HLM: 2 levels 2-way interaction	Not Supported
H10c: The proportion of male raters will moderate the relationship between leader gender and self-ratings. When the proportion of male raters is high, female leaders will give themselves lower agentic and communal development ratings (i.e. needing more agentic and communal development) than will male leaders. When the proportion of male raters is low, no differences are expected between male and female leaders with regard to self-ratings of agentic and communal development needs.	Leader gender[L1]	Proportion of male raters[L1]	Self communal ratings[L1] Self agentic ratings[L1]	MANOVA 2-way interaction	Not Supported
H11: Female leader self-ratings will be higher for communal skills (i.e. needing less development in communal skill areas) than male leader self-ratings of communal skills.	Leader gender[L1]	None	Self communal ratings[L1]	ANOVA 1-way	Supported

Table 13 (continued). Summary of Study Hypotheses, Analyses, and Results

Hypotheses	Independent Variables[a]	Moderators	Dependent Variables	Analyses	Results
H12: Leader gender will moderate the relationship between others' (combined boss, peer, and subordinate) effectiveness ratings and agentic development self-ratings such that the two sets of ratings will be more strongly, positively related for male than for female leaders. In other words, self ratings indicating a lesser need for development of agentic skills will be more strongly related to others' ratings indicating greater effectiveness for men than for women.	Self agentic ratings ^{L2}	Leader gender^{L2}	Others' effectiveness ratings^{L1}	HLM: 2 levels 2-way interaction	Not Supported
H13: Self-subordinate rater agreement on communal and agentic development ratings will be higher for female leaders than for male leaders.	Leader gender ^{L1}	None	Self-subordinate rater agreement on communal ratings ^{L1} Self-subordinate rater agreement on agentic ratings ^{L1}	MANOVA 1-way	Not Supported
H14: Leader level will moderate the relationship between leader gender and others'(combined boss, peer, and subordinate) ratings. At the executive level, female leaders will receive lower agentic and communal development ratings (i.e. needing more development in agentic and communal leadership skills) and lower effectiveness ratings (i.e. performing less effectively) than male leaders. At the middle-management level, no differences are expected between male and female leaders with regard to others' ratings.	Leader gender ^{L2}	Leader level ^{L2}	Others' communal ratings^{L1} Others' agentic ratings^{L1} Others' effectiveness ratings^{L1}	HLM: 2 levels 2-way interaction	Not Supported

Table 14. Means and Standard Deviations for Male and Female Leaders

Raters	Female Leaders									Male Leaders								
	Agentic			Commmunal			Effectiveness			Agentic			Communal			Effectiveness		
	N	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*	*N*	*Mean*	*SD*
Self	977	5.63	0.68	977	5.83	0.52	--	--	--	2695	5.66	0.64	2695	5.78	0.54	--	--	--
Female Subordinates	1798	5.64	0.92	1799	5.69	0.89	1627	5.90	1.14	2360	5.57	0.95	2361	5.56	0.95	2035	5.78	1.22
Male Subordinates	768	5.67	0.86	768	5.73	0.81	661	5.90	1.06	4700	5.53	0.91	4702	5.53	0.86	4193	5.72	1.13
Female Peers	1216	5.63	0.80	1217	5.69	0.78	1083	5.63	1.09	1415	5.49	0.88	1415	5.50	0.88	1181	5.47	1.17
Male Peers	1238	5.42	0.82	1238	5.53	0.79	1081	5.41	1.09	5169	5.43	0.84	5173	5.45	0.80	4557	5.37	1.13
Female Bosses	228	5.62	0.82	228	5.72	0.75	204	5.67	1.00	231	5.50	0.83	232	5.60	0.79	198	5.62	1.14
Male Bosses	510	5.57	0.77	512	5.63	0.71	452	5.62	0.97	1779	5.56	0.77	1782	5.60	0.71	1580	5.64	0.98

Table 15. Summary of HLM Descriptive Main Effects of Leader Gender and Leader Level

Independent Variables[a]	Dependent Variables	γ^b	Level 2 Parameter Estimates				
			Coef	SE	df	T[c]	R^2
Leader gender ^{L2}	Others' communal ratings ^{L1}	γ_{00}	5.5103	0.0087	3311	632.72*	
		γ_{01}	-0.1196	0.0194	3311	**-6.16***	0.0037
	Others' agentic ratings ^{L1}	γ_{00}	5.4806	0.0091	3310	604.91*	
		γ_{01}	-0.0741	0.0200	3310	**-3.70***	0.0013
	Others' effectiveness ratings ^{L1}	γ_{00}	5.5395	0.0131	2997	421.29*	
		γ_{01}	-0.0837	0.0291	2997	**-2.87***	0.0010
Leader level ^{L2}	Others' communal ratings ^{L1}	γ_{00}	5.5198	0.0105	2305	526.05*	
		γ_{01}	-0.0276	0.0210	2305	**-1.31**	0.0002
	Others' agentic ratings ^{L1}	γ_{00}	5.5199	0.0107	2305	517.53*	
		γ_{01}	0.0474	0.0214	2305	**2.22***	0.0005
	Others' effectiveness ratings ^{L1}	γ_{00}	5.5819	0.0158	2076	352.80*	
		γ_{01}	0.0443	0.0316	2076	**1.40**	0.0001

[a] = Leader gender is coded 1 (female) and 2 (male). Leader level is coded 1 (middle management) and 2 (executive). All variables were grand mean centered.

[b] = γ_{00} is the intercept variance coefficient and γ_{01} is the slope variance coefficient.

[c] = $p<.05$. Values in bold are the slope significance tests.

Table 16. Summary of Analysis of Variance Descriptive Main Effects of Leader Gender, Leader Level, and Rater Gender

Dependent Variable	Source	SS	df	MS	F[a]	R^2
Self communal ratings	Leader gender	2.21	1	2.21	7.70*	
	Error	1056.13	3685	0.29		
	Total	124820.32	3687			0.002
Self agentic ratings	Leader gender	0.65	1	0.65	1.53	
	Error	1569.03	3685	0.43		
	Total	119187.40	3687			0.000
Self communal ratings	Leader level (middle vs. executive)	3.17	1	3.17	11.00*	
	Error	755.79	2621	0.29		
	Total	89571.40	2623			0.004
Self agentic ratings	Leader level (middle vs. executive)	16.95	1	16.95	42.19*	
	Error	1052.73	2621	0.40		
	Total	86734.83	2623			0.016
Others' communal ratings	Rater gender	56.37	1	56.37	82.78*	
	Error	14014.19	20581	0.68		
	Total	653111.14	20583			0.004
Others' agentic ratings	Rater gender	67.96	1	67.96	93.43*	
	Error	14969.42	20581	0.73		
	Total	648072.97	20583			0.005
Others' effectiveness ratings	Rater gender	137.95	1	137.95	107.62*	
	Error	26379.60	20581	1.28		
	Total	675126.10	20583			0.005

[a] = $p<.05$

Table 17. Results of the Null Model: Testing for Between Group Variance

H	Independent Variables	Moderators	Dependent Variables	Level 1 and Level 2 Variance Components				Level 3 Variance Components			
				u_0/ r_0 [a]	df	χ^{2} [b]	ICC	u_{00}	df	χ^{2} [b]	ICC
1	Leader gender L2	Industry L3	Others' development ratings L1	0.1455	2525	4915.99*	0.1944	0.0853	50	955.52*	0.1140
			Others' effectiveness ratings L1	0.3263	2219	5168.86*	0.2317	0.1205	48	583.61*	0.0856
2	Leader gender L2	Functional area L3	Others' development ratings L1	0.1661	2744	9951.31*	0.2407	0.0029	9	50.33*	0.0042
			Others' effectiveness ratings L1	0.3730	2472	10033.06*	0.2770	0.0048	9	36.60*	0.0036
3	Leader gender L2	Proportion of male raters L2	Others' development ratings L1	0.1745	3312	12666.61*	0.2489	n/a	n/a	n/a	n/a
			Others' effectiveness ratings L1	0.3855	2998	125262.48*	0.2825	n/a	n/a	n/a	n/a
4	Others' development ratings L1	Leader gender L2	Others' effectiveness ratings L1	0.3857	2998	12548.80*	0.2825	n/a	n/a	n/a	n/a
5a	Leader gender L2	Industry L3	Others' communal ratings L1	0.1545	2525	4847.36*	0.1951	0.0888	50	875.53*	0.1121
5b	Leader gender L2	Functional area L3	Others' communal ratings L1	0.1751	2744	9856.73*	0.2396	0.0037	9	55.13*	0.0051
5c	Leader gender L2	Proportion of male raters L2	Others' communal ratings L1	0.1835	3312	12633.42*	0.2477	n/a	n/a	n/a	n/a
6a	Leader gender L2	Industry L3	Male bosses' communal ratings L1	0.1680	1465	2192.86*	0.3072	0.0702	43	247.66*	0.1284
			Male bosses' agentic ratings L1	0.1346	1465	1814.43*	0.2007	0.1082	43	270.57*	0.1613
			Male bosses' effectiveness ratings L1	0.3055	1252	1675.02*	0.2787	0.1965	43	232.00*	0.1793
6b	Leader gender L2	Functional area L3	Male bosses' communal ratings L1	0.2111	1645	3077.00*	0.4180	0.0026	9	19.07*	0.0051
			Male bosses' agentic ratings L1	0.5479	1644	81372.93*	0.9062	0.0062	9	27.31*	0.0103
			Male bosses' effectiveness ratings L1	0.3380	1447	2408.54*	0.3527	0.0123	9	33.37*	0.0128

[a] = For the 2-level and 3-level models, the symbols for the between group variance components are u_0 and r_0, respectively.

[b] = This is the chi-square significance test for each between group variance component; $p<.05$.

Table 17 (continued). Results of the Null Model: Testing for Between Group Variance

H	Independent Variables	Moderators	Dependent Variables	Level 1 and Level 2 Variance Components				Level 3 Variance Components			
				u_0/ r_0[a]	df	χ^2[b]	ICC	u_{00}	df	χ^2[b]	ICC
6c	Leader gender[L2]	Proportion of male raters[L2]	Male bosses' communal ratings[L1]	0.1929	1907	3327.69*	0.3838	n/a	n/a	n/a	n/a
			Male bosses' agentic ratings[L1]	0.1592	1905	2730.95*	0.2667	n/a	n/a	n/a	n/a
			Male bosses' effectiveness ratings[L1]	0.3426	1690	2822.17*	0.3587	n/a	n/a	n/a	n/a
7a	Leader gender[L2]	Industry[L3]	Female bosses' communal ratings[L1]	0.0001	290	297.91	0.0001	0.0846	31	68.89*	0.1253
			Female bosses' agentic ratings[L1]	0.0994	290	362.85*	0.1322	0.0728	31	65.79*	0.0968
			Female bosses' effectiveness ratings[L1]	0.4471	235	394.38*	0.3389	0.1973	30	73.12*	0.1495
7b	Leader gender[L2]	Functional area[L3]	Female bosses' communal ratings[L1]	0.0006	366	366.40	0.0010	0.0005	9	12.74	0.0008
			Female bosses' agentic ratings[L1]	0.0546	366	404.33	0.0821	0.0133	9	19.03*	0.0200
			Female bosses' effectiveness ratings[L1]	0.3438	316	463.06*	0.2985	0.0638	9	27.60*	0.0554
7c	Leader gender[L2]	Proportion of male raters[L2]	Female bosses' communal ratings[L1]	0.0550	440	474.95	0.0931	n/a	n/a	n/a	n/a
			Female bosses' agentic ratings[L1]	0.0795	440	493.49*	0.1165	n/a	n/a	n/a	n/a
			Female bosses' effectiveness ratings[L1]	0.5022	384	701.03*	0.4411	n/a	n/a	n/a	n/a
8a	Leader gender[L2]	Industry[L3]	Female peers' communal ratings[L1]	0.1463	1141	1680.73*	0.2128	0.0488	42	160.97*	0.0710
			Female peers' agentic ratings[L1]	0.1611	1141	1713.48*	0.2229	0.0669	42	200.42*	0.0925
			Female peers' effectiveness ratings[L1]	0.3129	959	1478.57*	0.2479	0.1070	42	149.75*	0.0848
8b	Leader gender[L2]	Functional area[L3]	Female peers' communal ratings[L1]	0.2027	1353	2255.22*	0.2812	0.0104	9	31.79*	0.0144
			Female peers' agentic ratings[L1]	0.1983	1353	2222.14*	0.2703	0.0047	9	21.67*	0.0064
			Female peers' effectiveness ratings[L1]	0.4055	1179	2070.37*	0.3081	0.0082	9	19.58*	0.0062

Table 17 (continued). Results of the Null Model: Testing for Between Group Variance

H	Independent Variables	Moderators	Dependent Variables	Level 1 and Level 2 Variance Components				Level 3 Variance Components			
				u_0/ r_0[a]	df	χ^2[b]	ICC	u_{00}	df	χ^2[b]	ICC
8c	Leader gender[L2]	Proportion of male raters[L2]	Female peers' communal ratings[L1]	0.1990	1559	2621.67*	0.2828	n/a	n/a	n/a	n/a
			Female peers' agentic ratings[L1]	0.2046	1559	2619.83*	0.2843	n/a	n/a	n/a	n/a
			Female peers' effectiveness ratings[L1]	0.3996	1368	2408.13*	0.3104	n/a	n/a	n/a	n/a
9	Bosses' and subordinates' communal ratings[L1] Bosses' and subordinates' agentic ratings[L1]	Rater type[L1] Leader gender[L2]	Bosses' and subordinates' effectiveness ratings[L1]	0.4038	2890	8464.57*	0.2992	n/a	n/a	n/a	n/a
10a	Leader gender[L1]	Industry[L2]	Self communal ratings[L1]	0.0347	50	380.18*	0.1149	n/a	n/a	n/a	n/a
			Self agentic ratings[L1]	0.0462	50	338.14*	0.1054	n/a	n/a	n/a	n/a
10b	Leader gender[L1]	Functional area[L2]	Self communal ratings[L1]	0.0059	9	74.74*	0.0204	n/a	n/a	n/a	n/a
			Self agentic ratings[L1]	0.0134	9	78.64*	0.0306	n/a	n/a	n/a	n/a
12	Self agentic ratings[L2]	Leader gender[L2]	Others' effectiveness ratings[L1]	0.3913	2859	12177.72*	0.2859	n/a	n/a	n/a	n/a
14	Leader gender[L2]	Leader level[L2]	Others' communal ratings[L1]	0.1825	2266	8693.08*	0.2461	n/a	n/a	n/a	n/a
			Others' agentic ratings[L1]	0.1900	2266	8593.40*	0.2433	n/a	n/a	n/a	n/a
			Others' effectiveness ratings[L1]	0.3867	2040	8782.19*	0.2843	n/a	n/a	n/a	n/a

Table 18. Summary of Results for 3-Level, 2-Way Interaction HLM Analyses: Hypotheses 1, 2, 5a-b, 6a-b, 7a-b, and 8a-b

H	Independent Variables[a]	Moderators	Dependent Variables	γ	Level 2 Parameter Estimates				Level 3 Parameter Estimates			
					Coef	SE	df	T[b]	Coef	SE	df	T[b]
1	Leader gender^{L2}	Industry^{L3}	Others' development ratings^{L1}	$\gamma_{000}/\gamma_{010}$	5.3129	0.0481	47	110.46*	-0.0457	0.0246	47	-1.86
				$\gamma_{001}/\gamma_{011}$	0.0007	0.0021	47	0.34	-0.0030	0.0009	47	**-3.02***
			Others' effectiveness ratings^{L1}	$\gamma_{000}/\gamma_{010}$	5.3444	0.0589	47	90.72*	0.0258	0.0319	47	0.81
				$\gamma_{001}/\gamma_{011}$	0.0016	0.0026	47	0.64	-0.0040	0.0010	47	**-3.88***
2	Leader gender^{L2}	Functional area^{L3}	Others' development ratings^{L1}	$\gamma_{000}/\gamma_{010}$	5.4853	0.0170	8	322.62*	-0.0745	0.0269	8	-2.77*
				$\gamma_{001}/\gamma_{011}$	0.0015	0.0007	8	2.19	0.0006	0.0015	8	**0.37**
			Others' effectiveness ratings^{L1}	$\gamma_{000}/\gamma_{010}$	5.5141	0.0268	8	205.57*	-0.0592	0.0356	8	-1.66
				$\gamma_{001}/\gamma_{011}$	-0.0000	0.0015	8	-0.02	-0.0014	0.0026	8	**-0.56**
5a	Leader gender^{L2}	Industry^{L3}	Others' communal ratings^{L1}	$\gamma_{000}/\gamma_{010}$	5.3209	0.0482	49	110.28*	-0.1163	0.0240	49	-4.85*
				$\gamma_{001}/\gamma_{011}$	0.0004	0.0021	49	0.19	-0.0021	0.0008	49	**-2.47***
5b	Leader gender^{L2}	Functional area^{L3}	Others' communal ratings^{L1}	$\gamma_{000}/\gamma_{010}$	5.4999	0.0185	8	296.73*	-0.0947	0.0316	8	-3.00*
				$\gamma_{001}/\gamma_{011}$	0.0015	0.0010	8	1.41	0.0008	0.0020	8	**0.39**
6a	Leader gender^{L2}	Industry^{L3}	Male bosses' communal ratings^{L1}	$\gamma_{000}/\gamma_{010}$	5.4967	0.0539	42	102.06*	-0.0790	0.0445	42	-1.78
				$\gamma_{001}/\gamma_{011}$	-0.0003	0.0023	42	-0.14	-0.0012	0.0017	42	**-0.73**
			Male bosses' agentic ratings^{L1}	$\gamma_{000}/\gamma_{010}$	5.4413	0.0628	42	86.60*	-0.0526	0.0491	42	-1.07
				$\gamma_{001}/\gamma_{011}$	-0.0019	0.0025	42	-0.79	-0.0034	0.0024	42	**-1.44**
			Male bosses' effectiveness ratings^{L1}	$\gamma_{000}/\gamma_{010}$	5.4737	0.0851	42	64.34*	0.0176	0.0690	42	0.26
				$\gamma_{001}/\gamma_{011}$	-0.0015	0.0032	42	-0.46	-0.0043	0.0027	42	**-1.60**
6b	Leader gender^{L2}	Functional area^{L3}	Male bosses' communal ratings^{L1}	$\gamma_{000}/\gamma_{010}$	5.6065	0.0237	8	236.63*	-0.0416	0.0427	8	-0.97
				$\gamma_{001}/\gamma_{011}$	0.0015	0.0014	8	1.10	0.0013	0.0029	8	**0.45**
			Male bosses' agentic ratings^{L1}	$\gamma_{000}/\gamma_{010}$	5.5481	0.0318	8	174.38*	-0.0376	0.0478	8	-0.79
				$\gamma_{001}/\gamma_{011}$	0.0008	0.0018	8	0.47	-0.0007	0.0032	8	**-0.23**
			Male bosses' effectiveness ratings^{L1}	$\gamma_{000}/\gamma_{010}$	5.6117	0.0423	8	132.65*	-0.0124	0.0670	8	-0.19
				$\gamma_{001}/\gamma_{011}$	-0.0025	0.0024	8	-1.04	-0.0042	0.0044	8	**-0.94**

[a] = Leader gender is coded 1 (female) and 2 (male). All variables were grand mean centered.
[b] = $p<.05$. Values in bold are the level-3 moderation significance tests.

Table 18 (continued). Summary of Results for 3-Level, 2-Way Interaction HLM Analyses: Hypotheses 1, 2, 5a-b, 6a-b, 7a-b, and 8a-b

H	Independent Variables[a]	Moderators	Dependent Variables	γ^b	Level 2 Parameter Estimates				Level 3 Parameter Estimates			
					Coef	SE	df	T[b]	Coef	SE	df	T[b]
7a	Leader gender[L2]	Industry[L3]	Female bosses' communal ratings[L1]	$\gamma_{000}/\gamma_{010}$	5.5810	0.0880	30	63.44*	-0.1152	0.1222	30	-0.94
				$\gamma_{001}/\gamma_{011}$	0.0008	0.0040	30	0.20	-0.0020	0.0036	30	**-0.56**
			Female bosses' agentic ratings[L1]	$\gamma_{000}/\gamma_{010}$	5.4635	0.0893	30	61.16*	-0.0875	0.1010	30	-0.87
				$\gamma_{001}/\gamma_{011}$	0.0015	0.0043	30	0.34	-0.0034	0.0043	30	**-0.78**
			Female bosses' effectiveness ratings[L1]	$\gamma_{000}/\gamma_{010}$	5.5136	0.1322	29	41.70*	-0.0606	0.1642	29	-0.37
				$\gamma_{001}/\gamma_{011}$	0.0042	0.0061	29	0.68	-0.0014	0.0051	29	**-0.28**
7b	Leader gender[L2]	Functional area[L3]	Female bosses' communal ratings[L1]	$\gamma_{000}/\gamma_{010}$	5.6384	0.0424	8	133.13*	-0.0463	0.0918	8	-0.50
				$\gamma_{001}/\gamma_{011}$	-0.0001	0.0032	8	-0.04	-0.0028	0.0102	8	**-0.27**
			Female bosses' agentic ratings[L1]	$\gamma_{000}/\gamma_{010}$	5.5320	0.0630	8	87.84*	-0.0254	0.1077	8	-0.24
				$\gamma_{001}/\gamma_{011}$	0.0002	0.0039	8	0.05	0.0031	0.0072	8	**0.43**
			Female bosses' effectiveness ratings[L1]	$\gamma_{000}/\gamma_{010}$	5.5611	0.1040	8	53.48*	0.0189	0.1525	8	0.12
				$\gamma_{001}/\gamma_{011}$	-0.0003	0.0061	8	-0.05	0.0110	0.0100	8	**1.09**
8a	Leader gender[L2]	Industry[L3]	Female peers' communal ratings[L1]	$\gamma_{000}/\gamma_{010}$	5.5095	0.0482	41	114.35*	-0.1687	0.0459	41	-3.68*
				$\gamma_{001}/\gamma_{011}$	-0.0001	0.0022	41	-0.03	-0.0023	0.0016	41	**-1.38**
			Female peers' agentic ratings[L1]	$\gamma_{000}/\gamma_{010}$	5.4480	0.0535	41	101.83*	-0.0933	0.0485	41	-1.92
				$\gamma_{001}/\gamma_{011}$	0.0004	0.0027	41	0.16	-0.0035	0.0017	41	**-2.07***
			Female peers' effectiveness ratings[L1]	$\gamma_{000}/\gamma_{010}$	5.4349	0.0711	41	76.42*	-0.0553	0.0716	41	-0.77
				$\gamma_{001}/\gamma_{011}$	0.0001	0.0029	41	0.05	-0.0027	0.0026	41	**-1.04**
8b	Leader gender[L2]	Functional area[L3]	Female peers' communal ratings[L1]	$\gamma_{000}/\gamma_{010}$	5.5422	0.0373	8	148.51*	-0.1590	0.0551	8	-2.88*
				$\gamma_{001}/\gamma_{011}$	-0.0008	0.0022	8	-0.35	-0.0019	0.0038	8	**-0.49**
			Female peers' agentic ratings[L1]	$\gamma_{000}/\gamma_{010}$	5.5294	0.0303	8	182.50*	-0.1636	0.0435	8	-3.76*
				$\gamma_{001}/\gamma_{011}$	-0.0021	0.0019	8	-1.13	0.0012	0.0033	8	**0.36**
			Female peers' effectiveness ratings[L1]	$\gamma_{000}/\gamma_{010}$	5.4969	0.0411	8	133.82*	-0.1651	0.0682	8	-2.42*
				$\gamma_{001}/\gamma_{011}$	-0.0032	0.0026	8	-1.25	-0.0030	0.0049	8	**-0.62**

Table 19. Summary of Results for 2-Level, 2-Way Interaction HLM Analyses: Hypotheses 3, 5c, 6c, 7c, 8c, 12, and 14

H	Independent Variables[a]	Moderators	Dependent Variables	γ	Level 2 Parameter Estimates			
					Coef	SE	df	T[b]
3	Leader gender ^{L2}	Proportion of male raters ^{L2}	Others' development ratings ^{L1}	γ_{00}	5.4986	0.0085	3309	647.34*
				γ_{01}	-0.1549	0.0414	3309	-3.75*
				γ_{02}	-0.0018	0.0013	3309	-1.40
				γ_{03}	0.0011	0.0007	3309	**1.51**
			Others' effectiveness ratings ^{L1}	γ_{00}	5.5395	0.0131	2995	421.44*
				γ_{01}	-0.1879	0.0638	2995	-2.95*
				γ_{02}	-0.0030	0.0020	2995	-1.53
				γ_{03}	0.0020	0.0011	2995	**1.78**
5c	Leader gender ^{L2}	Proportion of male raters ^{L2}	Others' communal ratings ^{L1}	γ_{00}	5.5103	0.0087	3309	632.82*
				γ_{01}	-0.1565	0.0424	3309	-3.69*
				γ_{02}	-0.0014	0.0013	3309	-1.03
				γ_{03}	0.0008	0.0008	3309	**1.05**
6c	Leader gender ^{L2}	Proportion of male raters ^{L2}	Male bosses' communal ratings ^{L1}	γ_{00}	5.6043	0.0156	1904	359.70*
				γ_{01}	-0.0836	0.0988	1904	-0.85
				γ_{02}	-0.0022	0.0028	1904	-0.78
				γ_{03}	0.0011	0.0016	1904	**0.67**
			Male bosses' agentic ratings ^{L1}	γ_{00}	5.5642	0.0168	1902	331.19*
				γ_{01}	-0.0962	0.1137	1902	-0.85
				γ_{02}	-0.0017	0.0034	1902	-0.51
				γ_{03}	0.0013	0.0019	1902	**0.72**
			Male bosses' effectiveness ratings ^{L1}	γ_{00}	5.6312	0.0228	1687	247.44*
				γ_{01}	-0.1631	0.1496	1687	-1.09
				γ_{02}	-0.0030	0.0043	1687	-0.68
				γ_{03}	0.0026	0.0024	1687	**1.06**

[a] = Leader gender is coded 1 (female) and 2 (male). All variables were grand mean centered.
[b] = $p<.05$. Values in bold are the level-2 moderation significance tests.

Table 19 (continued). Summary of Results for 2-Level, 2-Way Interaction HLM Analyses: Hypotheses 3, 5c, 6c, 7c, 8c, 12, and 14

H	Independent Variables[a]	Moderators	Dependent Variables	γ	Level 2 Parameter Estimates			
					Coef	SE	df	T[b]
7c	Leader gender[L2]	Proportion of male raters[L2]	Female bosses' communal ratings[L1]	γ_{00}	5.6585	0.0357	437	158.41*
				γ_{01}	-0.1077	0.1361	437	-0.79
				γ_{02}	-0.0018	0.0055	437	-0.33
				γ_{03}	0.0004	0.0033	437	**0.13**
			Female bosses' agentic ratings[L1]	γ_{00}	5.5589	0.0384	437	144.68*
				γ_{01}	-0.0787	0.1472	437	-0.54
				γ_{02}	-0.0024	0.0056	437	-0.44
				γ_{03}	0.0000	0.0034	437	**0.01**
			Female bosses' effectiveness ratings[L1]	γ_{00}	5.5589	0.0384	437	144.68*
				γ_{01}	-0.0787	0.1472	437	-0.54
				γ_{02}	-0.0024	0.0056	437	-0.44
				γ_{03}	0.0000	0.0034	437	**0.01**
8c	Leader gender[L2]	Proportion of male raters[L2]	Female peers' communal ratings[L1]	γ_{00}	5.5719	0.0185	1556	301.05*
				γ_{01}	-0.2854	0.0813	1556	-3.51*
				γ_{02}	-0.0054	0.0028	1556	-1.93*
				γ_{03}	0.0031	0.0018	1556	**1.75**
			Female peers' agentic ratings[L1]	γ_{00}	5.5510	0.0189	1556	293.29*
				γ_{01}	-0.2718	0.0814	1556	-3.34*
				γ_{02}	-0.0043	0.0029	1556	-1.49
				γ_{03}	0.0031	0.0018	1556	**1.71**
			Female peers' effectiveness ratings[L1]	γ_{00}	5.5361	0.0269	1365	205.80*
				γ_{01}	-0.3710	0.1262	1365	-2.94*
				γ_{02}	-0.0072	0.0042	1365	-1.72
				γ_{03}	0.0052	0.0027	1365	**1.96***
12	Self agentic ratings[L2]	Leader gender[L2]	Others' effectiveness ratings[L1]	γ_{00}	5.5351	0.0133	2856	416.59*
				γ_{01}	0.1449	0.0812	2856	1.79
				γ_{02}	-0.3016	0.2663	2856	-1.13
				γ_{03}	0.0379	0.0464	2856	**0.82**

Table 19 (continued). Summary of Results for 2-Level, 2-Way Interaction HLM Analyses: Hypotheses 3, 5c, 6c, 7c, 8c, 12, and 14

H	Independent Variables[a]	Moderators	Dependent Variables	γ	Level 2 Parameter Estimates			
					Coef	SE	df	T[b]
14	Leader gender [L2]	Leader level [L2]	Others' communal ratings [L1]	γ_{00}	5.5220	0.0104	2263	529.07*
				γ_{01}	-0.1432	0.0715	2263	-2.002*
				γ_{02}	-0.0159	0.0841	2263	-0.19
				γ_{03}	-0.0017	0.0471	2263	**-0.04**
			Others' agentic ratings [L1]	γ_{00}	5.5191	0.0107	2263	515.26*
				γ_{01}	-0.1585	0.0723	2263	-2.19*
				γ_{02}	-0.0117	0.0855	2263	-0.14
				γ_{03}	0.0365	0.0480	2263	**0.76**
			Others' effectiveness ratings [L1]	γ_{00}	5.5836	0.0158	2037	352.97*
				γ_{01}	-0.2081	0.1087	2037	-1.92*
				γ_{02}	-0.0466	0.1305	2037	-0.36
				γ_{03}	0.0532	0.0726	2037	**0.73**

Table 20. Summary of Results for 2-Level, 2-Way Interaction HLM Analyses: Hypotheses 4 and 10a-10b

H	Independent Variables[a]	Moderators	Dependent Variables	γ	Level 1 Parameter Estimates				Level 2 Parameter Estimates			
					Coef	SE	df	T[b]	Coef	SE	df	T[b]
4	Others' development ratings[L1]	Leader gender[L2]	Others' effectiveness ratings[L1]	γ_{00}/γ_{10}	5.5633	0.0059	2997	945.73*	1.1333	0.0062	2997	182.94*
				γ_{01}/γ_{11}	0.0285	0.0130	2997	2.189*	0.0104	0.0137	2997	**0.76**
10a	Leader gender[L1]	Industry[L2]	Self communal ratings[L1]	γ_{00}/γ_{10}	5.7301	0.0323	49	177.67*	-0.0816	0.0276	49	-2.96*
				γ_{01}/γ_{11}	0.0023	0.0016	49	1.41	-0.0021	0.0012	49	**-1.83**
			Self agentic ratings[L1]	γ_{00}/γ_{10}	5.5856	0.0380	49	146.97*	0.0256	0.0292	49	0.88
				γ_{01}/γ_{11}	0.0016	0.0020	49	0.81	-0.0038	0.0013	49	**-2.96***
10b	Leader gender[L1]	Functional area[L2]	Self communal ratings[L1]	γ_{00}/γ_{10}	5.7680	0.0279	8	206.93*	-0.0145	0.0303	8	-0.49
				γ_{01}/γ_{11}	0.0004	0.0015	8	0.26	-0.0015	0.0020	8	-0.72
			Self agentic ratings[L1]	γ_{00}/γ_{10}	5.6170	0.0380	8	147.64*	0.0571	0.0399	8	1.43
				γ_{01}/γ_{11}	-0.0007	0.0021	8	-0.35	-0.0005	0.0026	8	**-0.19**

[a] = Leader gender is coded 1 (female) and 2 (male). All variables were grand mean centered.
[b] = $p<.05$. Values in bold are the level-2 moderation significance tests.

Table 21. Summary of Results for 2-Level, 3-Way Interaction HLM Analysis: Hypothesis 9

H	Independent Variables[a]	Moderators	Dependent Variables	γ	Level 2 Parameter Estimates			
					Coef	SE	df	T[b]
9	Bosses' and subordinates' communal ratings [L1] Bosses' and subordinates' agentic ratings [L1]	Rater type [L1] Leader gender [L2]	Bosses' and subordinates' effectiveness ratings [L1]	γ_{00}	5.6967	0.0068	2889	841.66*
				γ_{01}	0.0237	0.0156	2889	1.52
				γ_{10}	-0.2139	0.1032	2889	-2.07*
				γ_{11}	0.2779	0.2275	2889	1.22
				γ_{20}	0.6810	0.0417	2889	16.33*
				γ_{21}	0.0795	0.0941	2889	0.85
				γ_{30}	0.3953	0.0383	2889	10.33*
				γ_{31}	-0.0354	0.0864	2889	-0.41
				γ_{40}	-0.0761	0.0315	2889	-2.41*
				γ_{41}	-0.0961	0.0718	2889	**-1.34**
				γ_{50}	0.0887	0.0283	2889	3.13*
				γ_{51}	0.0530	0.0646	2889	**0.82**

[a] = Leader gender is coded 1 (female) and 2 (male). All variables were grand mean centered.
[b] = $p<.05$. Values in bold are the level-2, 3-way moderation significance tests.

Table 22. Analysis of Variance Results for Hypotheses 10c, 11, and 13

H	Dependent Variable	Source	SS	df	MS	F[a]	R^2
10c	Self communal ratings	Leader gender	3.06	1	3.06	10.70*	
		% of male raters	1.46	2	0.73	2.55	
		Leader Gender * % of male raters	1.51	2	0.76	**2.65**	
		Error	1052.64	3681	0.29		
		Total	124820.32	3687			0.005
	Self agentic ratings	Leader gender	0.04	1	0.04	0.10	
		% of male raters	2.38	2	1.19	2.80	
		Leader gender * % of male raters	1.00	2	0.50	**1.17**	
		Error	1564.53	3681	0.43		
		Total	119187.40	3687			0.003
11	Self communal ratings	Leader gender	2.21	1	2.21	**7.70***	
		Error	1056.13	3685	0.29		
		Total	124820.32	3687			0.002
13	Self-subordinate rater agreement on communal ratings	Leader gender	0.12	1	0.12	**0.12**	
		Error	768.57	2837	0.27		
		Total	1808.70	2839			0.000
	Self-subordinate rater agreement on agentic ratings	Leader gender	1.81	1	1.81	**6.52***	
		Error	785.57	2837	0.28		
		Total	1949.83	2839			0.002

[a] = $p<.05$. Values in bold are the tests of significance for the relationships of interest.

Figure 1. Denison Leadership Development Model (Denison & Neale, 2001)

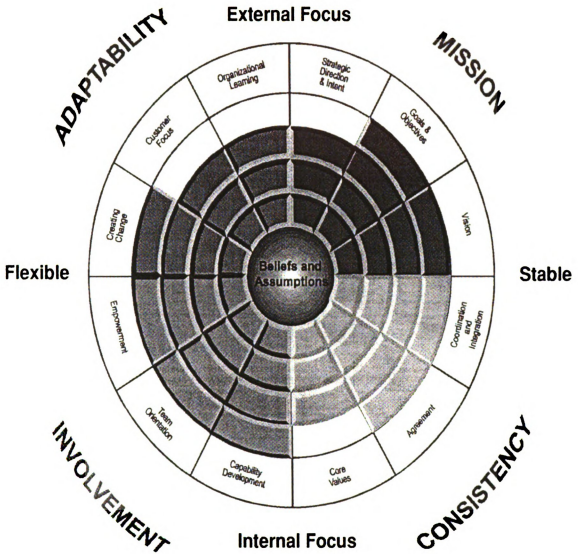
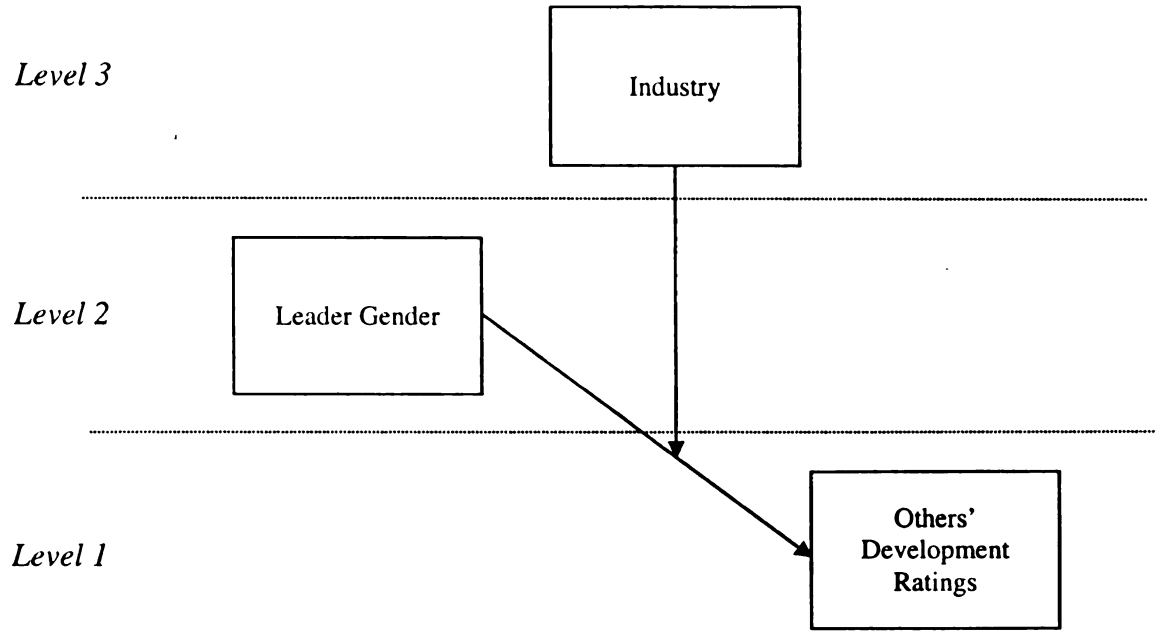


Figure 2. Example 3-Level, 2-Way Interaction HLM Model for Hypotheses 1, 2, 5a-b, 6a-b, 7a-b, and 8a-b (Taken from Hypothesis 1)



Level-1 Model
Others' Development Ratings = $\pi_0 + e$

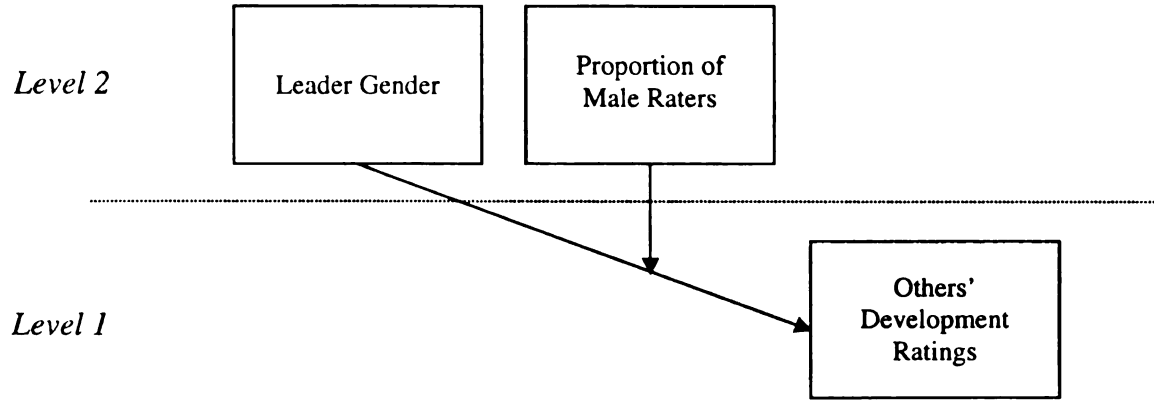
Level-2 Model
$\pi_0 = \beta_{00} + \beta_{01}$ (Leader Gender) + r_0

Level-3 Model
$\beta_{00} = \gamma_{000} + \gamma_{001}$ (% Women in Industry) + u_{00}
$\beta_{01} = \gamma_{010} + \gamma_{011}$ (% Women in Industry) + u_{01}

Where:

Others' Development Ratings = example of level-1 outcome variable
Leader Gender = example of level-2 predictor variable
% Women in Industry = example of level-3 predictor variable
π_0 = mean ratings for Leader Gender within Industry
β_{00} = level-2 intercept
β_{01} = level-2 slope
γ_{000} = grand mean
γ_{001} = level-3 slope
γ_{010} = level-3 intercept
γ_{011} = level-3 slope (cross level 2-way interaction: test of moderation hypotheses)
e = within group variance
r_0 = between Leader Gender variance
u_{00}, u_{01} = between Industry variance

Figure 3. Example 2-Level, 2-Way Interaction HLM Model for Hypotheses 3, 5c, 6c, 7c, 8c, 12, and 14 (Taken from Hypothesis 3)



Level-1 Model

Others' Development Ratings = $\beta_0 + r$

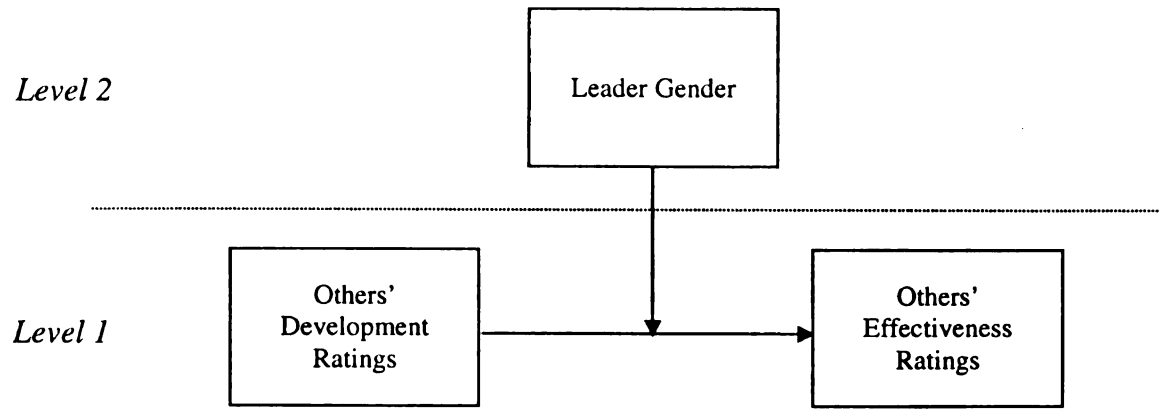
Level-2 Model

$\beta_0 = \gamma_{00} + \gamma_{01}$ (Leader Gender) + γ_{02} (Proportion of Male Raters) + γ_{03} (Leader Gender*Proportion of Male Raters) + u_0

Where:

Others' Development Ratings = example of level-1 outcome variable
Leader Gender = example of level-2 predictor variable
Proportion of Male Raters = example of level-2 predictor variable
Leader Gender*Proportion of Male Raters = example of level-2 interaction term
β_0 = mean ratings for Leader Gender
γ_{00} = grand mean
γ_{01}, γ_{02} = level-2 slope (cross level main effects of level-2 predictors on the level-1 outcome variable)
γ_{03} = level-2 slope (single level 2-way interaction on level-1 outcome variable: test of moderation hypotheses)
r = variance within Leader Gender
u_0 = variance between Leader Gender

Figure 4. Example 2-Level, 2-Way Interaction HLM Model for Hypotheses 4 and 10a-b (Taken from Hypothesis 4)



Level-1 Model

Others' Effectiveness Ratings = $\beta_0 + \beta_1$ (Others' Development Ratings) + r

Level-2 Model

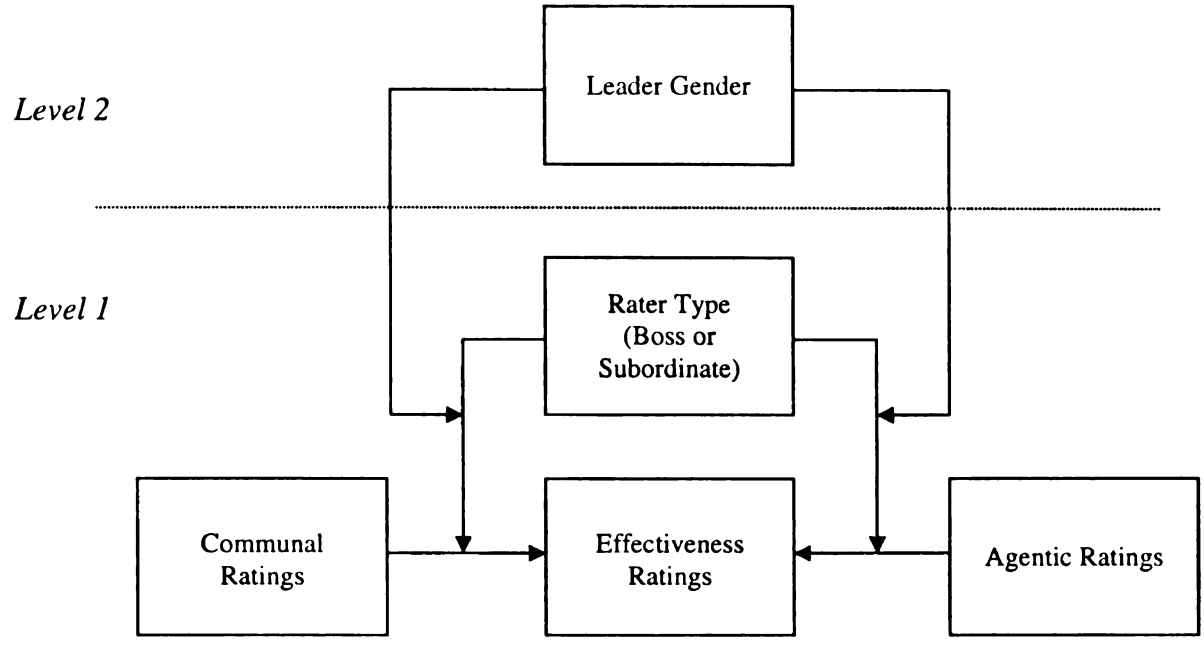
$\beta_0 = \gamma_{00} + \gamma_{01}$ (Leader Gender) + u_0

$\beta_1 = \gamma_{10} + \gamma_{11}$ (Leader Gender) + u_1

Where:

Others' Effectiveness Ratings = example of level-1 outcome variable
Others' Development Ratings = example of level-1 predictor
Leader Gender (in β_0 equation) = example of level-2 predictor of level-1 intercept
Leader Gender (in β_1 equation) = example of level-2 predictor of level-1 slopes
β_0 = mean ratings for Leader Gender
β_1 = level-1 intercept (main effect of level-1 predictor on outcome variable)
γ_{00} =grand mean
γ_{01} = level-2 slope
γ_{10} = level-2 intercept
γ_{11} = level-2 slope (cross level 2-way interaction: test of moderation hypotheses)
r = variance within Leader Gender
u_0, u_1 = variance between Leader Gender

Figure 5. 2-Level, 3-Way Interaction HLM Model for Hypotheses 9



Level-1 Model

Effectiveness Ratings = $\beta_0 + \beta_1$ (Rater Type) + β_2 (Communal Ratings) + β_3 (Agentic Ratings) + β_4 (Rater Type*Communal Ratings) + β_5 (Rater Type*Agentic Ratings) + r

Level-2 Model

$\beta_0 = \gamma_{00} + \gamma_{01}$ (Leader Gender) + u_0
$\beta_1 = \gamma_{10} + \gamma_{11}$ (Leader Gender) + u_1
$\beta_2 = \gamma_{20} + \gamma_{21}$ (Leader Gender) + u_2
$\beta_3 = \gamma_{30} + \gamma_{31}$ (Leader Gender) + u_3
$\beta_4 = \gamma_{40} + \gamma_{41}$ (Leader Gender) + u_4
$\beta_5 = \gamma_{50} + \gamma_{51}$ (Leader Gender) + u_5

Where:

Effectiveness Ratings = level-1 outcome variable
Rater Type, Communal Ratings, Agentic Ratings = level-1 predictors
Rater Type*Communal Ratings, Rater Type*Agentic Ratings = level-1 interactions
Leader Gender = level-2 predictor variable
β_0 = mean ratings for Leader Gender
$\beta_1, \beta_2, \beta_3$ = level-1 intercepts (main effects of level-1 predictors on outcome variable)
β_4, β_5 = level-1 intercepts (single level 2-way interactions)
γ_{00} = grand mean
$\gamma_{10}, \gamma_{20}, \gamma_{30}, \gamma_{40}, \gamma_{50}$ = level-2 intercepts
$\gamma_{01}, \gamma_{11}, \gamma_{21}, \gamma_{31}$ = level-2 slopes (cross level 2-way interactions)
γ_{41}, γ_{51} = level-2 slopes (cross level 3-way interactions: test of moderation hypotheses)
r = variance within Leader Gender
$u_0, u_1, u_2, u_3, u_4, u_5$ = variance between Leader Gender

Figure 6. Example 2 and 3-Level HLM Null Models

2-Level HLM Null Model

Level-1 Model

Others' Communal Ratings = $\beta_0 + r$

Level-2 Model

$\beta_0 = \gamma_{00} + u_0$

Where:

Others' Communal Ratings = example of level-1 outcome variable
β_0 = mean ratings for Leader Gender
γ_{00} = grand mean
r = variance within Leader Gender
u_0 = variance between Leader Gender

3-Level HLM Null Model

Level-1 Model

Others' Communal Ratings = $\pi_0 + e$

Level-2 Model

$\pi_0 = \beta_{00} + r_0$

Level-3 Model

$\beta_{00} = \gamma_{000} + u_{00}$

Where:

Others' Communal Ratings = example of level-1 outcome variable
π_0 = mean ratings for Leader Gender within Industry
β_{00} = mean ratings for Industry
γ_{000} = grand mean
e = variance within groups
r_0 = variance between Leader Gender within Industry
u_{00} = variance between Industry

Appendix A. Denison Leadership Development Survey Demographics Questions (Denison & Neale, 1996)

Sex
1. Female
2. Male
3. Prefer not to respond

Age
1. Under 20
2. 20-29
3. 30-39
4. 40-49
5. 50-59
6. 60 or over
7. Prefer not to respond

Education (mark highest level)
1. High school
2. Some college
3. Associate's/Technical degree
4. Bachelor's degree
5. Some graduate work
6. Master's degree
7. Doctoral degree
8. Other:_____
9. Prefer not to respond

Years With Organization
1. Less than 6 months
2. 6 months to 1 year
3. 1 to 2 years
4. 2 to 4 years
5. 4 to 6 years
6. 6 to 10 years
7. 10 to 15 years
8. More than 15 years
9. Prefer not to respond

Salary (Annual)
1. $25,000 or less
2. 25,001 to 35,000
3. 35,001 to 50,000
4. 50,001 to 75,000
5. 75,001 to 100,000
6. 100,001 to 150,000
7. 150,001 to 200,000
8. 200,001 plus
9. Prefer not to respond

Ethnic Background
1. Asian
2. African American
3. Hispanic
4. White/Caucasian
5. Other
6. Prefer not to respond

Appendix B. Procedure for Administering and Reporting the DLDS (taken from Denison & Neale, 2001)

Who should participate? The *Denison Leadership Development Survey* is designed for leaders and managers within all types of organizations. The participant does not have to be at any particular organizational level or hold a specific title. The important criterion is that the participant is acting in some type of leadership *role* within the organization, whether that be as CEO, manager, supervisor, team leader, or special task force leader. The survey is designed to look at the participant's leadership skills and highlight areas of strength, as well as areas that would benefit from focused developmental efforts.

What criteria should a participant utilize when choosing respondents? The participant should distribute the surveys to people who work closely with him/her, and whose opinions and judgment s/he respects. The basic package for the *Denison Leadership Development Survey* includes 25 Description by Others Surveys. The participant or survey facilitator should contact us if more than 25 surveys are desired. If the participant desires results of peers and direct reports/subordinates to be displayed separately, s/he must be sure that a minimum of three surveys for each group is returned for scoring to ensure the confidentiality of individual respondents.

What will the participant see in the Summary Report? Depending on how many survey questionnaires the participant distributed and which co-workers s/he chose to respond to the survey, the Summary Report will contain as few as two profiles (Self and Combined Other), or as many as five profiles (Self, Combined Others, Direct Reports, Peers and Boss.) Three surveys for each category are necessary to protect the anonymity and confidentiality of each individual respondent.

Self and Combined Other Profile Sheet: The first profiles in the Summary Report are the Self and Combined Other profiles. At minimum the participant will always receive these two profiles. They appear together on the same page to enable comparison. This provides the participant with a visual image of how s/he perceives ability to lead, how co-workers perceive those abilities, and where the differences and similarities lie. *Self* - This survey contains participant's own responses about his/her leadership skills. The data in the resulting profile and line-item report come from a single survey. *Combined Other* - This is a composite of all the responses of "Others" completing the survey. At least three completed surveys are required to generate this profile and line-item data.

Peer and Direct Report Profiles: A participant's Summary Report may include one or both of the following profiles depending on the breakdown of the surveys received. *Peers* - This is a composite of co-workers who are at approximately the same level within the organization as the participant, or other colleagues. This category can also include customers, suppliers, etc. A minimum of three surveys is needed to generate this profile and line-item detail. *Direct Reports* - This is a composite of co-workers who report directly to the participant. A minimum of three surveys is needed to generate this profile and line-item detail. As with previous profiles, these profiles allow the participant to quickly compare skill area assessments from several vantage points within the

organization. These different perspectives provide an excellent foundation for individual development. Research on 360° feedback suggests that direct reports will generally rate a leader higher than peers.

Boss Profile: The final profile included in the Summary Report is based on the survey responses of the participant's boss(es). This profile is generally comprised of the responses from one individual, but in a matrix with two bosses or when the participant requests a survey from his/her boss' boss, the profile may represent data from more than one individual. This profile is provided only if the boss has given permission for his/her responses to be displayed separately. In some instances, the boss may find it undesirable to have data reported in a separate profile. In that case, the participant will not receive a separate boss profile and the boss data will be included only in the Combined Other profile.

Line-Item Report: In addition to the graphic profiles, the Summary Report includes a Line-Item Report that lists each item in the survey, groups them by index and provides percentile scores for each item broken down by respondent (self, combined other, boss, peers and/or direct reports). A participant's Summary Report will contain a minimum of two breakdowns on the Line-Item Report - Self and Combined Other. The other breakdowns are provided only when there are three or more surveys for that category. The Line-Item Reports are color-coded to match the *Denison Leadership Model* (the Involvement pages are printed in green, the Consistency pages in yellow, Adaptability pages in blue and Mission pages in red). You will find a "+" and "-" sign next to the highest and lowest 10 line-item scores. The Line-Item Reports are most helpful for developing individual action plans because they highlight scores for specific leadership skills and practices. It is an important part of the Summary Report because it helps to identify the specific areas of strength and weakness that are not detailed in the graphic profile.

Overview of how the system works: Each participant in a study is e-mailed an invitation letter with their own unique **3Click™ link.** By clicking on the link, they can define their raters and relationships (e.g., Boss, Direct Report, Peers), take the "self" survey, track survey counts and download their own report (optional). Raters receive an e-mail from the participant (sent out by the **3Click™ system**), which explains the study and provides a link to the on-line survey. Reminder letters are automatically sent to only those raters who have not completed the survey. Participants can see their survey counts by relationship category (but they can not identify which participants have chosen to respond or not respond). The survey data is saved in a database system. If participants or raters disconnect from the survey (to continue later or because of internet connection issues), their partial survey data is saved and they can finish the survey at a later time. Electronic reports can be requested by the participants and/or, optionally, the project administrator can have the reports produced in batch mode for all the participants.

Appendix C. Description of Denison Leadership Development Survey Items (taken from Denison & Neale, 2001)

Involvement: *Building human capability, ownership, and responsibility.* Individual managers who know how to create "high-involvement" strongly encourage others to be involved and create an environment of experimentation and exploration, as well as a sense of ownership and responsibility. Highly-involved individual managers depend on informal, voluntary, and implicit leadership skills to move their work group or organization forward rather than formal, explicit, bureaucratic directives. Out of this sense of ownership grows a greater commitment to the organization, an increasing capacity for leadership, and a sense of autonomy. Receptivity to the ideas of others increases leadership quality and improves implementation of new ideas. The skills of the Involvement Trait are:

- Empowers People
- Builds Team Orientation
- Develops Organizational Capability

Empowers People – *Creates an environment where individuals have the authority, initiative and ability to develop and manage their own work. Employees have a sense of ownership and responsibility toward the organization.* Survey items that comprise the Empowerment skills are:

- Sees that decisions are made at the lowest possible level.
- Shares information so that everyone gets the information s/he needs.
- Creates an environment where everyone feels that his/her effort can make a difference.
- Involves everyone in shaping the plans and decisions that affect him/her.
- Conveys confidence in people's ability to do their job.
- Ensures that the necessary resources are available to do the job.
- Encourages others to take responsibility.
- Delegates authority so that others can do their work more effectively.

Builds Team Orientation – *Places value towards working cooperatively toward common goals and knows how to use team effort to get work done. Establishes a sense of mutual accountability for the accomplishment of goals.* The survey items that comprise the Builds Team Orientation skills are:

- Builds effective teams that get the job done.
- Encourages effective teamwork by others.
- Knows how to use a team approach to solve problems.
- Knows when to use a team approach to solve problems.
- Fosters teamwork within the work unit.
- Knows how to design work so that it can be done by a team.
- Values the contributions of the people s/he works with.
- Acknowledges and celebrates team accomplishments.

Develops Organizational Capability – *Continually focuses on the development of employees' skills and knowledge to meet ongoing organizational needs. Effectively uses the diverse capabilities of the workforce.* The survey items that comprise the Develops Organizational Capability skills are:

- Builds the capabilities of employees into an important source of competitive advantage. • Knows how to utilize the diversity of the work force.
- Coaches others in the development of their skills.
- Is sensitive and responsive to diversity issues when dealing with others.
- Helps direct reports create realistic development plans.
- Uses rewards and recognition to motivate good performance.
- Develops direct reports/subordinates so that they are ready for promotion.
- Builds employee skills so that the organization always has good "bench strength."

Trait Two: Consistency

Defining the values and systems that are the basis of strong leadership. Consistency provides a central source of integration, coordination, and control. Consistent individual managers develop a mindset and a set of operations that create an internal system of governance based
on consensus. They have highly committed employees, key central values, a distinct method of doing business, a tendency to promote from within, and a clear set of "do's and don'ts."

Consistency produces leadership based on a shared system of beliefs, values, and symbols that are widely understood by members of a work group or organization. Implicit control systems based on internalized values can be a more effective means of achieving
coordination and integration than external-control systems that rely on explicit rules and regulations. The power of leadership consistency is particularly apparent when organizational members encounter unfamiliar situations, when it enables leadership to react in a predictable way to an unpredictable environment by emphasizing a few general, value-based principles on which actions can be grounded. The skills of the Consistency Trait are:

- Defines Core Values
- Works to Reach Agreement
- Manages Coordination and Integration

Defines Core Values – *Communicates and lives by a set of nonnegotiable core values. Helps to define the work group or organization's culture, values and ethics; and helps employees learn to apply the organization's values when dealing with customers,
stakeholders and other employees.* The survey items that comprise the Defines Core Values skills are:

- Does the right thing even when it is not popular.
- Practices what s/he preaches.
- Has an ethical code that guides his/her behavior.
- Helps define the organization's culture, values and ethical standards.

- Helps employees learn to apply the organization's values when dealing with customers, stakeholders and other employees.
- Lives up to promises and commitments.
- Has earned the confidence and trust of others.
- Clearly articulates a set of fundamental beliefs that are not negotiable.

Works to Reach Agreement – *Helps to reconcile differences when they occur by actively promoting constructive discussion of conflicting ideas, incorporating diverse points of view into decisions, and working toward win-win solutions.* The survey items that comprise the Works to Reach Agreement skills are:

- Helps people to reach consensus, even on difficult issues
- Works to find alternatives that will benefit all when confronted with a disagreement.
- Helps people in his/her organization be effective at reaching agreement on key issues.
- Promotes constructive discussion among people with conflicting ideas.
- Is willing to compromise when necessary in order to reach agreement.
- Works toward win/win solutions when disagreements occur.
- Reconciles differences by seeking to clarify and understand other's points of view.

Manages Coordination and Integration – *Ensures that different functions or units of the work group or organization are able to work together well to achieve common goals. Establishes necessary contacts and coordinates resources in other groups to prevent organizational boundaries from interfering with getting work done.* The survey items that comprise the Coordination and Integration skills are:

- Works hard to foster the alignment of goals across all functional areas.
- Builds coordination across departmental boundaries.
- Uses informal networks to get things done.
- Builds relationships with key people in other functions and levels.
- Helps create an environment that facilitates coordination of projects across functional units.
- Makes certain that things do not fall between the cracks.
- Builds support for ideas through contact with other departments.
- Establishes mechanisms that facilitate effective cross-functional communication.

Trait Three: Adaptability - *Translating the demands of the organizational environment into action.* Successful individual managers hold a system of norms and beliefs that support his/her capacity to receive and interpret signals from the environment, and to translate them into internal behavioral changes that increase the work group or organization's chances for survival, growth, and development. Three aspects of adaptability influence an individual manager's effectiveness. First is the ability to perceive and respond to the external environment. Successful individual managers are very focused on their customers and their competitors. Second is the ability to respond to internal customers, regardless of level, department, or function. Third is the capacity to restructure and reinstitutionalize a set of behaviors and processes that allow the organization and its employees to adapt. Without this ability to implement adaptive response, an organization cannot be effective. The skills of the Adaptability Trait are:

- Creates Change
- Promotes Organizational Learning
- Emphasizes Customer Focus

Creates Change – *Knows the organizational environment and quickly reacts to current trends, and anticipates future changes. Continually creates adaptive and innovative ways to meet changing needs.* The survey items that comprise the Creates Change skills are:

- Continuously looks for new and better ways to do work.
- Encourages creative thinking.
- Challenges the way that things have always been done and looks for a better way.
- Champions change that goes beyond the scope of his/her job.
- Challenges organizational practices that are nonproductive.
- Foresees problems before they arise.
- Serves as a model that creates change in other parts of the organization.
- Generates innovative ideas and solutions to problems.

Emphasizes Customer Focus – *Driven to clearly understand the present and future needs of the customer, seeks ongoing input from the customer, continuously strives to improve customer service, and ensures that all employees are driven by a concern to satisfy the customer.* The survey items that comprise the Emphasizes Customer Focus skills are:

- Encourages direct contact with customers.
- Responds quickly and effectively to customer feedback.
- Ensures that employees have a deep understanding of customer wants and needs.
- Uses customer comments and recommendations to change organizational practices.
- Actively seeks feedback from customers.
- Continuously tries to improve service to customers.
- Incorporates customer input into the planning process.
- Recognizes the need to respond quickly to customer concerns.

Promotes Organizational Learning – *Encourages innovation, risk taking and continuous improvement. Sees mistakes as opportunities for gaining knowledge and developing capabilities.* The survey items that comprise the Promotes Organizational Learning skills are:

- Deals constructively with failures and mistakes.
- Views failure as an opportunity for learning and improvement.
- Creates a working environment in which learning is an important objective.
- Openly accepts criticism without being defensive.
- Works well under conditions of ambiguity and uncertainty.
- Knows the strengths and weaknesses of the competition.
- Encourages others to learn about the best practices in the industry.
- Helps others to understand the big picture

Trait Four: Mission

Defining a meaningful long-term direction for the organization The individual manager's mission provides purpose and meaning by defining a social role and external goals for his/her functional area or unit. It provides a clear direction and goals that serve to define an appropriate course of action for the individual manager and his/her employees. The individual manager is able to align the mission and goals for his/her functional area or unit to the mission and goals of the organization. A sense of mission allows an individual manager to inspire, to direct activities, and to formulate strategy by envisioning a desired future state. Being able to translate his/her mission into action contributes to both short and long-term commitment to the organization. Success is more likely when individual managers and organizations are goal directed. The skills of the Mission Trait are:

- Defines Strategic Direction and Intent
- Defines Goals and Objectives
- Creates Shared Vision

Defines Strategic Direction – *Communicates the organization's overall strategies so that everyone can see the relationship between their work and the accomplishment of the work group or organization's goals. Effectively implements short and long-term strategies to meet organizational needs.* The survey items that comprise the Defines Strategic Direction and Intent skills are:

- Provides employees with a clear mission that gives direction to their work.
- Implements strategies by developing clear goals, objectives and tactics.
- Focuses on long-term strategies, rather than quick fix "band- aid" solutions.
- Effectively allocates resources in line with strategic priorities.
- Helps define strategies and tactics that keep his/her organization competitive.
- Has a clear strategy for the future of his/her own part of the organization.
- Is able to meet short-term demands without losing sight of the long-term strategy.
- Communicates a clear and compelling rationale for the business strategy.

Defines Goals and Objectives – *Encourages high employee accountability in setting and accomplishing organizational goals. Communicates a clear set of goals and objectives that can be linked to the mission, vision and strategy of the work group or organization.* The survey items that comprise the Defines Goals and Objectives skills are:

- Sets clear goals that are ambitious, but realistic.
- Holds individuals and teams accountable for achieving goals and objectives.
- Provides clear directions and priorities for employees.
- Establishes high standards of performance.
- Involves employees in the goal-setting process.
- Tracks progress against stated goals.
- Effectively communicates the goals and objectives of the organization.
- Aligns goals and objectives with the strategy and vision.

Creates Shared Vision – *Helps create a shared view of a desired future state for his/her organizational unit. Inspires others with this vision, translates it into everyday activities and engages others to ensure buy-in and commitment.* The survey items that compromise the Creates Shared Vision skills are:

- Helps create a shared vision of what the organization will be like in the future. Communicates the organizational vision to his/her employees.
- Uses the vision to create excitement and motivation for employees.
- Realizes short-term goals without compromising long-term vision.
- Organizes work so that everyone sees the connection between the vision and his/her day-to-day activities.
- Translates the vision into reality in a way that helps guide individual action.
- Inspires others with his/her vision of the future.
- Engages others in ways that ensure buy-in and commitment.

Appendix D. Communal and Agentic Development Items

Communal Items (32 total)

1. Sees that decisions are made at the lowest possible level.
2. Shares information so that everyone gets the information s/he needs.
3. Creates an environment where everyone feels that his/her effort can make a difference.
4. Involves everyone in shaping the plans and decisions that affect them.
6. Conveys confidence in people's competence to do their job.
7. Encourages others to take responsibility.
10. Encourages effective teamwork by others.
11. Knows how to use a team approach to solve problems.
12. Knows when to use a team approach to solve problems.
13. Fosters teamwork within the work unit.
14. Knows how to design work so that it can be done by a team.
15. Values the contributions of the people s/he works with.
16. Acknowledges and celebrates team accomplishments.
19. Coaches others in the development of their skills.
20. Is sensitive and responsive to diversity issues when dealing with others.
23. Develops his/her own people so that they are ready for promotion.
27. Has an ethical code that guides his/her behavior.
29. Helps employees learn to apply the organization's values when dealing with others.
33. Helps people to reach consensus, even on difficult issues.
34. Works to find alternatives that will benefit all when confronted with a disagreement.
35. Helps people in his/her organization be effective at reaching agreement on key issues.
36. Incorporates diverse points of view when making decisions.
37. Promotes constructive discussion among people with conflicting ideas.
38. Is willing to compromise when necessary in order to reach agreement.
39. Works toward win/win solutions when disagreements occur.
40. Reconciles differences by seeking to clarify and understand other's points of view.
42. Builds coordination across departmental boundaries.
48. Establishes mechanisms that facilitate effective cross-functional communication.
59. Ensures that employees have a deep understanding of customer wants and needs.
61. Actively seeks feedback from customers.
63. Incorporates customer input into the planning process.
85. Involves employees in the goal-setting process so goals and objectives are understood and shared.

Appendix D (continued). Communal and Agentic Development Items

Agentic Items (22 total)

32. Clearly articulates a set of fundamental beliefs that are not negotiable.
49. Continuously looks for new and better ways to do work.
51. Challenges the way that things have always been done and looks for a better way.
52. Champions change that goes beyond the scope of his/her job.
53. Challenges organizational practices that are nonproductive.
55. Serves as a model that creates change in other parts of the organization.
70. Knows the strengths and weaknesses of the competition.
73. Provides employees with a clear mission that gives meaning and direction to their work.
74. Implements strategies by developing clear goals, objectives, and tactics.
75. Focuses on long-term strategies, rather than quick fix "band-aid" solutions.
76. Effectively allocates resources in line with strategic priorities.
77. Helps define strategies and tactics that keep his/her organization competitive.
78. Has a clear strategy for the future of his/her own part of the organization.
79. Is able to meet short-term demands without losing sight of the long-term strategy.
80. Communicates a clear and compelling rationale for the business strategy.
81. Sets clear goals that are ambitious, but realistic.
82. Holds individuals and teams accountable for achieving goals and objectives.
83. Provides clear directions and priorities for employees.
84. Establishes high standards of performance.
86. Tracks progress against stated goals.
88. Aligns goals and objectives with the strategy and vision.
94. Translates the vision into reality in a way that helps guide individual action.

Appendix E. Denison Leadership Development Survey Effectiveness Items (Denison & Neale, 1996)

1. Overall, this individual is a highly effective leader.
2. This individual's leadership style serves as a role model for others in the organization.
3. This individual has great potential as a future leader in our organization.
4. Overall, this individual is one of the most capable leaders in our organization.
5. This individual develops high quality relationships with internal and external customers.
6. This individual and his/her organization are consistently high performers.
7. This individual is capable of leading the organization through future changes and transitions.

Appendix F. Denison Leadership Development Survey Organizational Level Options (Denison & Neale, 1996)

Organization Level

1. Non-management
2. Line management (supervising non-management personnel)
3. Middle management (managing managers)
4. Senior Management
5. Executive/Senior Vice President
6. CEO/President
7. Owner
8. Prefer not to respond

Appendix G. Denison Leadership Development Survey Functional Area Options and Corresponding DOC Codes (Denison & Neale, 1996)

1. Finance and Accounting
   (DOC Code: Financial Managers; 52.1% women)

2. Engineering
   (DOC Code: Engineers; 10.4% women)

3. Manufacturing and Production
   (DOC Code: Production Occupations; 24.6% women)

4. Research and Development
   (DOC Code: Scientists; 34.3% women)

5. Sales and Marketing
   (DOC Code: Sales Occupations; 49.4% women)

6. Purchasing
   (DOC Code: Purchasing Managers; 42.6% women)

7. Human Resources
   (DOC Code: Personnel and Labor Relations Managers; 68.2% women)

8. Administration
   (DOC Code: Executive, Administrative, and Managerial; 46% women)

9. Support Staff
   (DOC Code: Administrative Support Occupations, Including Clerical; 78.7% women)

10. Professional Staff
    (DOC Code: Professional Speciality; 53.7% women)

11. Prefer not to respond