

# Gender and the Effectiveness of Leaders: A Meta-Analysis

Alice H. Eagly, Steven J. Karau, and Mona G. Makhijani  
Purdue University

This article presents a synthesis of research on the relative effectiveness of women and men who occupy leadership and managerial roles. Aggregated over the organizational and laboratory experimental studies in the sample, male and female leaders were equally effective. However, consistent with the assumption that the congruence of leadership roles with leaders' gender enhances effectiveness, men were more effective than women in roles that were defined in more masculine terms, and women were more effective than men in roles that were defined in less masculine terms. Also, men were more effective than women to the extent that leader and subordinate roles were male-dominated numerically. These and other findings are discussed from the perspective of social-role theory of sex differences in social behavior as well as from alternative perspectives.

As women gain greater access to leadership and managerial roles in organizations, it becomes increasingly important to understand the nature and extent of the similarities and differences between female and male leaders. One important set of issues concerns the effectiveness of leaders, when effectiveness is understood in terms of leaders' facilitation of a group or organization's ability to meet its goals (e.g., Hunt, 1991; Mumford, 1986). The specific questions we address in this article are whether one sex is generally more effective in leadership roles and whether there are conditions that tend to produce sex differences in the effectiveness of leaders and managers.<sup>1</sup>

These issues may be highly consequential for an organization's productivity. For example, if women and men are equally effective in leading and managing, then discriminatory barriers limiting women's access to leadership roles—the much-discussed *glass ceiling* (see A. M. Morrison, White, & Van Velsor, 1987)—would reduce an organization's productivity by removing a substantial proportion of the available pool of managerial talent. Although the glass ceiling may be cracking, some skepticism about women's ability to lead remains, especially among male executives (Sutton & Moore, 1985), and many people still express a preference for a male boss (Gallup, 1990; Rubner, 1991; Simon & Landis, 1989). Given such attitudes, it is worthwhile to determine empirically whether women and men are

equally successful when they occupy leadership or managerial roles.

The effectiveness of male and female leaders has been discussed by quite a few social and organizational psychologists (e.g., Denmark, 1993; Hollander, 1992; A. M. Morrison & Von Glinow, 1990; Powell, 1993; Ragins & Sundstrom, 1989; Riger & Galligan, 1980). Most authors have disavowed any pervasive sex difference in effectiveness. For example, Hollander (1992) stated that there is no difference in effectiveness, although “women begin with an initial hurdle to attaining legitimacy” (p. 72). Powell (1993) noted that, “Women and men do not differ in their effectiveness as leaders, although some situations favor women and others favor men” (p. 175). Bass (1990) acknowledged some evidence that male leaders were evaluated more favorably than female leaders but attributed this trend to observers' biases and stereotyped expectations. In general, there is no clear consensus in earlier discussions of whether the effectiveness of male and female leaders differs in general or in particular circumstances. Moreover, prior efforts to invoke the relevant empirical literature have been buttressed by citations of only a small proportion of the available studies, which were selected by unknown criteria (e.g., Bartol & Martin, 1986; Bass, 1990; Dobbins & Platz, 1986). To address these issues on a more systematic basis, we quantitatively synthesized the large number of studies that have compared the effectiveness of women and men who occupy leadership and managerial roles.

The aggregation of findings produced by this synthesis addresses the very general question of whether the male and female leaders included in this research literature differed in leadership effectiveness. In addition, this synthesis identifies charac-

---

Alice H. Eagly, Steven J. Karau, and Mona G. Makhijani, Department of Psychological Sciences, Purdue University. Steven J. Karau is now at the Department of Psychology, Virginia Commonwealth University. Mona G. Makhijani is now at A. C. Nielsen, Northbrook, Illinois.

This research was supported by National Science Foundation Grants BNS-8605256 and BNS-8807495. We thank Rebecca Henry, David Hoffman, Blair Johnson, Bruce Klonsky, and Wendy Wood for their comments on a draft of the article.

A table showing the effect sizes and study characteristics for each study included in the meta-analysis is available from Alice H. Eagly.

Correspondence concerning this article should be addressed to Alice H. Eagly, Department of Psychological Sciences, 1364 Psychological Sciences Building, Purdue University, West Lafayette, Indiana 47907-1364. Electronic mail may be sent via Internet to [eagly@psych.purdue.edu](mailto:eagly@psych.purdue.edu).

---

<sup>1</sup> In this article, the *sexes* denotes the grouping of people into female and male categories. The terms *sex differences* and *similarities* as well as *sex-related* or *sex-correlated differences* are applied to describe the results of comparing these two groups. Consistent with ordinary usage in psychology, these terms are not intended to give priority to any class of causes that may underlie differences and similarities. The term *gender* refers to the meanings that societies and individuals ascribe to female and male categories. Therefore, gender is an appropriate term to apply to products of the culture, such as gender stereotypes and gender roles (see Eagly, 1987).

teristics of studies that moderate tendencies for men or women to be more effective. Identifying these characteristics should facilitate progress in understanding the social and psychological processes that may account for women's and men's effectiveness as leaders and managers. Exploring the reasons why women or men may be differentially effective is thus a secondary purpose of this project.

### *Theoretical Analyses of Sex Differences in Leadership Effectiveness*

*Social-role theory.* Among the theories that yield predictions about the effectiveness of male and female leaders is Eagly's (1987) social-role theory of sex differences in social behavior, which maintains that as a general tendency people are expected to engage in activities that are consistent with their culturally defined gender roles. Social pressures external to individuals generally favor gender role consistent behavior, and, to some extent, people may internalize cultural expectations about their sex and consequently be intrinsically motivated to act in a manner consistent with their gender roles. Any such pressures favoring behavior congruent with one's gender role could be problematic for women occupying leadership or managerial roles because of the alignment of these social roles with stereotypically male qualities and therefore with the male gender role (Heilman, Block, Martell, & Simon, 1989; Schein, 1973). Although the definition of managerial roles as requiring qualities traditionally regarded as masculine may have moderated somewhat, at least among women (Brenner, Tomkiewicz, & Schein, 1989; Frank, 1988; Russell, Rush, & Herd, 1988), female managers may often face a degree of role conflict, as numerous organizational theorists have argued (Bass, 1990; Bayes & Newton, 1978; Kruse & Wintermantel, 1986; P. Y. Martin, 1992; O'Leary, 1974; Ragins & Sundstrom, 1989). By fulfilling people's expectations concerning leaders or managers, women may violate conventions concerning appropriate female behavior.

To the extent that female leaders violate their associates' gender expectancies, they may be subjected to prejudiced reactions, which may include biased performance evaluations and negative preconceptions about future performance. Consistent with research on self-fulfilling prophecies (Geis, 1993; Miller & Turnbull, 1986), negative preconceptions about women's leadership can diminish their performance. One mechanism that may mediate an effect of negative expectancies on performance is a lowering of female leaders' self-confidence (Heilman, Simon, & Repper, 1987; Lenny, 1977; Powell, 1993). Although competent female managers may win over skeptics in the long run and overcome any lack of self-confidence, male leaders may have an advantage over female leaders and may be somewhat more effective on the average because they are less likely to be subjected to prejudiced reactions.

A deeper analysis from a social-role perspective suggests that this role conflict analysis is somewhat too simple. If conflict between a woman's gender role and her leadership role could be alleviated, any decline in performance might be averted. One factor that should affect degree of role conflict is the extent to which particular leadership roles are given a masculine definition. Although leadership roles may in general be aligned more strongly with the male gender role than the female gender role, roles within certain occupational categories or certain types of

organizations may be defined in more androgynous terms. For example, expectations for clerical or nursing supervisors may be more androgynous than those for military officers or industrial foremen. Expectations for administrators of elementary schools may be more androgynous than those for managers in business organizations.

Role conflict for female leaders might also be minimized by adopting a relatively feminine leadership style that would meet people's traditional expectations about female behavior. This idea is consistent with Eagly, Makhijani, and Klonsky's (1992) synthesis of experiments that examined evaluations of female and male leaders whose behavior had been made equivalent by the researchers (e.g., Lee & Alvares, 1977; Rosen & Jerdee, 1973). This synthesis showed that female and male leaders were evaluated equally favorably when they adopted equivalent leadership styles that were stereotypically feminine (i.e., democratic or interpersonally oriented). In contrast, female leaders were devalued relative to their male counterparts when they adopted equivalent leadership styles that were stereotypically masculine, particularly when this style was autocratic and directive.

These meta-analytic findings on prejudice suggest that female leaders may have the possibility of shaping their behavior to minimize gender-role violations. As Sheppard (1992) argued, many female managers may strive to display behavior that is both sufficiently businesslike and professional that they are credible as managers and sufficiently feminine that they do not challenge associates' assumptions about gender. Consistent with this reasoning, Eagly and Johnson's (1990) synthesis of studies that compared the leadership styles of women and men who occupied the same leadership or managerial role showed that, on the average, female leaders, including managers in organizations, adopted a relatively democratic and participative style consistent with the female gender role.

Adopting a feminine leadership style may not provide women with a sure route to unbiased evaluations of their competence as leaders. In certain circumstances even those female leaders who avoid masculine leadership styles may be at risk for prejudiced evaluations and lowered effectiveness. Suggesting such circumstances are Eagly et al.'s (1992) findings that the devaluation of women in leadership roles was stronger when leaders occupied especially male-dominated roles and when their evaluators were men. Women's occupancy of highly male-dominated leadership roles no doubt produces a more severe violation of people's expectancies about women than does occupancy of less male-dominated roles, perhaps because roles in which women are rare become particularly strongly identified with masculine characteristics (see Gutek & Cohen, 1987; Williams, 1989). Moreover, male evaluators, more than female evaluators, may experience female leaders as a more threatening intrusion because leadership is traditionally a male prerogative. These findings lead us to predict that women may be less effective than men (or at least *perceived* as less effective) in leadership roles that are rarely occupied by women and in which most of their evaluators are male.

*Alternative perspectives.* Perspectives other than social-role theory also yield predictions about women's and men's effectiveness as leaders. A *structural* perspective takes into account only the formal role structure of groups and organizations and suggests that leadership or managerial roles provide powerful guides to behavior, quite aside from the sex of the role

occupant (e.g., Kanter, 1977). Particularly in organizational settings, people develop expectations about the behavior of leaders or managers (Phillips & Lord, 1982), and these specific expectations should be important determinants of behavior, far more important than expectations based on gender. This structural view suggests that men and women who occupy the same leadership role elicit similar reactions from others and are equally effective, as long as they have equivalent access to status and power. In the absence of sex-correlated differences in individual traits or predispositions that may influence the quality of leaders' role performances (House, 1991; Hunt, 1991), the sex of role occupants should have little or no influence on leaders' effectiveness.

Empirical research has revealed the inadequacy of a structural perspective that fails to acknowledge that people's expectations about managers' behavior depend to some extent on managers' gender (e.g., Heilman et al., 1989; Russell et al., 1988). This phenomenon has been expressed by the concept of *gender-role* (or "*sex-role*") *spillover*, which is a carryover into the workplace of gender-based expectations for behavior (Guterk & Morasch, 1982; see also Nieva & Guterk, 1981). To some extent, gender spills over or contaminates roles that might more properly be defined solely in terms of a genderless organizational hierarchy.

A second alternative perspective takes into account the *differential selection* of men and women for leadership. This argument follows from the assumption that women more than men face formidable barriers to achieving positions of leadership. Indeed, Eagly and Karau's (1991) meta-analysis of studies of initially leaderless groups showed that even in the relatively unstructured settings of this research, men were more likely to emerge as leaders. A preference for men in managerial roles in organizations, especially at higher levels, has also been documented (Bowman, Worthy, & Greyser, 1965; Sutton & Moore, 1985). Moreover, perhaps because women anticipate a glass ceiling, they may be less likely to apply for leadership positions than equally qualified men (Powell & Butterfield, 1994). Because of the barriers that women face in achieving leadership roles, whatever the source of these barriers, those women who nonetheless attain these roles may be more qualified and competent than their male counterparts. To the extent that the adage that a woman has to be "twice as good as a man" is valid, women may be superior performers, especially in the longer run as they may erode negative preconceptions about their competence (see Craig & Jacobs, 1985).

This selection bias argument suggests that women who actually occupy leadership roles should excel relative to their male counterparts. Yet, the imposition of higher standards for women than men would weaken as women's access to leadership roles improves. The selection bias argument thus implies that women's competence advantage should have lessened over time, given their considerably improved access to managerial careers (Guterk, 1993; Jacobs, 1992) and the existence of circumstances under which affirmative action policies favor female candidates over male candidates (Powell & Butterfield, 1994). Moreover, differential selection would be less relevant to organizations with less formidable barriers against female managers.

Yet another perspective, that of *contingency theories* of leadership, suggests that leaders' effectiveness depends on their style

of leading in interaction with features of the situation (see reviews by Bass, 1990; Chemers, 1987; Yukl & Van Fleet, 1992). From this standpoint, women and men may differ in effectiveness, at least in some situations, to the extent that they have chronically different leadership styles. Contingency theories of course raise the controversial issue of whether women and men differ in leadership style.

Sex-related differences in leadership style were investigated in Eagly and Johnson's (1990) earlier synthesis of 162 leadership studies that produced 370 comparisons between men and women. Leadership styles tended to be somewhat gender stereotypical in laboratory experiments and slightly gender stereotypical in assessment studies that investigated the leadership styles of people not selected to occupy leadership roles (e.g., samples of employees or students). In these laboratory experiments and assessment studies, women tended to manifest relatively interpersonally oriented and democratic styles, whereas men tended to manifest relatively task-oriented and autocratic styles. In contrast, sex differences were more limited in organizational studies assessing managers' styles: The only demonstrated difference between female and male managers was that women adopted a somewhat more democratic or participative style and a less autocratic or directive style than did men. In view of the subsequent meta-analytic finding that people react especially negatively to women who adopt an autocratic and directive style (Eagly et al., 1992), it appears that women less frequently adopt the style that produces particularly prejudiced evaluations of their behavior. Sex-related differences in leadership style may thus reflect prejudice directed toward female leaders who adopt more masculine styles, consistent with such styles' violations of the norms associated with the female gender role. However, these sex-related differences could also be influenced by various other causes, such as (a) personality and ability differences, especially women's greater social skill and interest in other people (Eagly & Wood, 1991); (b) the learning of different styles of influence in sex-segregated play groups (e.g., Maccoby, 1990); or (c) biologically grounded differences between the sexes (e.g., Kenrick & Trost, 1993).

Any such sex difference in leadership style could be consequential for leaders' effectiveness because contingency theorists have generally focused on aspects of style that are inclusive of this distinction between participative and directive leadership. Yet, the details of the theories' predictions differ, and all of these theories predict that relations between leadership style and effectiveness are moderated by situational variables. For example, Fielder's contingency theory (Fiedler, 1967; Fiedler & Chemers, 1984) suggests that directive, task-oriented managerial behavior would be effective only in certain situations (particularly with simple tasks in relatively structured situations when the leader has good relationships with subordinates as well as in especially difficult situations that lack all of these features). Although Vroom and Yetton's decision-making model (Vroom & Jago, 1988; Vroom & Yetton, 1973) is in general agreement with Fiedler's predictions, it includes additional moderating variables such as the likelihood of conflict among subordinates. In contrast, House's (1971) path-goal theory suggests that a directive style would be effective to the extent that tasks are ambiguous and therefore would benefit from the leader's directive structuring (see also House & Mitchell, 1974). Drenth and Koopman (1984) argued that a participative style is

generally facilitative for short-term, routine decisions, but that its effects are more variable for longer-term, strategic decisions. Given the variety and complexity of these contingency theories, we cannot formally test their power to account for any observed sex differences in leaders' effectiveness. Moreover, the studies that provide comparisons between male and female leaders were not designed to test contingency theories and therefore rarely measured or manipulated the moderating variables crucial to these theories (or even included information that would allow the entire study to be classified with respect to these variables). Nonetheless, these theories' propositions may aid in interpreting some of the relations we obtain with moderating variables such as type of organization.

*Predictions for meta-analysis.* As we have explained, several theoretical frameworks offer predictions about the relative effectiveness of female and male leaders and suggest that effects would likely depend on features of the situation. Although social-role theory suggests that men may be more effective than women in leadership roles, any differences should be small in view of female leaders' demonstrated tendency to adopt leadership styles that are likely to minimize role conflict (Eagly & Johnson, 1990). However, men may fare better than women in leadership roles that have been defined in particularly masculine terms. In contrast, the structural theory assumption that organizational roles override any effects of gender roles argues for no sex differences in effectiveness, as long as male and female leaders occupy the same role (and do not differ in other role-relevant attributes such as seniority in the role). In addition, the selection bias assumption that women must meet higher standards than men to qualify for leadership roles suggests that women may be more effective than men, especially in the longer run and especially if they had to overcome sexist barriers to attaining these roles. The implications of the tendency for women to lead in a somewhat different style remain less clear because, as contingency theories predict, the effectiveness of leadership styles appears to interact with a number of moderating variables (e.g., task ambiguity or complexity) that cannot be adequately represented in this meta-analysis. In view of the predictions that follow from these several theoretical frameworks, any pervasive superiority of male or female leaders is unlikely. Instead, situational and organizational variables should moderate any sex-related differences that are obtained.

### *Design of the Synthesis*

*Assessment of leader effectiveness.* The issue of how leaders' effectiveness should be assessed has been addressed repeatedly by organizational psychologists (e.g., Bass, 1990; Hunt, 1991; Tsui, 1984; Yukl & Van Fleet, 1992). In most such discussions, psychologists eschewed defining effectiveness in terms of particular behavioral tendencies or styles of leadership because they regarded the relationship between style and effectiveness as an issue of considerable complexity in and of itself. However, there is some consensus that effective leaders facilitate the ability of a group or organization to meet its goals and maintain itself over time (e.g., Hunt, 1991; Mumford, 1986). Therefore, effectiveness is ordinarily regarded as an outcome of leaders' behavior rather than a particular type of behavior. Although there is no single method of assessing this outcome that can be free of hazards, a number of types of measures provide relevant informa-

tion. Indeed, most organizational psychologists advocate multiple criteria of effectiveness. Reflecting this preference, empirical studies have commonly assessed effectiveness by varied methods, including evaluative ratings of leaders' performances (by superiors, subordinates, peers, or leaders themselves), ratings of subordinates' satisfaction with leaders, and measures of group and organizational productivity. Reputational ratings are clearly relevant to determining leaders' effectiveness, even though persons in any given position (i.e., superiors, subordinates, peers, leaders) may provide biased or one-sided judgments of effectiveness. Although all subjective ratings of performance are potentially vulnerable to various kinds of biases (e.g., Frone, Adams, Rice, & Instone-Noonan, 1986; Landy & Farr, 1980; Schmidt, Ones, & Hunter, 1992; Uleman, 1991), there is some evidence that perceptions of leadership are consensual and on the whole moderately accurate (Malloy & Janowski, 1992). System-level outcome measures such as group or organizational productivity are relevant as well but do not directly or purely reflect the quality of particular leaders' performances, especially in large organizations. Although the diversity of methods that have been used to assess effectiveness should thus be regarded as a strength of this research area, this variety complicates the task of integrating research. Therefore, we included in our meta-analysis most measures that researchers have regarded as assessing effectiveness. We then coded the studies in the meta-analysis on a number of these measures' features to see if our findings proved generalizable across differing methods of assessing effectiveness.

*Types of studies of leaders' effectiveness.* In the research literature on leaders' effectiveness, there are two basic types of studies: a relatively small number of laboratory experiments, usually conducted with college students, and a much larger number of organizational studies, conducted with managers in a variety of types of organizations. These organizational studies compared male and female managers who occupied the same managerial role or type of role (e.g., middle manager). For example, in a study representative of the organizational sample, Tsui and Gutek (1984) studied middle managers in a multicompany corporation specializing in the computer, data services, education, and finance industries. A stratified random sample overrepresenting female managers produced a sample of 295 managers, each of whom selected one superior, two subordinates, and two peers to evaluate them. These associates and the manager completed questionnaires that elicited their ratings of the manager's effectiveness. Also, formal performance appraisals by the manager's superior were made available by the corporation.

In laboratory experiments, leaders are usually randomly appointed to lead fellow students in solving one or more problems in a single experimental session. For example, Jacobson and Effertz (1974) had 72 college students form 3-person groups that carried out a task requiring subordinates to place dominoes into a predetermined pattern known only to the leader. The experimenter selected this leader from the group to fulfill the requirements of a factorial design in which male or female leaders worked with two subordinates who were either both male or both female. The leader and the subordinates rated the leader's effectiveness, and the mean number of dominoes correctly placed by the subordinates provided an objective measure of performance. Although, as this example illustrates, labora-

tory experiments are very different from organizational studies, most of the theoretical principles that we reviewed should apply to leadership in experimental groups as well as in organizations.

*Congeniality of leadership roles for men and women.* As in Eagly and Johnson's (1990) project on leadership style, we found considerable variation in the extent to which the leadership roles investigated in this research would be perceived as congenial for women or men. As we argued earlier in this article, for leadership roles that are defined in relatively androgynous terms, conflict between the leader role and the female role should be minimized. In such roles women and men should be equally effective, whereas in roles that are defined in especially masculine terms, men may be more effective than women.

We labeled the fit between gender roles and particular leadership roles the *gender congeniality* of the leadership role. To enable us to take account of this consideration for the leadership roles that were examined in the research literature, we obtained questionnaire ratings of each role and analyzed these judgments to estimate the extent to which women or men were more interested in each role and believed themselves more competent to perform it. In addition, because people tend to associate task-oriented qualities with men and interpersonally oriented qualities with women (Deaux & Lewis, 1984; Eagly & Steffen, 1984), we also assessed the extent to which each role was judged to require these gender-stereotypic qualities. These strategies allowed us to determine the extent to which the ascription of gender-stereotypic qualities to leadership roles related to any sex differences in leaders' effectiveness.

## Method

### *Sample of Studies*

Computer-based information searches were conducted using the keywords *leadership performance* and *leadership effectiveness* as well as *leader*, *leadership*, *manager*, and *supervisor* when paired with terms such as *gender*, *sex*, *sex differences*, and *women*. These keywords were searched in the following databases from the beginning point of each database through the end of 1989: *Psychological Abstracts* (PsycINFO), *Dissertation Abstracts International* (DISS), Educational Resources Information Center (ERIC), and a worldwide business and management database (ABI/INFORM). References were also located in extensive earlier searches conducted for related meta-analyses (Eagly & Johnson, 1990; Eagly & Karau, 1991; Eagly, Karau, Miner, & Johnson, 1994; Eagly et al., 1992). Additionally, we searched through numerous review articles, chapters, and books, as well as the reference lists of all located studies.

Criteria for including studies in the sample were that (a) the study compared male and female leaders, managers, supervisors, officers, department heads, or coaches on a measure of leadership effectiveness; (b) participants were at least 14 years old, from the United States or Canada, and not sampled from abnormal populations;<sup>2</sup> (c) the study assessed the effectiveness of at least five leaders of each sex; (d) the sex of the leader and the sex of the subordinates were not completely confounded (as they would be in a study that compared female and male leaders of single-sex groups); and (e) the reported results were sufficient either to calculate a sex-of-participant effect size or to determine the statistical significance or direction of the sex difference. This last criterion eliminated studies that provided only a multiple regression equation in which sex appeared as one of the predictors as well as studies that provided only a multivariate analysis of variance on leadership effectiveness combined with other measures (e.g., Bruning & Snyder, 1981; A. C. Smith, 1982).

Studies included in the sample used one or more of the following types of effectiveness measures: (a) measures of leaders' effectiveness, performance, motivation to perform well, or leadership ability that were based on self-ratings or ratings by supervisors, peers, subordinates, or judges; (b) ratings (by self or others) of satisfaction with leaders or satisfaction with leaders' performance; (c) coding or counting of effective leadership behaviors by trained judges or observers; and (d) measures of organizational productivity or group performance. Studies that reported only measures of salary and promotion (e.g., Camden & Witt, 1983) were not included because these outcomes are determined by a variety of factors (e.g., selection by supervisors, job tenure) that do not necessarily reflect performance. Studies that reported only measures of influence, dominance, likability, or interpersonal attractiveness were also excluded (e.g., Wahrman & Pugh, 1974). Ratings of effectiveness also had to pertain to the general quality of the leaders' performance, rather than to specific stylistic features (e.g., consideration, initiation of structure). Thus, studies were excluded if leadership effectiveness was merely inferred on the basis of a measure of leadership style (e.g., Davenport, 1977; McMahon-Dumas, 1981). Moreover, studies were omitted if they reported only subordinates' global job satisfaction (without also measuring their satisfaction with the leader) or leaders' self-ratings of their job satisfaction (e.g., Klawitter, 1985; Perritt, 1983). In addition, studies were excluded if they examined only stereotypes of leaders or peoples' attitudes toward male and female leaders in general (e.g., Heilman et al., 1989).

Studies were omitted if leaders were selected to equalize their performance or their status on a personality or attitudinal variable (e.g., an index of masculinity or femininity) that probably correlated with both sex and leader effectiveness (e.g., Hatch, 1987; Stake, 1981); accurate estimation of any sex difference in leadership effectiveness is not possible from such studies. If leaders were selected for high or low performance, the proportion of male and female leaders selected was used as an effectiveness measure when the total distribution of male and female leaders was known. In these cases, however, effectiveness measures obtained from the resulting samples of leaders were not included because any differences in women's and men's scores would have been reduced by the selection into high- and low-performing groups.

Studies were eliminated if male and female leadership performance was standardized or made equivalent (e.g., through vignettes that varied only the leader's sex, or trained leaders or confederates who varied by sex; Lee & Alvares, 1977; Rosen & Jerdee, 1973). Also, studies of non-management employees performing "in-basket" exercises not involving group interaction were excluded because the participants in these studies do not assume a clear leadership role (e.g., Moses, 1973; Moses & Boehm, 1975). Studies reporting teachers' performance were also excluded (e.g., Arons, 1980) because the teachers did not occupy a clear leadership position. Finally, studies of the leadership effectiveness of participants in T-groups, encounter groups, and therapy groups were excluded (e.g., T. L. Morrison & Stein, 1985) because leadership in such groups is provided, at least in part, by experts or group facilitators.

### *Variables Coded From Each Study*

The following general information was coded from each report: (a) date of publication; (b) publication form (journal article, book or book chapter, dissertation, unpublished document); (c) percentage of male authors; (d) sex of first author; (e) number of observations;<sup>3</sup> (f) number of judgments (or discrete data records in cases of objective measures)

<sup>2</sup> Among studies meeting all of the inclusion criteria, none examined leaders younger than college students.

<sup>3</sup> The number of observations (*n*) for the statistical analysis typically represented either the number of leaders or the number of raters (e.g., subordinates) who described the leaders.

aggregated into each data point in the study's statistical analysis;<sup>4</sup> (g) confounding of male–female comparison with variables such as seniority, education, and age (controlled by matching; known to be confounded on some variables; unknown whether confounded and confounding likely; unknown whether confounded and confounding unlikely);<sup>5</sup> (h) setting of study (laboratory or organization [business, educational, government or social service, military, or miscellaneous]); and (i) size of group or organization in which leadership occurred (laboratory group of given size; small organization, defined as less than 500 long-term participants; large organization; mixed or unknown).

In addition, the following characteristics of the leaders and their roles were coded: (a) level of leadership (first or line; second or middle; third or higher; ambiguous, mixed, or unknown); (b) age of leaders; (c) percentage of men among leaders (estimated from census tables and other information if not stated in report); (d) percentage of men among subordinates (also estimated if not stated in report; unknown when subordinates not specifically identified); (e) basis of selection of organizational leaders (random sample or entire population; unsuccessful random sample [i.e., effort to obtain random sample or entire population but less than 80% participation], selected nonrandomly [e.g., self-selected] unclear); and (f) basis for laboratory leadership (appointed randomly; appointed based on qualifications; emerged; mixed or unclear).

Among the attributes we coded of the measures of leadership effectiveness, the most important was the type of effectiveness measure. Measures were first classified as objective performance measures or subjective measures. Examples of objective measures include groups' speed of completing a geometric puzzle (Eskilson, 1975), Wagner and Sternberg's (1985) test of managers' tacit knowledge of business (Horgan & Simeon, 1988), workers' productivity defined by the proportion of their production goals they met and their absenteeism (Szilagy, 1980), and seasonal win–loss records of athletic teams (Young, 1981).

The measures coded as subjective were classified as measures of performance or of satisfaction with leaders. The subjective performance measures were further coded into four classes:

1. Effectiveness: Examples include (a) a rating scale anchored by *poor leader* and *outstanding leader* (Day & Stogdill, 1972, p. 355) and (b) performance evaluations of cadets' leadership, given on 20 dimensions such as "delegates authority effectively" and "keeps troops motivated" (Mohr, Rowan, & Reidy, 1978, p. 3).

2. Ability: Examples include (a) ratings of 29 leadership skills such as "training officers/committee members" and "planning club/group activities," given on scales anchored by *excellent* and *poor* (Couch, 1980, pp. 114–115), and (b) ratings of nine managerial abilities such as decisiveness, leadership, and organizing and planning, given on nine scales anchored by *outstanding* and *poor* (Friedman, 1981, p. 67).

3. Effort or motivation to perform well: Examples include (a) agreement ratings on 12 motivational items such as "I like to start work on new things" and "I become discouraged easily in my work" (Waetjen, Schuerger, & Schwartz, 1979, p. 89) and (b) ratings of the effort managers "put forth on the job" on a scale anchored by *slightly below that of others* to *far exceeds that of others* (Renwick, 1977, p. 407).

4. Other or mixed: Examples include (a) frequency and intensity of feeling reported on the personal accomplishment subscale of the Maslach Burnout Inventory (Maslach & Jackson, 1981), which assesses "feelings of competence and successful achievement in one's work with people" (Caccese & Mayerberg, 1984, p. 281); (b) subordinates' ratings of their own performance on two items anchored by *excellent* and *poor* (Terborg & Shingledecker, 1983, p. 817); and (c) supervisors' ratings of their subordinates' performance in six areas such as "motivation to work hard," "potential for promotion," and "overall performance" on scales anchored by *outstanding* and *very poor* (Kipnis & Schmidt, 1988, p. 535).

Examples of the measures of satisfaction with leaders include (a) the satisfaction with supervisors scale of P. C. Smith, Kendall, and Hulin's (1969) Job Descriptive Index (E. F. Adams, 1978), (b) ratings of the extent to which subordinates would "like to continue working for this

supervisor" (Arnett, Higgins, & Priem, 1980, p. 143), and (c) ratings of the extent to which subordinates were satisfied with the "helpfulness of your supervisor in getting the job done" (Gupta, Jenkins, & Beehr, 1983, p. 178).

Finally, the following additional attributes of the measures of leadership effectiveness were coded: (a) identity of raters (leaders; supervisors of leaders; subordinates of leaders; peers of leaders; judges or trained observers; objective counting device, e.g., industrial productivity; mixed or unclear); (b) target of ratings, that is, whose performance was rated to infer leader's effectiveness (leaders; subordinates; other, mixed, or unclear);<sup>6</sup> (c) basis of selection of raters (random sample or entire population; unsuccessful random sample, i.e., effort to obtain random sample or entire population but less than 80% participation; nonrandom selection; unknown or irrelevant because leaders rated selves); (d) percentage of men among raters; and (e) reliability of effectiveness measure (reported value or unknown).<sup>7</sup>

All variables were independently coded by two of the authors, with a median agreement of 96% (estimated kappa = .92). "Size of group or organization in which leadership occurred" yielded the lowest agreement, 85% (kappa = .80). Disagreements were resolved by discussion.

### *Variables Constructed From Questionnaire Respondents' Judgments of Leadership Roles*

To generate measures of gender-relevant aspects of the leadership roles investigated in the organizational and laboratory studies, we conducted additional analyses using data obtained in a questionnaire study reported by Eagly and Johnson (1990). In that study, respondents rated a wide range of leadership roles on a number of gender-relevant dimensions. These data yielded ratings for all but five studies in the current sample.

In the questionnaire study, several measures were used to assess the degree to which the leadership roles were perceived as congenial for women and men. In particular, sex differences were calculated for respondents' self-reported competence to perform each role and interest

<sup>4</sup> For example, each data point might represent four observers' ratings of a leader on three items, yielding 12 judgments aggregated into each data point. To the extent that measures were based on multiple judgments of leaders' effectiveness, they should yield more reliable estimates of sex differences, in the manner that the number of items in a test relates to the reliability of the total test (e.g., Ghiselli, 1964).

<sup>5</sup> Relevant to the confounding issue is the fact that most organizational studies examined occupants of a given role (e.g., elementary school principals), but a few examined broader classifications of managers (e.g., middle managers of an organization). Especially with these broader classifications, the equivalence of the men and women in terms of attributes such as age, education, and job seniority is not ensured. Therefore, this aspect of our coding scheme took account of possible confounding between sex and other attributes of leaders. Some organizational studies lessened confounding by establishing samples of male and female leaders who were matched on various attributes, and other studies included data revealing the presence or absence of confounding. When such data were absent in organizational studies, we ordinarily coded confounding as unknown and likely because the dramatic gains in women's representation in managerial careers in recent years means that female managers tend to be younger and to have less seniority than male managers (Jacobs, 1992). However, when such data were absent and the leaders were students or military cadets, we coded confounding as unknown and unlikely.

<sup>6</sup> Because subordinates were rarely the unique target of ratings, this variable is not reported further.

<sup>7</sup> Reliability was reported for only 27% of the measures, precluding any corrections or weighting based on reliability information.

in performing each role as well as respondents' beliefs about differences in average men's and women's interest in performing each role.<sup>8</sup> Other measures assessed respondents' judgments of the extent to which each role requires interpersonally oriented ability and task-oriented ability.

*Respondents and procedure.* The sample consisted of 125 female and 181 male Purdue University undergraduates who received partial course credit for participation. Respondents participated in groups of about 15 and in sessions conducted by a female or male experimenter. Each respondent completed one of three versions of a questionnaire that took approximately 1 hr to complete. Each of the three versions contained brief descriptions of each of 119 leadership roles investigated in organizational or laboratory studies. Examples of descriptions used for organizational studies are (a) principal of an elementary school, (b) manager in a large government service agency, (c) middle manager in a manufacturing firm, and (d) coach of a boys' high-school basketball team. Examples of descriptions used for laboratory studies are (a) manager of a simulated engineering department of a large oil refinery for which the manager is given the responsibility of bolstering productivity; (b) leader of a laboratory group playing a simulated business game (a small manufacturing enterprise); and (c) leader of a laboratory group that is trying to solve a visual puzzle; the leader knows the solution but is not allowed to use visual means to tell group members.

In one version of the questionnaire, respondents judged the roles in response to two questions eliciting self-reports of their competence and interest in relation to each role: (a) How competent would you be as a [role description given]? and (b) How interested would you be in becoming a [role description given]? In a second version of the questionnaire, respondents judged the roles in response to two questions assessing their beliefs about women's and men's interest in the roles: (a) How interested would the average woman be in becoming a [role description given]? and (b) How interested would the average man be in becoming a [role description given]?<sup>9</sup> In a third version of the questionnaire, respondents judged the roles in response to two questions assessing their beliefs about the abilities each role required: (a) How much *ability to cooperate and get along with other people* is needed to be an effective [role description given]? and (b) How much *ability to direct and control people* is needed to be an effective [role description given]?

All ratings were made on 15-point scales. Each version of the questionnaire was divided into two parts, both of which elicited respondents' judgments of all of the roles in relation to one of the questions. The order of the two parts was counterbalanced. Within each part, the descriptions of the behaviors appeared in one of two random orders.

*Analysis of ratings.* For the two questions in the first version of the questionnaire, mean scores for each role were computed separately for female and male respondents. For each role, the female respondents' mean was subtracted from the male respondents' mean to yield a mean sex difference, which was standardized by dividing it by the pooled (within-sex) standard deviation. For the two questions in the second version of the questionnaire, the respondents' mean rating of the average woman for each role was subtracted from their mean rating of the average man to yield a mean stereotypic sex difference, which was standardized by dividing it by the standard deviation of the differences between the paired ratings. For the two questions in the third version of the questionnaire, a mean of all the respondents' ratings of each role was calculated. These five mean scores thus described each of the leadership roles in the organizational and laboratory studies. For studies reporting findings aggregated over several roles, ratings of the relevant roles were averaged (e.g., Gupta et al., 1983).

### Computation and Analysis of Effect Sizes

The effect size calculated was  $g$ , the difference between the leadership effectiveness of the men and women, divided by the pooled standard deviation (see Hedges & Olkin, 1985). A positive sign indicates that men were more effective than women, and a negative sign indicates that women were more effective than men.

Four studies presented data separately for leaders from two or more organizations or management levels within an organization (e.g., Osborn & Vicars, 1976). In these cases, each sample of leaders was treated as a separate study. Moreover, for the three laboratory studies that used experimental manipulations to create groups of subordinates that differed in sex composition (e.g., all male, all female, mixed sex; e.g., Eskilson, 1975), the experimental conditions differing in sex composition were treated as separate studies when sufficient findings were presented. Although the resulting partition of some of the studies created some (rather minimal) nonindependence in our data set, the questions we desired to address could not be answered without proceeding in this way. Using this strategy, 87 documents produced 96 studies.

*Computation of effect sizes.* To reduce computational error, the effect sizes were calculated independently by two of the authors with the aid of a computer program (Johnson, 1989). The statistical significance and direction of the sex comparisons were also recorded; this information provided the only record of the sex comparisons for studies that contained insufficient data to calculate effect sizes. If the data report was sufficient, separate effect sizes were calculated for different types of measures of effectiveness and for different identities of raters. The computation of  $g$  was based on (a) means and standard deviations or error terms for 83 of the effect sizes; (b)  $F$ s,  $t$ s, or  $z$ s for 20; (c) correlations or chi-squares for 14; (d) proportions of men and women judged as high or low in performance for 5;<sup>10</sup> and (e) exact  $ps$  or level  $ps$  (e.g.,  $p < .05$ ) for 3. In total, 125 effect sizes were computed.

The pooled standard deviation, which is the denominator of the effect size, was estimated, whenever possible, only from the portion of each study's data entering into the effect size. When the pooled standard deviation was estimated from the mean square error (MSE) of an analysis of variance (ANOVA), this error term was reconstituted by adding into the sums of squares error all (available) between-groups sums of squares except that for sex. One-way designs were approximated by this procedure, which has been recommended by Hedges and Becker (1986) and Glass, McGaw, and Smith (1981).

To adequately address all of our research questions, it was necessary to aggregate the effect sizes in various ways within each study that yielded more than one effect size. Initially, we calculated separate effect sizes for each of the seven identities of raters within each of the six types of effectiveness measures (see *Variables Coded From Each Study* for these rater identities and measure types). However, if a study yielded more than one of these effect sizes for a given rater identity within a particular type of measure, these effect sizes were aggregated from the outset. On the basis of preliminary analyses (see Results) and for simplicity of exposition, these effect sizes were then aggregated across the seven rater identities to produce overall effect sizes for each of the six types of measures (with each study contributing no more than one effect size for each type of measure). We then aggregated these effect sizes across these types of measures in stages. First, effect sizes were aggregated across the four types of measures that were subjective measures of performance, to allow a comparison between these subjective perfor-

<sup>8</sup> In this article, the term *respondents* designates people who participated in the questionnaire study and not those who participated in the studies included in the meta-analysis.

<sup>9</sup> For the questions in the first and second versions of the questionnaire, respondents were told to assume that they (or the average man and woman) could obtain or had obtained any training or education required for the role.

<sup>10</sup> These calculations were based on the proportion of male versus female leaders who were rated as high or low on performance or whose performance was judged to be successful or unsuccessful. Proportions were transformed to effect sizes by treating each proportion as the mean of a distribution of 0s and 1s (McNemar, 1962). Thus, the effect size was the difference between the male and female proportions divided by the pooled standard deviation of the male and female samples of 0s and 1s.

Table 1  
Summary of Study Characteristics

Variable and class	Value	Variable and class	Value
Median publication year	1980	Level of leadership	
Publication form		First or line	43
Journal article	52	Second or middle	22
Book or book chapter	5	Third or higher	1
Dissertation or master's thesis	28	Ambiguous, mixed, or unknown	30
Unpublished document	11	Median age of leaders	30
Median percentage of male authors	50	Laboratory	20
Median no. observations	112	Organizational	38
Median no. judgments aggregated into each observation	8	Median percentage of men among leaders	60
Confounding of male-female comparison		Laboratory	50
Controlled via matching	12	Organizational	73
Known	27	Median percentage of men among subordinates	48
Unknown and likely	34	Laboratory	50
Unknown and unlikely	23	Organizational	28
Setting of study		Basis of selection of organizational leaders	
Laboratory	22	Random sample	24
Organization	74	Unsuccessful random sample	26
Business	22	Nonrandom	4
Educational	21	Unclear	20
Governmental or social service	7	Basis for laboratory leadership	
Military	10	Appointed randomly	14
Miscellaneous	14	Emerged	5
Size of group or organization		Unclear or mixed	3
Median no. members in laboratory group	3	Mean respondent judgments of roles	
Size of organization		Competence sex difference <sup>a</sup>	0.06
Small	13	Interest sex difference	0.00
Large	27	Stereotypic interest difference	0.11*
Mixed or unknown	34	Interpersonal ability rating <sup>b</sup>	10.56
		Task ability rating	10.50

Note. For categorical variables, numbers in the table represent frequency of sex comparisons in each class. Summaries of continuous variables are based on reports for which information was available on each variable.

<sup>a</sup> For the first three variables constructed from judgments of leadership roles, values are positive for differences in the masculine direction (greater male estimates of competence and of interest; ascription of greater interest to average men). <sup>b</sup> For the last two variables constructed from judgments of the leadership roles, values are larger to the extent that a role was judged to require more interpersonal or task ability (on 15-point scales with 15 indicating high ability).

\* Differs significantly ( $p < .05$ ) from 0.00 (exactly no difference).

mance measures and the objective performance measures. Second, the effect sizes for objective and subjective performance were aggregated into overall performance effect sizes. Third, these overall performance effect sizes were combined with the effect sizes for the measures that assessed satisfaction with leaders to produce *study-level effect sizes*. This last step was accompanied by analyses showing that the characteristics of the studies predicted the performance effect sizes and the study-level effect sizes in a similar manner. These study-level effect sizes were used in most of the analyses that we report in this article, in order to satisfy the independence assumption of meta-analytic statistics. All combined effect sizes were calculated using Rosenthal and Rubin's (1986) suggested formula.<sup>11</sup>

When the raters who provided the leadership effectiveness measure were different from the leaders or managers rated (e.g., they were subordinates) and these ratings were reported separately for male and female raters, effect sizes were calculated separately for the male and female raters (and the significance and direction of the sex comparisons were recorded separately). These additional effect sizes supplemented the effect sizes that were combined over both sexes of raters and used in most analyses.

*Analysis of effect sizes.* Following Hedges and Olkin's (1985) procedures, the  $g$ s were converted to  $d$ s by correcting them for bias. All mean  $d$ s were computed with each effect size weighted by the reciprocal of its variance, a procedure that gives more weight to effect sizes that are more reliably estimated. To determine whether each set of  $d$ s shared a com-

mon effect size (i.e., was consistent across the studies), we calculated a homogeneity statistic  $Q$ , which has an approximate chi-square distribution with  $k - 1$  degrees of freedom, where  $k$  is the number of effect sizes.

In the absence of homogeneity, we accounted for variability in heterogeneous effect sizes by calculating categorical and continuous models that related the effect sizes to the attributes of the studies (Hedges & Olkin, 1985). Calculation of a categorical model provides a between-classes effect,  $Q_B$ , and a test of the homogeneity of the effect sizes within each class,  $Q_{wi}$ . The continuous models are least squares simple linear and multiple regressions, calculated with each effect size weighted by the reciprocal of its variance. Each such model yields a test of the significance of each predictor as well as a test of model specification,  $Q_E$ , which evaluates whether significant systematic variation remains unexplained in the regression model.

As a supplementary analysis, we attained homogeneity among the effect sizes by identifying outliers and sequentially removing those that

<sup>11</sup> The between-measures correlations that we used when implementing this formula were estimated from correlations and coefficient alphas given in a number of studies in the sample. The resulting correlations were .56 for subjective performance, .38 for overall performance (combined over objective and subjective performance), and .46 for satisfaction with leader. In addition, .36 was estimated for aggregating the performance and satisfaction measures into study-level effect sizes.



reduced the homogeneity statistic by the largest amount (see Hedges & Olkin, 1985). Studies yielding effect sizes identified as outliers were then examined to determine whether they appeared to differ methodologically from the other studies. In addition, the central tendency of the effect sizes was recomputed after removal of the outliers.

## Results

### Characteristics of the Studies

Before presenting the sex differences reported in the research on leadership effectiveness, we display the characteristics of the studies in our sample so that our meta-analytic findings can be interpreted in relation to the studies' typical methodologies. As shown by the central tendencies of the characteristics listed in Table 1, studies typically (a) were published relatively recently, (b) were published as journal articles, (c) were authored by men and women in equal numbers, (d) based the statistical analysis on a moderate number of observations, and (e) aggregated a modest number of judgments into each data point. The comparisons of female and male leaders were sometimes confounded with other variables such as age and job seniority; in the majority of the studies, some confounding was likely (although unknown) or known to have occurred. Although a minority of the studies were carried out in experimental laboratories on small groups, the majority were carried out in organizations, most of which were dedicated to business or education and were either large or of mixed or unknown size.

The leaders or managers assessed in the studies were most commonly first-level (i.e., line) leaders (or supervisors), although in substantial numbers of studies leadership was either at a middle level or an ambiguous, mixed, or unknown level. In the organizational studies compared with the laboratory experiments, the leaders were older, the leadership roles were more heavily male dominated, and the subordinates were more likely to be female. The methods by which the leaders were selected varied considerably, with the most common methods being successful or unsuccessful random samples in the organizational studies and random appointments to leadership in the laboratory experiments.

The means for the last five characteristics represent the variables constructed from our questionnaire respondents' judgments of the leadership roles examined in the studies. As shown by these means, self-reported competence to carry out these roles and interest in occupying them were equal for female and male respondents. However, the average man was judged to be more interested in occupying the roles than the average woman. Also, the roles were thought to require "quite a lot" of both interpersonal and task ability.<sup>12</sup>

### Overall Sex Differences in Leader Effectiveness

The summary of the study-level effect sizes given in Table 2 allows one to determine whether, on the whole, male and female leaders differed in effectiveness. In this table (and subsequent tables), an overall sex difference is suggested by a mean effect size that differed significantly from the no-difference value of 0.00 (as indicated by a confidence interval that does not include 0.00). The mean weighted effect size averaged across all 76 studies that yielded study-level effect sizes did not differ from 0.00. The other estimates of central tendency, the mean unweighted

Table 2  
Summary of Study-Level Effect Sizes

Criterion	Value
Known effect sizes	
<i>n</i>	76
Mean weighted <i>d</i> ( <i>d</i> <sub>+</sub> ) <sup>a</sup>	-0.02
95% CI for <i>d</i> <sub>+</sub>	-0.05/0.02
Homogeneity ( <i>Q</i> ) of <i>d</i> s comprising <i>d</i> <sub>+</sub> <sup>b</sup>	269.85***
Mean unweighted <i>d</i>	-0.03
95% CI for mean unweighted <i>d</i>	-0.13/0.07
Median effect size	-0.07
Known effect sizes excluding outliers	
No. removed outliers	12 (.16)
<i>n</i>	64
Mean weighted <i>d</i> ( <i>d</i> <sub>+</sub> )	-0.12
95% CI for <i>d</i> <sub>+</sub>	-0.16/-0.08
Homogeneity ( <i>Q</i> ) of <i>d</i> s comprising <i>d</i> <sub>+</sub>	86.30

Note. Effect sizes are positive for greater effectiveness of male leaders and negative for greater effectiveness of female leaders. *n* = sample size; CI = confidence interval; *d* = effect size; *d*<sub>+</sub> = mean weighted effect size; *Q* = homogeneity of effect sizes.

<sup>a</sup> Effect sizes were weighted by the reciprocal of the variance. <sup>b</sup> Significance indicates rejection of the hypothesis of homogeneity.

\*\*\* *p* < .001.

effect size and the median effect size, were also very close to 0.00.<sup>13</sup>

As shown by the homogeneity statistic given in Table 2, the hypothesis of homogeneity was rejected. Removal of a moderate number of outliers produced a homogeneous set of effect sizes with a weighted mean indicating that female leaders were slightly more effective than male leaders.<sup>14</sup> Consistent with the shift of the mean effect size in favor of women when outliers

<sup>12</sup> The mean ratings on these 15-point scales fell in the range anchored by the term *quite a lot*. Suggesting that our student respondents were able to discriminate between the interpersonal and task requirements of leadership roles, mean ratings showed that some roles were thought to require more interpersonal than task ability (e.g., school administrators and leaders of laboratory discussion groups) and others were thought to require more task than interpersonal ability (e.g., military officers). On all of these indexes of gender congeniality, a few organizational roles were particularly masculine (e.g., military leaders, athletic coaches), and a few were particularly feminine (e.g., administrative roles in elementary schools).

<sup>13</sup> Not considered in these estimates of central tendency were 20 non-significant comparisons that could not be represented as effect sizes because of a lack of sufficient information. The null results reported in Table 2 would merely remain null if these comparisons were given the value of 0.00 (indicating exactly no sex difference) and included in the aggregation.

<sup>14</sup> The removed studies, listed in the order in which they were removed, are the following: Mohr et al. (1978), ROTC cadets attending a summer camp; Yoder, Adams, and Hicks (1986), squad leaders for cadets at U.S. Military Academy at West Point; Caccese and Mayerberg (1984), head coaches at colleges (all sports, in larger colleges and universities); Rice, Yoder, Adams, Priest, and Prince (1984), first-year cadets at U.S. Military Academy at West Point; Deaux (1979), Sample 2, supervisors in California branches of a national retail chain store; Mohr and Downey (1977), newly commissioned U.S. Army lieutenants at-

Table 3  
*Categorical Model for Identity of Raters*

Identity of raters	Between-classes effect ( $Q_B$ )	$n$	Mean weighted effect size ( $d_{i+}$ )	95% CI for $d_{i+}$		Homogeneity within each class ( $Q_{wi}$ ) <sup>a</sup>
				Lower	Upper	
	104.13***					
Leaders		34	0.14 <sub>a</sub>	0.08	0.20	198.62***
Supervisors of leaders		16	0.07 <sub>a</sub>	0.01	0.13	89.58***
Subordinates of leaders		40	-0.12 <sub>b</sub>	-0.16	-0.07	50.67
Peers of leaders		7	0.25 <sub>a</sub>	0.15	0.35	85.81***
Judges		11	-0.19 <sub>b</sub>	-0.30	-0.08	19.82*
Objective counting		13	-0.07 <sub>ab</sub>	-0.23	0.09	9.36
Mixed or unclear		4	0.31 <sub>a</sub>	0.17	0.46	21.17***

*Note.* Effect sizes are positive for greater effectiveness of male leaders and negative for greater effectiveness of female leaders. Some studies are represented by more than one effect size because the effect sizes were not aggregated over identities of raters and types of measure in order to represent each study only once (see *Computation of effect sizes* in Method). Differences between means that do not have a subscript in common are significant at the .05 level or beyond. CI = confidence interval.

<sup>a</sup> Significance indicates rejection of the hypothesis of homogeneity.

\*  $p < .05$ . \*\*\*  $p < .001$ .

were removed, all of the 12 outlying effect sizes favored male leaders over female leaders. Moreover, six of these outlying studies examined military officers or cadets, and two of the other outlying studies examined leaders in traditionally masculine roles (coaches, Outward Bound instructors).

Among the 82 studies that allowed the direction of the difference to be discerned, 35 (or 43%) favored the male leaders over the female leaders. Because this proportion did not differ from .50 (the proportion expected under the null hypothesis), this measure of central tendency, like the others we have presented, suggests that women and men did not differ in general in their effectiveness as leaders.

#### *Tests of Categorical Models for Identity of Raters and Type of Measure*

To help interpret the aggregated study-level effect sizes for sex differences in effectiveness, we sought to determine whether these differences varied according to (a) the identity of the raters who provided the effectiveness measure and (b) the type of measure. The analyses, shown in Table 3 and Table 4, present categorical models that examined between-classes effects for identity of raters and type of measure.

The categorical model for identity of raters proved significant (see Table 3). Judges as well as subordinates of leaders favored women in their ratings, whereas four other categories of raters

(namely, leaders, supervisors of leaders, peers of leaders, and mixed or unclear) favored men. However, interpretation of these differences is hampered by the fact that the classes of raters producing the most divergent mean effect sizes (peers of leaders and mixed or unclear) contained relatively few effect sizes. Having examined variation in the effect sizes due to the identities of the raters, we then aggregated them over these identities, within each type of measure.

The classification of the resulting effect sizes by type of measure appears in Table 4. The first categorical model, which examined the various types of subjective performance measures (see Method for details), suggested that men fared well relative to women on measures of leaders' ability, in comparison with the other types of measures. In general, as shown by the second model, studies' findings did not vary with the subjective versus objective nature of the performance measures. The third model, which compared performance and satisfaction measures, was significant and showed that women fared well compared with men when satisfaction with leaders was assessed, whereas men fared slightly better than women when performance measures were used. However, these models are difficult to interpret because they were substantially changed by the removal of outlying effect sizes from those classes that were not already homogeneous.<sup>15</sup>

Having examined the impact of type of measure, we then aggregated the effect sizes over type of measure to produce study-level effect sizes. Consequently, in the remaining analyses (as in the overall aggregations shown in Table 2) each study was represented by one effect size to meet the independence criterion of meta-analytic statistics.

tending a training course; Instone, Major, and Bunker (1983), student supervisors of laboratory work groups participating in a simulation of a production task; Quinn (1977), elementary school principals; Yoder and Adams (1984), graduates of U.S. Military Academy at West Point one year after graduation in the first class that admitted women; Rice, Bender, and Vitters (1980), first-year cadets at U.S. Military Academy at West Point participating as group leaders in a laboratory experiment; Riggins (1985), instructors and assistant instructors at Colorado Outward Bound school; H. W. Smith and George (1980), Study 2, group leaders for laboratory task involving writing titles for stories for a woman's magazine.

<sup>15</sup> Although the first model, for type of subjective measure of performance, remained significant, the mean effect size for measures of ability no longer differed from 0.00. Instead, men fared well on measures of motivation, whereas women fared well on measures of effectiveness and on measures in the other and mixed category. The third model (the comparison of performance and satisfaction measures), which was also significant before the removal of outliers, became nonsignificant after the removal.

### Impact of Moderating Variables on the Relative Effectiveness of Female and Male Leaders

Numerous study attributes were examined as predictors of the study-level effect sizes, with the outliers included in the data set. Models that were especially interesting or that produced significant results appear in Table 5 and Table 6. Models were also calculated on various subgroups of effect sizes, including organizational studies only and performance measures only. Because these additional analyses were for the most part quite similar to those computed on the entire group of effect sizes, they are not reported here.

**Setting of study.** The first model in Table 5 shows that sex differences in leaders' effectiveness did not vary depending on whether the study's setting was an organization or an experimental laboratory. However, the particular type of organization did relate to the effect sizes, as indicated by the second model. Consistent with the analysis of outliers presented in the preceding subsection, the military studies deviated strongly from all other classes of studies. Men fared significantly better than women in the military studies, and the mean effect size for these studies differed significantly from the means for the other types of organizations and the laboratory groups. The mean effect sizes for two classes of studies (education, and government or social service) significantly favored female leaders, and the mean effect size for one class (business) marginally favored female leaders.

**Level of leadership.** Classification of the studies by the hierarchical level of leadership also produced a significant model. As shown in Table 5, effectiveness comparisons favored men for first-level or line leadership, favored women for second or middle-level leadership, and favored neither sex in the remaining category of studies, for which the level of leadership was ambiguous, mixed, or unknown. That these trends were not an artifact of any confounding between the military studies and level of leadership was shown by the similar results we obtained

with the military studies excluded from the analysis; the model remained significant despite this exclusion.

**Percentage of men among leaders and subordinates and congeniality of leadership roles for men and women.** As Table 6 indicates, the effect sizes related positively to the percentage of men among the leaders and the percentage of men among the subordinates. These findings mean that comparisons on leader effectiveness favored men more and women less to the extent that the leadership role was male-dominated and that the subordinates were male.

The remaining continuous models in Table 6 predicted sex differences in leaders' effectiveness from our questionnaire measures of the gender congeniality of the leadership roles. In general, male leaders fared well in roles thought to be congenial to men, and female leaders fared well in roles thought to be congenial to women. Specifically, effectiveness comparisons favored male leaders over female leaders to the extent that (a) male (compared with female) respondents rated themselves as more competent in the role, (b) male respondents rated themselves as more interested in occupying the role, (c) respondents of both sexes judged the average man more interested in occupying the role than the average woman, and (d) respondents of both sexes judged that the role required relatively low levels of interpersonal ability and relatively high levels of task ability. Analogously, effectiveness comparisons favored female leaders over male leaders to the extent that these conditions were reversed.

The findings presented in Table 6 help clarify the sharp divergence of the military studies from other classes of studies (see Table 5). Indeed, the studies classified as military differed from the other studies in the sample on all seven of the variables shown in Table 6. Thus, the military studies examined leadership roles with larger percentages of men among the leaders and the subordinates than the other classes of studies ( $ps < .001$ ). In addition, the military roles were more congenial for men (relative to women) than the other roles, as assessed on the five ques-

Table 4  
Categorical Models for Type of Measure

Variable and class	Between-classes effect ( $Q_b$ )	$n$	Mean weighted effect size ( $d_{i+}$ )	95% CI for $d_{i+}$		Homogeneity within each class ( $Q_{wi}$ ) <sup>a</sup>
				Lower	Upper	
Type of subjective measure of performance	50.99***					
Effectiveness		36	-0.01 <sub>a</sub>	-0.06	0.04	114.43***
Ability		22	0.28 <sub>b</sub>	0.22	0.34	145.47***
Motivation		8	0.01 <sub>a</sub>	-0.08	0.10	34.53***
Other and mixed		7	0.08 <sub>a</sub>	-0.03	0.18	33.84***
Subjective versus objective measures of performance	0.69					
Subjective		57	0.05	0.01	0.09	230.13***
Objective		14	-0.02	-0.17	0.14	19.14
Performance versus satisfaction measures	31.39***					
Performance		65	0.05	0.01	0.09	237.30***
Satisfaction		17	-0.16	-0.21	-0.10	23.53

*Note.* Effect sizes are positive for greater effectiveness of male leaders and negative for greater effectiveness of female leaders. Some studies are represented by more than one effect size because the effect sizes were not aggregated over types of measures in order to represent each study only once (see *Computation of effect sizes* in Method). CI = confidence interval. Differences between means that do not have a subscript in common are significant at the .05 level or beyond.

<sup>a</sup> Significance indicated rejection of the hypothesis of homogeneity.

\*\*\*  $p < .001$ .

Table 5  
Categorical Models Predicting Study-Level Effect Sizes

Variable and class	Between-classes effect ( $Q_B$ )	$n$	Mean weighted effect size ( $d_{i+}$ )	95% CI for $d_{i+}$		Homogeneity within each class ( $Q_{wi}$ ) <sup>a</sup>
				Lower	Upper	
Setting of study	1.81					
Organization		56	-0.03	-0.06	0.01	235.05***
Laboratory group		20	0.07	-0.06	0.20	32.99**
Type of organization	92.92***					
Business		15	-0.07 <sub>a</sub>	-0.14	0.00	41.88***
Education		16	-0.11 <sub>a</sub>	-0.18	-0.04	61.29***
Government or social service		6	-0.15 <sub>a</sub>	-0.25	-0.05	6.77
Military		7	0.42 <sub>b</sub>	0.32	0.52	15.58*
Miscellaneous <sup>b</sup>		10	-0.05 <sub>a</sub>	-0.16	0.05	9.84
Not organizational (laboratory group)		20	0.07 <sub>a</sub>	-0.06	0.20	32.99*
Level of leadership	67.44***					
First or line		35	0.19 <sub>a</sub>	0.13	0.26	115.05***
Second or middle <sup>c</sup>		20	-0.18 <sub>b</sub>	-0.24	-0.12	42.11**
Ambiguous, mixed, or unknown		21	-0.03 <sub>c</sub>	-0.09	0.02	45.23**

Note. Effect sizes are positive for greater effectiveness of male leaders and negative for greater effectiveness of female leaders. Differences between means that do not have a subscript in common are significant at the .05 level or beyond. CI = confidence interval.

<sup>a</sup> Significance indicates rejection of the hypothesis of homogeneity. <sup>b</sup> Includes hospitals and clinics, recreational camps, and mixed settings.

<sup>c</sup> Category includes one study with higher-level managers, described as "executives."

\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

tionnaire-derived indexes of gender-relevant aspects of the leadership roles ( $ps < .01$  or smaller). Specifically, the tendency for the male (vs. female) questionnaire respondents to rate themselves relatively more competent and interested was greater for the military roles than the other roles, and the tendency for the respondents to rate men as more interested than women was

also greater for the military roles. In addition, the respondents rated the military roles as requiring less interpersonal ability and more task ability than the other leadership roles.

The findings shown in Table 6 are also relevant to the effects of level of leadership (see Table 5). The first-level (or line) leadership roles, in which men were more effective than women, differed from the second-level (or middle management) roles, in which women were more effective than men, on all seven of the variables shown in Table 6. Although the second-level roles were more male-dominated than the first-level roles ( $p < .001$ ), they had a marginally smaller proportion of male subordinates ( $p < .11$ ). In addition, the first-level roles were more congenial for men (relative to women) than the second-level roles, as assessed on four of the five indexes of the gender-relevant aspects of the leadership roles ( $ps < .05$  or smaller). The exception was the ratings of the extent to which roles required task ability: Respondents rated the second-level roles as requiring more task ability ( $p < .025$ ) and more interpersonal ability ( $p < .0001$ ) than the first-level roles.

*Characteristics of studies' methods.* Also of interest was whether studies in which leaders' sex was known to be confounded with personal attributes such as age and job seniority (or was likely to have been so confounded) would favor men more than other studies did, especially for studies with an organizational setting. No such trend appeared; in fact, studies in which confounding was likely (but unknown) favored women significantly more than other classes of studies.

The relation of the effect sizes to the sex of the raters who produced the measure of leader effectiveness was also examined. The effect sizes that were calculated separately for male raters and female raters (see Method) did not differ, nor did either of these classes of effect sizes differ from the effect sizes

Table 6  
Continuous Models Predicting Study-Level Effect Sizes

Predictor	$b$	$\beta$
Percentage of men among leaders ( $n = 76$ )	0.00** <sup>a</sup>	0.19
Percentage of men among subordinates ( $n = 48$ )	0.00** <sup>b</sup>	0.22
Respondent judgments of roles ( $n = 70$ )		
Competence sex difference <sup>c</sup>	0.40***	0.52
Interest sex difference	0.29***	0.37
Stereotypic interest difference	0.30***	0.54
Interpersonal ability rating <sup>d</sup>	-0.09**	0.20
Task ability rating	0.11***	0.28

Note. Models are weighted least squares simple linear regressions calculated with weights equal to the reciprocal of the variance for each effect size. Effect sizes are positive for greater effectiveness of male leaders and negative for greater effectiveness of female leaders. Because of missing data,  $n$  varied across the analyses.  $b$  = unstandardized regression coefficient,  $\beta$  = standardized regression coefficient.

<sup>a</sup>  $b = 0.0027$ ,  $SE(b) = 0.00091$ . <sup>b</sup>  $b = 0.0022$ ,  $SE(b) = 0.00098$ . <sup>c</sup> For the first three variables constructed from judgments of the leadership roles, values are positive for differences in the masculine direction (greater male estimates of competence and of interest; ascription of greater interest to average men). <sup>d</sup> For the last two variables constructed from judgments of the leadership roles, values are larger to the extent that a role was judged to require more interpersonal or task ability.

\*  $p < .05$ . \*\*  $p < .01$ . \*\*\*  $p < .001$ .

that were based on both sexes of raters. However, because these separate effect sizes could be calculated for only 10 studies, the results of this analysis may well be unreliable. Therefore, we predicted the entire group of study-level effect sizes from the proportion of each study's raters who were male (the median proportion was .50). This analysis yielded a significant tendency for studies with higher proportions of men among the raters to produce effect sizes that favored male leaders ( $b = .0023$ ;  $\beta = .23$ ;  $p < .001$ ).<sup>16</sup>

The level of data aggregation (i.e., the number of observations aggregated into each data point) produced a significant model. The effectiveness comparisons tended to favor women over men to the extent that a greater number of judgments (i.e., items) were aggregated into each of the observations underlying the effect sizes ( $b = -0.0025$ ;  $\beta = -.23$ ;  $p < .001$ ).

*Other characteristics.* Finally, analyses treating other study characteristics as independent variables either did not yield significant models or were uninterpretable because of missing data, small sample sizes, or confounding with other variables. Among the characteristics producing null findings were the sex of the authors and publication year, attributes that in other meta-analyses on gender-relevant hypotheses have sometimes related to effect sizes (e.g., Eagly & Johnson, 1990; Eagly & Karau, 1991; Thomas & French, 1985; Wood, 1987).

*Multiple regression models.* Given that the study characteristics were not independent of one another, we estimated multiple regression models that examined the simultaneous impact of several of the continuous and categorical variables. The relatively large number of categorical and continuous variables that produced significant one-way models restricted our efforts to test multiple regression models because the number of potential predictors was quite large in relation to the number of effect sizes. Moreover, information was incomplete on some of our predictors, requiring classifications of studies as unknown or unclear in relation to some variables.

One of the most informative models entered as predictors the percentage of men among the leaders, gender congeniality (a composite of competence sex difference, interest sex difference, and stereotypic interest difference), interpersonal ability rating, task ability rating, and military setting (vs. all other organizational and laboratory settings, dummy coded). The significant predictors in this model were gender congeniality,  $p < .05$ , and military setting,  $p < .001$ . As reflected in its multiple correlation of .62, this model was moderately effective in accounting for variability in the magnitude of the effect sizes, although the test of model specification,  $Q_E$ , showed that it cannot be regarded as correctly specified.

## Discussion

When all of the studies in our sample were aggregated, female and male leaders did not differ in effectiveness. The various measures of central tendency showed no overall sex difference. This finding is important in applied terms because it suggests that, despite barriers and possible handicaps in functioning as leaders, the women who actually serve as leaders and managers are in general succeeding as well as their male counterparts. Even if female leaders do behave somewhat differently than male leaders, they appear to be equally effective. On a theoretical basis, this null outcome is not surprising in view of the con-

flicting predictions that arose from the several perspectives we reviewed. Indeed, if there are some reasons to think that female leaders may excel and other reasons to think that male leaders may excel, neither sex may possess an overall advantage in effectiveness.

### *Conditions That Favor Male Leaders Versus Conditions That Favor Female Leaders*

Despite our overall finding that women and men were equivalent in their effectiveness as leaders, this generalization is not appropriate in all, or even most, of the settings in which investigators have examined leadership. Very often the sex of the leader or manager does make a difference. The need for identifying conditions that favored one sex over the other was initially revealed by the inconsistency of the findings across the studies—that is, by the effect sizes' heterogeneity, indicating that they did not share a common mean. Still, to place the findings' inconsistency in context, readers should note that this synthesis did not produce especially inconsistent findings in comparison with other quantitative syntheses on psychological topics. The finding that homogeneity was obtained by removing 16% of the effect sizes, a percentage well in line with other meta-analyses (Hedges, 1987), suggests that findings in this area were no more inconsistent than those in typical research areas in psychology. Nevertheless, the homogeneity analysis called for more detailed scrutiny of the studies' findings, which did indeed reveal some conditions under which men fared better than women and others under which women fared better than men.

*Gender-congeniality of leadership roles.* Our efforts to specify the conditions that were associated with advantage for each sex produced the most striking finding of our meta-analysis—that leadership roles defined in relatively masculine terms favored male leaders and that roles defined in relatively feminine terms favored female leaders. Specifically, sex differences in leaders' effectiveness were significantly correlated with the congeniality of their roles for men or women, as indexed by our questionnaire respondents' judgments of competence and interest in relation to the roles (see Table 6). Male leaders tended to be more effective than female leaders to the extent that a leadership role was more congenial to men; female leaders tended to be more effective than male leaders to the extent that a leadership role was more congenial to women. Furthermore, women were more effective than men in leadership roles that were feminine in the sense that our respondents judged they required considerable interpersonal ability, defined as the ability to cooperate and get along with other people. Men were more effective than women in roles that were masculine in the sense that our respondents judged they required considerable task ability, defined as the ability to direct and control people. In general, leaders of each sex were particularly effective when they were in a

<sup>16</sup> This analysis raises the issue of whether the identity of the raters (i.e., as leaders, supervisors, subordinates, etc.) might be related to their sex distribution, possibly accounting for the findings shown in Table 3. Indeed, the proportion of raters who were male did vary with raters' identities, with the proportion of men being highest among the supervisors and lowest among the subordinates. Because only this contrast was significant post hoc, it is unlikely that the sex distribution of the raters provides a complete account of the findings shown in Table 3.

leadership role regarded as congruent with their gender. These findings suggest that being "out of role" in gender-defined terms may produce a decline in leaders' actual or perceived effectiveness.

The extent to which leader and subordinate roles were male dominated numerically also related to sex differences in effectiveness. Specifically, male leaders tended to be more effective than female leaders to the extent that the role was especially male dominated and associated with male subordinates. Moreover, to the extent that leaders' subordinates were male, the role was rated as more congenial to men on our questionnaire-derived indexes of gender congeniality,  $ps < .001$ . However, the male dominance of the leadership role did not relate significantly to the gender congeniality of the role in our sample of leadership studies. These findings relating the congeniality of the leadership roles to the male dominance of leader and subordinate roles thus provided some additional support for the idea that gender congruence shapes leaders' effectiveness.

*Type of organization.* Also very informative about the conditions that favor leaders of one sex over those of the other sex is our classification of organizational studies according to type of organization. This analysis showed that findings from one type of organization, namely, the military, differed from those of all other types of organizations. In fact, only military organizations yielded findings that significantly favored male leaders. Although this result needs to be interpreted cautiously, given the relatively small number of military studies in our sample, it provides additional evidence of the utility of our gender-role congruence interpretation.

Both gender congeniality and prejudicial attitudes may contribute to men's judged effectiveness in military settings. The role of military officer (and cadet) has been defined in exceedingly masculine terms, as social scientists have argued (e.g., Arkin & Dobrofsky, 1978). Our synthesis provided empirical evidence of this masculine definition of military roles by showing that the military studies differed from the other studies on all of our measures of the gender congeniality of the leadership roles. Military roles were thus shown to be particularly masculine on several indexes as well as especially male dominated on a numerical basis. Other research has consistently shown that male military officers and cadets generally hold hostile and sexist attitudes toward female officers and cadets, although there is also some evidence of growing tolerance (Stevens & Gardner, 1987; P. J. Thomas, Perry, & David, 1994). Prejudiced attitudes would surely set up difficult conditions for female leaders.

The special status of the military studies was also indicated by the results of our outlier analysis, which identified the studies that differed most strongly from the overall mean effect size. This analysis revealed that all of the outlying studies produced sex differences that strongly favored men. Many of these deviating studies were conducted on military personnel, and some on leaders in other male-dominated roles (e.g., head coaches at colleges).

Our analysis also identified types of organizations that favored female leaders more than male leaders. Although no type of organization produced a very substantial advantage for women (parallel to the substantial advantage that men possessed in the military studies), several types of organizations produced weak tendencies for women to be more effective than men, namely, business, education, and government or social ser-

vice. Because these studies outnumbered military studies, women's small advantage in these organizations outweighed men's larger advantage in military organizations, to produce our null overall effect size. It should be noted that our findings did not differ between the several types of organizations that slightly favored women, despite many obvious differences in such organizations' goals and structures (e.g., differences between organizations in business and education).

*Level of leadership.* Our analysis also established that the hierarchical level of leadership related to sex differences in leaders' effectiveness. Men fared better than women in first-level (or line) positions; women fared better than men in second-level (or middle management) positions despite the tendency for these roles to be more male dominated numerically. These interesting findings also reflected the gender congeniality of leadership roles, with first-level positions being perceived as relatively congenial to men and second-level positions as relatively congenial to women. Relevant to interpreting these findings is research suggesting that the skills required to fulfill managerial roles vary with the hierarchical level of the role (e.g., Alexander, 1979; Pavett & Lau, 1983). Most provocative from the perspective of this research synthesis is Paolillo's (1981) finding that lower-level management ordinarily requires technical skills, whereas middle-level management places a relatively greater burden on human relations skills that allow a manager to build cooperative effort and to motivate and develop subordinates. Consistent with research suggesting that women are, on the average, more socially skilled than men (see Eagly & Wood, 1991), women may tend to be particularly suited to middle managerial roles. An alternate explanation for our findings for level of leadership is that the prejudice against female bosses (e.g., Gallup, 1990) may pertain especially to relationships of direct supervision because role conflict may be more acute when gender roles and superior-subordinate roles are in conflict in a face-to-face relationship. These accounts, of course, remain speculative because the relation obtained in this synthesis between the hierarchical level of leadership roles and sex differences in effectiveness is correlational and thus warrants further scrutiny in primary research. Moreover, our findings do not allow any extrapolation to hierarchical levels higher than middle management. As more women gain access to the upper levels of management, appropriate research would allow the effectiveness of female and male executives to be compared.

### *Ambiguity of the Criterion in Studies of Leaders' Effectiveness*

This synthesis of empirical literature provides some information concerning the methods that researchers have typically used to assess the effectiveness of leaders. As we indicated early in this article, there is no single superior method of assessing effectiveness, and multiple measures are frequently advocated. Consistent with this view, our classification of the types of measures used in effectiveness research revealed diversity in methods. However, subjective measures far outnumbered objective measures and, among subjective measures, ratings of leaders' performance outnumbered ratings of subordinates' personal satisfaction with leaders. This favoring of subjective ratings of leaders' performance raises questions of validity because such judgments do not provide pure measures of leaders' actual per-

formance. Indeed, subjective performance ratings' vulnerability to bias is well documented (e.g., Landy & Farr, 1980).

To shed some light on how these assessment issues may have influenced the findings of our synthesis, we first classified our data according to the identity of the raters (e.g., supervisors and subordinates), but this analysis proved inconclusive. That this variable failed to produce a readily interpretable pattern of findings is compatible with Landy and Farr's (1980) conclusion that research examining raters' identity has not produced consistent effects.

More provocative in terms of showing the possible impact of raters' characteristics is our finding that the sex of the raters was correlated with studies' outcomes. The higher the proportion of men among the raters, the more raters' effectiveness ratings favored men over women. This finding is entirely consistent with Eagly et al.'s (1992) synthesis of experimental studies of evaluation of leaders. This prior meta-analysis thus showed that female leaders were more strongly devalued relative to their (experimentally equated) male counterparts when their evaluators were men rather than women. Whereas female raters did not favor one sex over the other, male raters significantly favored men over equivalent women. If these earlier findings are taken into account, a reasonable interpretation of the tendency for male raters to favor men in leader effectiveness studies is that male raters manifested some promale or antifemale bias.

Our analyses classifying the effect sizes by type of measure in general revealed an absence of strong effects of this variable. Although these analyses suggested that men were especially favored on subjective ratings of ability and that women fared especially well on subordinates' ratings of their satisfaction with their leader, these trends failed to survive the removal of outlying effect sizes. These analyses thus remained somewhat inconclusive.

A more critical concern for the interpretation of our meta-analysis is the extent to which our major findings represent raters' gender bias rather than genuine differences in effectiveness. Does the tendency for each sex to fare better than the other in relatively gender-congruent settings reflect a tendency for raters to give prejudiced reactions to men or women who are out of role and thus to rate their performance on the basis of gender congruence rather than actual effectiveness? Unfortunately our meta-analytic data are not very informative about this matter, although they did show that the aggregated findings based on objective measures of effectiveness such as group productivity did not differ from the findings based on the subjective measures. Separate analyses predicting the subjective and objective performance measures from study characteristics might have shed more light on these issues, but unfortunately too few objective measures were available to allow them to be analyzed separately (and their values were homogeneous). Yet, the tendency for our gender congeniality effects to reflect gender bias at least to some extent (rather than true differences in effectiveness) would be consistent with Eagly et al.'s (1992) finding that men were rated more favorably than women in male-dominated leadership roles, even though the studies in that literature had experimentally equated the portrayal of leadership by women and men.

### *Reflections on Theoretical Analyses*

Although our review gives priority to the descriptive purpose of determining the conditions under which female and male

leaders differ (and do not differ) in effectiveness, the findings are provocative in relation to our theoretical analysis of gender's impact on effectiveness. As explained early in the article, a simple version of social-role theory suggests that men are more effective than women because the definition of leadership roles in terms of male-stereotypic qualities places female leaders in a situation of role conflict. Another simple principle, the strict structural analysis in terms of organizational hierarchy, predicts that women and men occupying the same leadership role do not differ in effectiveness because leadership roles override gender roles. In contrast, a third simple principle, the differential selection analysis, predicts that female leaders are more successful than male leaders because they have to be more competent than their male counterparts to attain leadership roles.

Based on the overall aggregation of the effect sizes, which produced no overall sex difference in effectiveness, some readers might be tempted to conclude in favor of the structural approach. Alternatively, other readers might be tempted to conclude from the null finding that the selection principle, which should produce advantage for female leaders, counteracted the gender-role principle, which should produce advantage for male leaders. These conclusions, plausible in terms of our overall aggregated finding, are decidedly less attractive in view of the various analyses predicting the effect sizes from characteristics of the studies. These additional analyses suggest that gender does matter, with leaders achieving greater effectiveness in gender-congruent environments. These findings thus favor the elaborated version of social-role theory that we discussed early in this article. Moreover, the findings suggest that leadership or managerial roles may be defined in a more masculine or feminine fashion, depending on the organizational context of management. It is the fit between leaders' gender and the specifics of the leadership role that influences leaders' effectiveness.

The theoretical principles that fare less well in view of the totality of our findings are the structural perspective and the selection-bias perspective, as long as they are confined to the simple principles that we introduced early in this article. From a strictly structural viewpoint, there should be no sex differences in leaders' effectiveness, neither overall nor in specific categories of studies. However, male and female leaders' effectiveness did differ under the conditions that we identified. Yet, proponents of the structural perspective might elaborate it by arguing that women and men occupying the same leadership or managerial role would indeed differ in effectiveness if they differed in their status, power, or experience, even though their role was formally equivalent. In this synthesis of leadership effectiveness studies, we attempted to take account of this consideration by coding the extent to which leaders' sex and other attributes were confounded (see Method), but this variable did not produce results suggesting that likely forms of confounding (e.g., female managers being younger and less experienced than male managers) produced advantages for men. Nonetheless, the confounding hypothesis, which has enjoyed some support in primary research (Liden, 1985), could be elaborated to incorporate some features of social-role theory if, for example, probable restrictions on female leaders' autocratic and directive displays of power were taken into account. Such elaborations would enable the perspective to encompass at least some aspects of the findings we have presented.

From the selection-bias perspective, which argues for female

leaders' greater competence because of their more stringent selection, women should fare better than men overall and be particularly successful to the extent that they are rarer in the leadership role and presumably selected by standards that are especially stringent. However, our findings showed the opposite tendency—men were more effective than women in particularly male-dominated roles. Nonetheless, a selection-bias theorist might argue that the very factors that produced the glass ceiling in the first place, such as negative expectancies about women's competence, would also operate to compromise the effectiveness of the women who successfully broke through that ceiling, nullifying the gains from their intrinsically greater competence. In addition, the contemporary presence of affirmative action policies favoring women in some contexts could weaken women's self-confidence and performance (e.g., Heilman, Rivero, & Brett, 1991), although such outcomes are surely not inevitable (Graves & Powell, 1994). More generally, an elaborated selection-bias argument could produce more subtle predictions, some of which would incorporate features of social-role theory.

Contingency theories of leadership do not provide much guidance for interpreting our findings because they emphasize moderating conditions that could not be assessed in the present data set (e.g., quality of leader-member relations) and predictions that are more molecular than those we investigated. Nonetheless, some comments are in order in view of Eagly and Johnson's (1990) finding that in general female leaders adopted a more democratic and participative style than male leaders, even in organizational settings. Could this sex difference in leaders' style be responsible for some of the relationships we observed? Possibly a democratic style is nonoptimal (and surely nontraditional) in military organizations but more acceptable in other settings, particularly in educational organizations and government and social service organizations, where women fared best relative to men. Yet Eagly and Johnson also showed that to the extent that a role was highly male dominated, the tendency weakened for women (vs. men) to be more democratic and participative. Thus, women in military settings may not adopt the democratic leadership style that is more typical of women than men.

Complicating this discussion of leadership style are additional findings of Eagly and Johnson's (1990) meta-analysis on sex differences in leadership style. This earlier synthesis found that leaders who were out of role in gender terms (according to the same questionnaire-derived gender-congeniality measures we reported in Table 6) were less task oriented in their leadership style. Men and women emphasized task accomplishment when occupying a leadership role that was regarded as congruent with their own gender or that was numerically dominated by leaders of their own sex. In general, these findings paralleled the effectiveness findings produced in the present synthesis. Leaders of each sex were more effective to the extent they were in a leadership role regarded as congruent with their own gender or that was numerically dominated by their own sex. The mediation of our effectiveness findings may thus be that it is difficult for leaders to be effectively task oriented (or be perceived as task oriented) when they are in a domain that is more congenial to the other gender than to their own gender. Leaders and managers who are out of role may tend to have some difficulty in comfortably guiding and directing the task-relevant activities of their subordinates; they may be somewhat deficient in the knowledge

and authority required to direct subordinates to attain task-relevant goals and, as a consequence, may suffer some loss of confidence. Lowered effectiveness may follow mainly from this deficit in out-of-role leaders' ability to guide groups and organizations toward their task-relevant goals.

### *Conclusion*

Our findings are provocative in view of women's increasing occupancy of managerial roles and other leadership roles. In the aggregate, this trend may have little impact on organizations' success in view of our finding that female leaders and managers are no more or less effective than male leaders and managers. Still, some gains should be realized merely from selecting managers from a pool of candidates that would be substantially enlarged through the inclusion of women, enabling organizations to implement more stringent selection criteria. Any fears that organizations would be in peril because they allow women to take the reins of leadership are unjustified from the perspective of our synthesis.

Our findings should not, however, be taken as evidence for equality of performance in most settings, given our evidence that female and male leaders *are* differentially effective in many settings. Of course, even in these settings, factors other than gender are doubtless extremely important in determining effectiveness—for example, a leader's skill and training, and the match between the situation and the leader's style. Nonetheless, women fared poorly in settings in which leadership was defined in highly masculine terms, especially in military settings. Men fared slightly worse than women in settings in which leadership was defined in less masculine terms, especially in educational organizations and in governmental and social service organizations. Although these findings remain modest in size, they suggest a pervasive gendering of leadership roles that can operate to the disadvantage of women or men. Were gender entirely unimportant in organizations and groups, it would be irrelevant to leaders' effectiveness, and men and women would fare equally well as leaders throughout all types of organizations and groups. Our project instead suggests that gender role expectations spill over onto leadership roles within organizations and groups and produce important consequences for the effectiveness of leaders. The mechanisms by which these consequences are produced deserve careful scrutiny in primary research. Knowledge of mechanisms would be especially important for women who lead in settings defined in relatively masculine terms and for men who lead in settings defined in relatively less masculine (or more feminine) terms because such individuals may be at risk for diminished effectiveness on the basis of their gender.

If the trend of the past decades continues so that more women enter managerial roles, including at the highest levels, the shaping of people's expectations according to the gender of leaders and managers should weaken. Indeed, research suggests that some weakening has already occurred in the traditional tendency to define leadership and management in masculine terms (e.g., Brenner et al., 1989; Frank, 1988; Russell et al., 1988), and this trend may continue as women gain access to higher-level management. Thus, the entry of a substantial proportion of women into managerial roles may itself change the perception of these roles in an androgynous direction. Alternatively, for reasons unrelated to women's participation in management,



many organizations may be changing to favor styles of management that are less autocratic and more participative, producing managerial roles that tend to be more congenial to women (see Offerman & Gowing, 1990). Given these trends and our finding that women tended to be even slightly more effective than men in roles defined in less masculine terms, there is reason to expect that women will fare at least as well as men in an increasing range of leadership and managerial roles.

### References

References marked with an asterisk indicate studies included in the meta-analysis.

- \*Adams, E. F. (1978). A multivariate study of subordinate perceptions of and attitudes toward minority and majority managers. *Journal of Applied Psychology*, 63, 277-288.
- \*Adams, J., Rice, R. W., & Instone, D. (1984). Follower attitudes toward women and judgments concerning performance by female and male leaders. *Academy of Management Journal*, 27, 636-643.
- 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*, 186-189.
- \*Archambeau, A. A. (1980). Sex-role flexibility as a predictor of subordinate evaluations of managers and subordinate job satisfaction (Doctoral dissertation, University of South Florida, 1980). *Dissertation Abstracts International*, 41, 1550B.
- Arkin, W., & Dobrofsky, L. R. (1978). Military socialization and masculinity. *Journal of Social Issues*, 34(1), 151-168.
- \*Arnett, M. D., Higgins, R. B., & Priem, A. P. (1980). Sex and least preferred co-worker score effects in leadership behavior. *Sex Roles*, 6, 139-152.
- Arons, E. L. (1980). Male and female administrative potential—Is there a difference? *National Association of Secondary School Principals Bulletin*, 64, 5-8.
- \*Baril, G. L., Elbert, N., Mahar-Potter, S., & Reavy, G. C. (1989). Are androgynous managers really more effective? *Group & Organization Studies*, 14, 234-249.
- \*Barker, C. J. F. (1983). The effect of selected demographic variables on Hollander's theory of idiosyncrasy credit (Doctoral dissertation, University of Maryland, 1982). *Dissertation Abstracts International*, 44, 1937A.
- \*Bartol, K. M. (1973). *Male and female leaders in small work groups*. East Lansing, MI: MSU Business Studies.
- Bartol, K. M., & Martin, D. C. (1986). Women and men in task groups. In R. D. Ashmore & F. K. Del Boca (Eds.), *The social psychology of female-male relations* (pp. 259-310). San Diego, CA: Academic Press.
- \*Bartol, K. M., & Wortman, M. S., Jr. (1975). Male versus female leaders: Effects on perceived leader behavior and satisfaction in a hospital. *Personnel Psychology*, 28, 533-547.
- Bass, B. M. (1990). *Bass & Stogdill's handbook of leadership: Theory, research, and managerial applications* (3rd ed.). New York: Free Press.
- Bayes, M., & Newton, P. M. (1978). Women in authority: A sociopsychological analysis. *Journal of Applied Behavioral Science*, 14, 7-20.
- \*Beard, J. E. (1989). A study of leadership styles and their relationship to management effectiveness among middle school principals in the state of Maryland. *Dissertation Abstracts International*, 49, 2460A. (University Microfilms No. 88-15, 563)
- Bowman, G. W., Worthy, N. B., & Greyser, S. A. (1965). Are women executives people? *Harvard Business Review*, 43(4), 14-28, 164-178.
- 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.
- Bruning, N. S., & Snyder, R. A. (1981). Predictions of work-related competence: Are they different for men vs. women? *American Institute for Decision Sciences Proceedings*, 1, 350-352.
- \*Caccese, T. M., & Mayerberg, C. K. (1984). Gender differences in perceived burnout of college coaches. *Journal of Sport Psychology*, 6, 279-288.
- Camden, C., & Witt, J. (1983). Manager communicative style and productivity: A study of female and male managers. *International Journal of Women's Studies*, 6, 258-269.
- Chemers, M. M. (1987). Leadership processes: Intrapersonal, interpersonal, and societal influences. In C. Hendrick (Ed.), *Review of personality and social psychology: Vol. 8. Group processes* (pp. 252-277). Newbury Park, CA: Sage.
- \*Couch, M. E. (1980). Self-assessment of leadership skills by 4-H volunteer leaders in the southern region of the United States (Doctoral dissertation, Texas Tech University, 1980). *Dissertation Abstracts International*, 41, 1327B.
- Craig, J. M., & Jacobs, R. R. (1985). The effect of working with women on male attitudes toward female firefighters. *Basic and Applied Social Psychology*, 6, 61-74.
- Davenport, I. W. (1977). Analysis of the perceived leader behavior of male and female elementary school principals (Doctoral dissertation, University of Missouri-Columbia, 1976). *Dissertation Abstracts International*, 37, 5476A.
- \*Day, D. R., & Stogdill, R. M. (1972). Leader behavior of male and female supervisors: A comparative study. *Personnel Psychology*, 25, 353-360.
- \*Deaux, K. (1979). Self-evaluations of male and female managers. *Sex Roles*, 5, 571-580.
- Deaux, K., & Lewis, L. L. (1984). Structure of gender stereotypes: Interrelationships among components and gender label. *Journal of Personality and Social Psychology*, 46, 991-1004.
- Denmark, F. L. (1993). Women, leadership, and empowerment. *Psychology of Women Quarterly*, 17, 343-356.
- Dobbins, G. H., & Platz, S. J. (1986). Sex differences in leadership: How real are they? *Academy of Management Review*, 11, 118-127.
- Drenth, P. J. D., & Koopman, P. L. (1984). A contingency approach to participative leadership: How good? In J. G. Hunt, D. Hosking, C. A. Schriesheim, & R. Stewart (Eds.), *Leaders and managers: International perspectives on managerial behavior and leadership* (pp. 303-315). New York: Pergamon Press.
- 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. (1991). Gender and the emergence of leaders: A meta-analysis. *Journal of Personality and Social Psychology*, 60, 685-710.
- Eagly, A. H., Karau, S. J., Miner, J. B., & Johnson, B. T. (1994). Gender and motivation to manage in hierarchic organizations: A meta-analysis. *Leadership Quarterly*, 5, 135-159.
- 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. (1991). Explaining sex differences in social behavior: A meta-analytic perspective. *Personality and Social Psychology Bulletin*, 17, 306-315.
- \*Eskilson, A. (1975). Sex composition and leadership in small groups (Doctoral dissertation, University of Illinois, 1974). *Dissertation Abstracts International*, 35, 5694B.
- \*Ezell, H. F., Odewahn, C. A., & Sherman, J. D. (1980). Perceived competence of women managers in public human service organizations: A comparative view. *Journal of Management*, 6, 135-144.

- \*Farokhi, F. R. (1987). Teachers' satisfaction with their principal and selected teachers' and principals' characteristics. *Dissertation Abstracts International*, 47, 3255A. (University Microfilms No. 86-28, 873)
- Fiedler, F. E. (1967). *A theory of leadership effectiveness*. New York: McGraw Hill.
- Fiedler, F. E., & Chemers, M. M. (1984). *Improving leadership effectiveness: The Leader Match concept* (2nd ed.). New York: Wiley.
- \*Francke, C. A. (1975). Perceived performance differences between women and men supervisors and implications for training (Doctoral dissertation, Michigan State University, 1975). *Dissertation Abstracts International*, 36, 1349A.
- Frank, E. J. (1988). Business students' perceptions of women in management. *Sex Roles*, 19, 107-118.
- \*Friedman, M. J. (1981). Differences in assessment center performance as a function of the race and sex of rates and the race of assessors (Doctoral dissertation, University of Tennessee at Knoxville, 1980). *Dissertation Abstracts International*, 41, 3924B.
- Frone, M. R., Adams, J., Rice, R. W., & Instone-Noonan, D. (1986). Halo error: A field study comparison of self- and subordinate evaluations of leadership process and leader effectiveness. *Personality and Social Psychology Bulletin*, 12, 454-461.
- \*Frosch, C. S. (1980). An analysis of the background profiles and self-assessed effectiveness of Oklahoma public secondary school administrators by race and gender (Doctoral dissertation, University of Oklahoma, 1980). *Dissertation Abstracts International*, 41, 2542A.
- \*Gallegos, P. I. (1988). Emerging leadership among Hispanic women: The role of expressive behavior and nonverbal skill (Doctoral dissertation, University of California, Riverside, 1987). *Dissertation Abstracts International*, 48, 3152B.
- Gallup, G. H. (1990). *The Gallup poll*. Wilmington, DE: Scholarly Resources.
- 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.
- Ghiselli, E. E. (1964). *Theory of psychological measurement*. New York: McGraw Hill.
- Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-analysis in social research*. Newbury Park, CA: Sage.
- \*Goktepe, J. R. (1986). The role of gender, androgyny and attraction in predicting the identity and effectiveness of emergent leaders. *Dissertation Abstracts International*, 47, 2229A. (University Microfilms No. 86-20, 779)
- \*Goktepe, J. R., & Schneider, C. E. (1988). Sex and gender effects in evaluating emergent leaders in small groups. *Sex Roles*, 19, 29-35.
- \*Goodman, M. S. (1981). A survey of the relationship of selected situational and personal variables to the performance of physical education chairpersons in institutions of higher education. *Dissertation Abstracts International*, 41, 2854A. (University Microfilms No. 80-27, 442)
- Graves, L. M., & Powell, G. N. (1994). Effects of sex-based preferential selection and discrimination on job attitudes. *Human Relations*, 47, 133-157.
- \*Gross, N., & Trask, A. E. (1976). *The sex factor and the management of schools*. New York: Wiley.
- \*Gupta, N., Jenkins, G. D., Jr., & Beehr, T. A. (1983). Employee gender, gender similarity, and supervisor-subordinate cross-evaluations. *Psychology of Women Quarterly*, 8, 174-184.
- Guterk, B. A. (1993). Changing the status of women in management. *Applied Psychology: An International Review*, 42, 301-311.
- Guterk, B. A., & Cohen, A. G. (1987). Sex ratios, sex role spillover, and sex at work: A comparison of men's and women's experiences. *Human Relations*, 40, 97-115.
- Guterk, B. A., & Morasch, B. (1982). Sex-ratios, sex-role spillover, and sexual harassment of women at work. *Journal of Social Issues*, 38(4), 55-74.
- \*Hamilton, E. E. (1983). Androgyny and leadership: An empirical field study of effective influence in decision-making groups. *Dissertation Abstracts International*, 43, 2379B. (University Microfilms No. 82-23, 210)
- Hatch, D. D. (1987). The relationship between self-monitoring and management performance in males and females in a field setting (Doctoral dissertation, Brigham Young University, 1987). *Dissertation Abstracts International*, 48, 917B.
- Hedges, L. V. (1987). How hard is hard science, how soft is soft science? The empirical cumulativeness of research. *American Psychologist*, 42, 443-455.
- Hedges, L. V., & Becker, B. J. (1986). Statistical methods in the meta-analysis of research on gender differences. In J. Hyde & M. C. Linn (Eds.), *The psychology of gender: Advances through meta-analysis* (pp. 14-50). Baltimore: Johns Hopkins University Press.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.
- Heilman, M. E., Block, C. J., Martell, R. F., & Simon, M. C. (1989). Has anything changed? Current characterizations of men, women, and managers. *Journal of Applied Psychology*, 74, 935-942.
- Heilman, M. E., Rivero, J. C., & Brett, J. F. (1991). Skirting the competence issue: Effects of sex-based preferential selection on task choices of women and men. *Journal of Applied Psychology*, 76, 99-105.
- Heilman, M. E., Simon, M. C., & Repper, D. P. (1987). Intentionally favored, unintentionally harmed? Impact of sex-based preferential selection on self-perceptions and self-evaluations. *Journal of Applied Psychology*, 72, 62-68.
- \*Hemphill, J. K., Griffiths, D. E., & Frederiksen, N. (1962). *Administrative performance and personality: A study of the principal in a simulated elementary school*. New York: Teachers College, Columbia University.
- Hollander, E. P. (1992). The essential interdependence of leadership and followership. *Current Directions in Psychological Science*, 1, 71-75.
- \*Holsclaw, C. A. (1967). The relationship of selected factors to administrative effectiveness in the principalship (Doctoral dissertation, University of Tennessee, 1967). *Dissertation Abstracts International*, 28, 2027A.
- \*Horgan, D. D., & Simeon, R. J. (1988). *Gender, mentoring, and tacit knowledge* (Report No. CG-021658). Memphis, TN: Memphis State University. (ERIC Document Reproduction Service No. ED 306 524)
- House, R. J. (1971). A path goal theory of leader effectiveness. *Administrative Science Quarterly*, 16, 321-338.
- House, R. J. (1991). The distribution and exercise of power in complex organizations: A meso theory. *Leadership Quarterly*, 2, 23-58.
- House, R. J., & Mitchell, T. R. (1974). Path goal theory of leadership. *Journal of Contemporary Business*, 3(4), 81-98.
- \*Hoyle, J. R., & Randall, R. S. (1967). Problem-attack behavior and sex, prior teaching experience, and college preparation of elementary school principals. *Journal of Experimental Education*, 36, 28-33.
- \*Hughes, H., Jr., Copeland, D. R., Ford, L. H., & Heidt, E. A. (1983). *Leadership and management education and training (LMET) course requirements for recruit company commanders and "A" school instructors* (Tech. Rep. No. 154, Report No. AD-A137306). Orlando, FL: Department of the Navy. (ERIC Document Reproduction Service No. ED 237 777)
- Hunt, J. G. (1991). *Leadership: A new synthesis*. Newbury Park, CA: Sage.
- \*Instone, D., Major, B., & Bunker, B. B. (1983). Gender, self-confidence, and social influence strategies: An organizational simulation. *Journal of Personality and Social Psychology*, 44, 322-333.
- \*Jacobs, J. A. (1992). Women's entry into management: Trends in earnings, authority, and values among salaried managers. *Administrative Science Quarterly*, 37, 282-301.
- \*Jacobson, M. B., & Effertz, J. (1974). Sex roles and leadership: Perceptions of the leaders 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, B. T. (1989). *DSTAT: Software for the meta-analytic review of research literatures*. Hillsdale, NJ: Erlbaum.
- Kanter, R. M. (1977). *Men and women of the corporation*. New York: Basic Books.
- Kenrick, D. T., & Trost, M. R. (1993). The evolutionary perspective. In A. E. Beall & R. J. Sternberg (Eds.), *The psychology of gender* (pp. 148-172). New York: Guilford Press.
- \*King, P. J. (1978). An analysis of teachers' perceptions of the leadership styles and effectiveness of male and female elementary school principals (Doctoral dissertation, University of Southern California, 1978). *Dissertation Abstracts International, 39*, 2658A.
- \*Kipnis, D., & Schmidt, S. M. (1988). Upward-influence styles: Relationship with performance evaluations, salary, and stress. *Administrative Science Quarterly, 33*, 528-542.
- Klawitter, P. A. (1985). *The relationship between principal's leadership style and teacher job satisfaction*. Unpublished doctoral dissertation, West Virginia University.
- \*Koberg, C. S. (1985). Sex and situational influences on the use of power: A follow-up study. *Sex Roles, 13*, 625-639.
- Kruse, L., & Wintermantel, M. (1986). Leadership Ms.-qualified: I. The gender bias in everyday and scientific thinking. In C. F. Graumann & S. Moscovici (Eds.), *Changing conceptions of leadership* (pp. 171-197). New York: Springer-Verlag.
- Landy, F. J., & Farr, J. L. (1980). Performance rating. *Psychological Bulletin, 87*, 72-107.
- Lee, D. M., & Alvares, 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.
- \*LeGrand, A. A. (1982). *Leadership style and effectiveness in summer camps* (Doctoral dissertation, University of Southern California, 1982). *Dissertation Abstracts International, 42*, 5055A.
- Lenny, E. (1977). Women's self-confidence in achievement settings. *Psychological Bulletin, 84*, 1-13.
- \*Liden, R. C. (1985). Female perceptions of female and male managerial behavior. *Sex Roles, 12*, 421-432.
- \*Lindholm, H. R. (1989). The effects of race and gender on supervisor-subordinate relationships: An integrative dyadic perspective. *Dissertation Abstracts International, 50*, 1142B. (University Microfilms No. 89-15, 276)
- Maccoby, E. E. (1990). Gender and relationships: A developmental account. *American Psychologist, 45*, 513-520.
- \*Maier, N. R. F. (1970). Male versus female discussion leaders. *Personnel Psychology, 23*, 455-461.
- \*Maier, N. R. F., & Sashkin, M. (1971). The contributions of a union steward vs. a time-study man in introducing change: Role and sex effects. *Personnel Psychology, 24*, 221-238.
- Malloy, T. E., & Janowski, C. L. (1992). Perceptions and metaperceptions of leadership: Components, accuracy, and dispositional correlates. *Personality and Social Psychology Bulletin, 18*, 700-708.
- \*Martin, C. R., Jr. (1972). Support for women's lib: Management performance. *Southern Journal of Business, 7*, 17-28.
- Martin, P. Y. (1992). Gender, interaction, and inequality in organizations. In C. L. Ridgeway (Ed.), *Gender, interaction, and inequality* (pp. 208-231). New York: Springer-Verlag.
- Maslach, C., & Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Occupational Behavior, 2*, 99-113.
- McMahon-Dumas, C. E. (1981). An investigation of the leadership styles and effectiveness dimensions of principals, and their relationship with reading gain scores of students in the Washington, DC, public schools. *Dissertation Abstracts International, 42*, 1881A. (University Microfilms No. 81-21, 448)
- McNemar, Q. (1962). *Psychological statistics* (3rd ed.). New York: Wiley.
- \*Millard, R. J., & Smith, K. H. (1985). Moderating effects of leader sex on the relation between leadership style and perceived behavior patterns. *Genetic, Social and General Psychology Monographs, 111*, 305-316.
- Miller, D. T., & Turnbull, W. (1986). Expectancies and interpersonal processes. *Annual Review of Psychology, 37*, 233-256.
- \*Moehlman, G. M. (1989). The relationship between leadership styles and faculty perceived effectiveness of principals in secondary vocational-technical schools in Michigan (Doctoral dissertation, Ohio State University, 1988). *Dissertation Abstracts International, 49*, 3344A.
- \*Mohr, E. S., & Downey, R. G. (1977). Are women peers? *Journal of Occupational Psychology, 50*, 53-57.
- \*Mohr, E. S., Rowan, G. P., & Reidy, R. F. (1978). *Women and ROTC summer camp, 1975* (U.S. Army Tech. Rep. No. 293, Report No. AD-A064553). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. (ERIC Document Reproduction Service No. ED 169 241)
- 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. (1987). *Breaking the glass ceiling: Can women reach the top of America's largest corporations?* Reading, MA: Addison-Wesley.
- Morrison, T. L., & Stein, D. D. (1985). Member reaction to male and female leaders in two types of group experience. *Journal of Social Psychology, 125*, 7-16.
- Moses, J. L. (1973). The development of an assessment center for the early identification of supervisory potential. *Personnel Psychology, 26*, 569-580.
- Moses, J. L., & Boehm, V. R. (1975). Relationship of assessment center performance to management progress of women. *Journal of Applied Psychology, 60*, 527-529.
- \*Muldrow, T. W., & Bayton, J. A. (1979). Men and women executives and processes related to decision accuracy. *Journal of Applied Psychology, 64*, 99-106.
- Mumford, M. D. (1986). Leadership in the organizational context: A conceptual approach and its applications. *Journal of Applied Social Psychology, 16*, 508-531.
- \*Munson, C. E. (1979). Evaluation of male and female supervisors. *Social Work, 24*, 104-110.
- Nieva, V. F., & Gutek, B. A. (1981). *Women and work: A psychological perspective*. New York: Praeger.
- \*Nygust, 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.
- Offerman, L. R., & Gowing, M. K. (1990). Organizations of the future: Changes and challenges. *American Psychologist, 45*, 95-108.
- O'Leary, V. E. (1974). Some attitudinal barriers to occupational aspirations in women. *Psychological Bulletin, 81*, 809-826.
- \*O'Leary, V. E., & Ickovics, J. R. (1987, April). *Who wants a woman boss? Only those who have one*. Paper presented at the annual meeting of the Eastern Psychological Association, Arlington, VA.
- \*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.
- Paolillo, J. G. P. (1981). Manager's self assessments of managerial roles: The influence of hierarchical level. *Journal of Management, 7*, 43-52.
- \*Pate, T. M., Jr. (1979). The relationship of the elementary principal's leadership style to his self-concept of his ability to communicate and his knowledge of his job (Doctoral dissertation, George Peabody College for Teachers, 1978). *Dissertation Abstracts International, 39*, 6445A.
- \*Patrick, J. F. (1978). The relationship between sex of supervisor, attitudes toward women, and the job satisfaction of female employees

- (Doctoral dissertation, University of Maryland, 1978). *Dissertation Abstracts International*, 39, 1028B.
- Pavett, C. M., & Lau, A. W. (1983). Managerial work: The influence of hierarchical level and functional specialty. *Academy of Management Journal*, 26, 170-177.
- \*Peregrine, P. E. (1983). The robustness of managerial life in schools: A test of contingency theory. *Dissertation Abstracts International*, 43, 2849A. (University Microfilms No. 83-00, 320)
- Perritt, L. J. C. (1983). Job satisfaction as a function of worker and supervisor characteristics (Doctoral dissertation, University of Kentucky, 1983). *Dissertation Abstracts International*, 44, 2002B.
- \*Peters, L. H., O'Connor, E. J., Weekley, J., Pooyan, A., Frank, B., & Erenkrantz, B. (1984). Sex bias and managerial evaluations: A replication and extension. *Journal of Applied Psychology*, 69, 349-352.
- \*Petty, M. M., & Bruning, N. S. (1980). A comparison of the relationships between subordinates' perceptions of supervisory behavior and measures of subordinates' job satisfaction for male and female leaders. *Academy of Management Journal*, 23, 717-725.
- \*Petty, M. M., & Lee, G. K., Jr. (1975). Moderating effects of sex of supervisor and subordinate on relationships between supervisory behavior and subordinate satisfaction. *Journal of Applied Psychology*, 60, 624-628.
- \*Petty, M. M., & Miles, R. H. (1976). Leader sex-role stereotyping in a female-dominated work culture. *Personnel Psychology*, 29, 393-404.
- \*Petty, M. M., Odewahn, C. A., Bruning, N. S., & Thomason, T. L. (1977). An examination of the moderating effects of supervisor sex and subordinate sex upon the relationships between supervisory behavior and subordinate outcomes in mental health organizations. *Academy of Management Proceedings*, 37, 408-412.
- Phillips, J. S., & Lord, R. G. (1982). Schematic information processing and perceptions of leadership in problem-solving groups. *Journal of Applied Psychology*, 67, 486-492.
- Powell, G. N. (1993). *Women & men in management* (2nd ed.). Newbury Park, CA: Sage.
- 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.
- \*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.
- \*Quinn, K. I. (1977). Self-perceptions of leadership behaviors and decision-making orientations of men and women elementary school principals in Chicago public schools (Doctoral dissertation, University of Illinois at Urbana-Champaign, 1976). *Dissertation Abstracts International*, 37, 6199A.
- \*Ragins, B. R. (1989). Power and gender congruency effects in evaluations of male and female managers. *Journal of Management*, 15, 65-76.
- Ragins, B. R., & Sundstrom, E. (1989). Gender and power in organizations: A longitudinal perspective. *Psychological Bulletin*, 105, 51-88.
- \*Renwick, P. A. (1977). The effects of sex differences on the perception and management of superior-subordinate conflict: An exploratory study. *Organizational Behavior and Human Performance*, 19, 403-415.
- \*Rice, R. W., Bender, L. R., & Vitters, A. G. (1980). Leader sex, follower attitudes toward women, and leadership effectiveness: A laboratory experiment. *Organizational Behavior and Human Performance*, 25, 46-78.
- \*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.
- \*Rice, R. W., Yoder, J. D., Adams, J., Priest, R. F., & Prince, H. T., II. (1984). Leadership ratings for male and female military cadets. *Sex Roles*, 10, 885-901.
- Riger, S., & Galligan, P. (1980). Women in management: An exploration of competing paradigms. *American Psychologist*, 35, 902-910.
- \*Riggins, R. D. (1985). An analysis of selected biographical and personality characteristics contributing to the leadership effectiveness of instructors and assistant instructors at the Colorado Outward Bound School: 1982-1983 (Doctoral dissertation, Indiana University, 1983). *Dissertation Abstracts International*, 45, 3744A.
- Rosen, B., & Jerdee, T. H. (1973). The influence of sex-role stereotypes on evaluations of male and female supervisory behavior. *Journal of Applied Psychology*, 57, 44-48.
- Rosenthal, R., & Rubin, D. B. (1986). Meta-analytic procedures for combining studies with multiple effect sizes. *Psychological Bulletin*, 99, 400-406.
- Rubner, M. B. (1991). More workers prefer a man in charge. *American Demographics*, 13, 11.
- Russell, J. E., Rush, M. C., & Herd, A. M. (1988). An exploration of women's expectations of effective male and female leadership. *Sex Roles*, 18, 279-287.
- \*Sadeghi-Nejad, M. M. (1989). The professional development of experienced school administrators. *Dissertation Abstracts International*, 49, 3581A. (University Microfilms No. 89-04, 213)
- Schein, V. E. (1973). The relationship between sex role stereotypes and requisite management characteristics. *Journal of Applied Psychology*, 57, 95-100.
- Schmidt, F. L., Ones, D. S., & Hunter, J. E. (1992). Personnel selection. *Annual Review of Psychology*, 43, 627-670.
- \*Schneier, C. E. (1978). The contingency model of leadership: An extension to emergent leadership and leader's sex. *Organizational Behavior and Human Performance*, 21, 220-239.
- \*Schneier, C. E., & Bartol, K. M. (1980). Sex effects in emergent leadership. *Journal of Applied Psychology*, 65, 341-345.
- \*Schulman, J. M. (1989). Power behavior, job satisfaction and leadership effectiveness in public school principals. *Dissertation Abstracts International*, 50, 1518A. (University Microfilms No. 89-21, 859)
- Sheppard, D. (1992). Women managers' perceptions of gender and organizational life. In A. J. Mills & P. Tancred (Eds.), *Gendering organizational analysis* (pp. 151-166). Newbury Park, CA: Sage.
- Simon, R. J., & Landis, J. M. (1989). Women's and men's attitudes about a woman's place and role. *Public Opinion Quarterly*, 53, 265-276.
- \*Simpson, M. W. (1985). Teacher satisfaction with leadership based on principal sex role orientation (Doctoral dissertation, University of Georgia, 1985). *Dissertation Abstracts International*, 46, 1474A.
- Smith, A. C., Jr. (1982). Correlates of first-line supervisory effectiveness (Doctoral dissertation, Virginia Polytechnic Institute and State University, 1981). *Dissertation Abstracts International*, 42, 4429A.
- \*Smith, H. W., & George, C. E. (1980). Gender effects in two gender-oriented appointed leader situations. *JSAS: Catalog of Selected Documents in Psychology*, 10, 58. (Ms. No. 2072)
- Smith, P. C., Kendall, L. M., & Hulin, C. L. (1969). *The measurement of satisfaction in work and retirement: A strategy for the study of attitudes*. Chicago: Rand McNally.
- \*Snodgrass, S. E., & Rosenthal, R. (1984). Females in charge: Effects of sex of subordinate and romantic attachment status upon self-ratings of dominance. *Journal of Personality*, 52, 355-371.
- Stake, J. E. (1981). Promoting leadership behaviors in low performance-self-esteem women in task-oriented mixed-sex dyads. *Journal of Personality*, 49, 401-414.
- \*Steckler, N. A., & Rosenthal, R. (1985). Sex differences in nonverbal and verbal communication with bosses, peers, and subordinates. *Journal of Applied Psychology*, 70, 157-163.
- Stevens, G., & Gardner, S. (1987). But can she command a ship? Acceptance of women by peers at the Coast Guard Academy. *Sex Roles*, 16, 181-188.
- Sutton, C. D., & Moore, K. K. (1985). Executive women—20 years later. *Harvard Business Review*, 63, 43-66.
- \*Szilagyi, A. D. (1980). Reward behavior by male and female leaders: A causal inference analysis. *Journal of Vocational Behavior*, 16, 59-72.

- \*Terborg, J. R., & Shingledecker, P. (1983). Employee reactions to supervision and work evaluation as a function of subordinate and manager sex. *Sex Roles, 9*, 813-824.
- Thomas, J. R., & French, K. E. (1985). Gender differences across age in motor performance: A meta-analysis. *Psychological Bulletin, 98*, 260-282.
- Thomas, P. J., Perry, Z. A., & David, K. M. (1994). *Fitness reports of naval warfare officers: A search for gender differences* (Rep. No. NPRDC-TR-94-10). San Diego, CA: Navy Personnel Research and Development Center.
- Tsui, A. S. (1984). A multiple-constituency framework of managerial reputational effectiveness. In J. G. Hunt, D. Hosking, C. A. Schriesheim, & R. Stewart (Eds.), *Leaders and managers: International perspectives on managerial behavior and leadership* (pp. 28-44). New York: Pergamon Press.
- \*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., & 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.
- Uleman, J. S. (1991). Leadership ratings: Toward focusing more on specific behaviors. *Leadership Quarterly, 2*, 175-187.
- Vroom, V. H., & Jago, A. G. (1988). *The new leadership: Managing participation in organizations*. Englewood Cliffs, NJ: Prentice Hall.
- Vroom, V. H., & Yetton, P. W. (1973). *Leadership and decision-making*. Pittsburgh, PA: University of Pittsburgh Press.
- \*Waetjen, W. B., Schuerger, J. M., & Schwartz, E. B. (1979). Male and female managers: Self-concept, success, and failure. *Journal of Psychology, 103*, 87-94.
- Wagner, R. K., & Sternberg, R. J. (1985). Practical intelligence in real-world pursuits: The role of tacit knowledge. *Journal of Personality and Social Psychology, 49*, 436-458.
- Wahrman, R., & Pugh, M. D. (1974). Sex, nonconformity and influence. *Sociometry, 37*, 137-147.
- \*Wakefield, D. G. (1984). *Sex and "power communication" in middle- and top-level administration* (Report No. CS-209909). Warrensburg: Central Missouri State University. (ERIC Document Reproduction Service No. ED 259 357)
- \*Wexley, K. N., & Pulakos, E. D. (1982). Sex effects on performance ratings in manager-subordinate dyads: A field study. *Journal of Applied Psychology, 67*, 433-439.
- Williams, C. L. (1989). *Gender differences at work: Women and men in nontraditional occupations*. Berkeley, CA: University of California Press.
- Wood, W. (1987). Meta-analytic review of sex differences in group performance. *Psychological Bulletin, 102*, 53-71.
- \*Yoder, J. D., & Adams, J. (1984). Women entering nontraditional roles: When work demands and sex-roles conflict: The case of West Point. *International Journal of Women's Studies, 7*, 260-272.
- \*Yoder, J. D., Adams, J., & Hicks, J. M. (1986). *When sex-roles and work roles conflict: A critical look at standards of evaluation*. Unpublished manuscript, Webster University.
- \*Young, J. C. (1981). *Comparisons of leader style, behaviors and effectiveness of male and female coaches* (Report No. SP-022744). College Park, MD: University of Maryland. (ERIC Document Reproduction Service No. ED 231 830)
- 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. 147-197). Palo Alto, CA: Consulting Psychologists Press.

Received March 16, 1994

Revision received August 8, 1994

Accepted August 18, 1994 ■

### New Editors Appointed, 1996-2001

The Publications and Communications Board of the American Psychological Association announces the appointment of three new editors for 6-year terms beginning in 1996. As of January 1, 1995, manuscripts should be directed as follows:

- For *Behavioral Neuroscience*, submit manuscripts to Michela Gallagher, PhD, Department of Psychology, Davie Hall, CB# 3270, University of North Carolina, Chapel Hill, NC 27599.
- For the *Journal of Experimental Psychology: General*, submit manuscripts to Nora S. Newcombe, PhD, Department of Psychology, Temple University, 565 Weiss Hall, Philadelphia, PA 19122.
- For the *Journal of Experimental Psychology: Learning, Memory, and Cognition*, submit manuscripts to James H. Neely, PhD, Editor, Department of Psychology, State University of New York at Albany, 1400 Washington Avenue, Albany, NY 12222.

Manuscript submission patterns make the precise date of completion of 1995 volumes uncertain. The current editors, Larry R. Squire, PhD, Earl Hunt, PhD, and Keith Rayner, PhD, respectively, will receive and consider manuscripts until December 31, 1994. Should either volume be completed before that date, manuscripts will be redirected to the new editors for consideration in 1996 volumes.