

Sex Effects in Evaluating Leaders: A Replication Study

Dafna N. Izraeli

Department of Sociology and Anthropology
Bar-Ilan University

Dove Izraeli

Faculty of Management
Tel Aviv University

This replication of the Bartol and Butterfield (1976) study examines the hypothesis that sex role stereotypes bias the performance evaluation of managerial behavior. Respondents, Israeli managers ($n = 1,020$), evaluated the performance of a manager as described in four vignettes, each representing a different managerial style. Sex of manager was manipulated by change of name. Contrary to the findings of the original study, multivariate analyses of variance revealed no overall main sex of manager effects. The differences are explained in terms of dissimilarities in sample, dependent variable measure, and culture. The need for better controls for comparative purposes is emphasized. In addition senior managers rated the female manager more favorably than lower level managers in three of the four leadership styles.

Bias against women in performance appraisal has been suggested as an explanation for their paucity in management. Most published studies support this view, but the evidence is not consistent. (For a review of this literature see Nieva & Gutek, 1980.) Studies by Bartol and Butterfield (1976) and Dipboye and Wiley (1977) yield different results on the effects of sex on performance evaluation of managers who use different behavioral styles. For example, Bartol and Butterfield (1976) found that subjects, when asked to rate the performance of a fictitious manager described in a short story as using an initiation-of-structure leadership style, rated a male manager higher than a female manager. When the manager was portrayed in another story as using a considerate leadership style, they rated a female manager higher than a male manager. These results support their hypothesis that sex role stereotypes influence the evaluation of leadership behavior. Initiating structure, a task-oriented behavior style, is deemed more appropriate for men than for women; consideration, a relationship-oriented behavior, is considered more appropriate for

women than for men. It appears that persons of each sex were rated on the basis of the congruence of his or her behavior with sex role expectations.

Jago and Vroom (1982) found that women, but not men, are penalized for autocratic behavior. They asked 68 male and 22 female managers to judge the participativeness of five to eight other members of mixed-sex training groups and to describe their affective response to each person. Males and females perceived to be participative were rated equally favorably. However, females perceived to be autocratic were negatively evaluated, whereas males perceived to be autocratic received modest but nonetheless positive evaluations.

The findings of these and other studies (Petty & Lee, 1975; Petty & Miles 1976; Haccoun, Haccoun, & Sallay, 1978) suggest an interaction between subject sex and leadership style in performance evaluation. Sex role expectations penetrate the permeable boundaries that separate the work place from the wider society and influence normative perceptions of behavior. Different criteria for evaluating that behavior are then used depending on the sex of the person.

Dipboye and Wiley (1977) similarly hypothesized that sex role stereotypes influence the evaluation of performance. College recruiters rated two candidates for a supervisory position, based on written resumes (which portrayed the candidates as highly qualified

This study was supported by a grant from the Ford Foundation and from the Israel Institute of Business Research at Tel Aviv University, Faculty of Management.

Requests for reprints should be sent to Dafna N. Izraeli, Department of Sociology, Bar-Ilan University, Ramat Gan, 52100 Israel.

for the position) and video presentations of interviews. They hypothesized that a female candidate using a passive interview style would be rated higher than a male candidate using the same style, whereas a female candidate using a moderately aggressive interview style would be rated lower than a man using the same style. Their hypotheses were derived from earlier studies, which show that passivity is judged to be more appropriate in women than men and aggressive behavior more appropriate for men than for women. Contrary to their predictions, the male and the female candidates were rated equally low when they used the passive style and equally high when they used the moderately aggressive style. No significant interactions were found.

Stitt, Schmidt, Price, and Kipnis (1983), like Jago and Vroom (1982), studied subordinate satisfaction and assessed productivity under varying conditions of leadership style (democratic-autocratic) and sex of leader. They found no sex effects among a sample of 678 male and female students. These findings support an alternative theory, namely that there are occupational subcultures in which norms reflecting occupational specific expectations determine how behavior will be evaluated. Within each occupational context certain behavioral styles are considered more effective than others (Powell & Butterfield, 1979; Stitt et al. 1983). Success for both men and women requires their adopting those styles rather than others. The compelling impact of the managerial context may explain why so few sex differences in attitudes and behavior have been found among practicing managers, especially when level in the hierarchy is controlled (for reviews, see Birsdall, 1980; Chapman, 1975; Donnell & Hall, 1980).

The Bartol and Butterfield (1976) results may have been influenced by the fact that the subjects, largely undergraduate students, were less responsive to cues about performance competence and more affected by sex role stereotypes than practicing managers would be. The present study is a replication based on a sample of practicing managers. It addresses the same problem that Bartol and Butterfield raised—do stereotypes about appropriate behavior of men and women affect managerial performance evaluation? Are dif-

ferent criteria used to rate the effectiveness of women and men in management?

We also sought to find out whether level of education and organizational rank of respondents influence performance evaluation and whether these variables influence the evaluation of the male and female managers in the same way. Terborg, Peters, Ilgen, and Smith (1977), in one of the only studies of the organizational and personal correlates of attitudes toward women as managers, found education to be a significant predictor of these attitudes. Organizational rank, however, as measured by hourly or salary pay classification and pay level, did not consistently relate to attitudes toward women as managers, although the direction found was that managers had more positive attitudes than lower level participants.

Method

Subjects

This study analyzes data from a larger study of 920 men and 100 women in 44 different management training courses given in various parts of Israel. Over 92% of the respondents were managers: 53 were junior managers, 310 were middle managers, and 436 were senior managers. A post facto comparison of the distribution of sample characteristics with those of the population of managers in Israel shows the sample to be representative in its sex, ethnic, and educational composition but somewhat younger in age. Forty-nine percent were between 18 and 34 years old; 48.7% were between 35 and 54, and 2.5% were aged 55+. Fifty-two percent had 12 years or less of education, 35.4% had 13 to 15 years, and 22.6% 16 and more years of education. The women in the sample (and in the population as well) were on the average younger and more educated than the men. The distribution among employing organizations is as follows: business, 61.3%; government, 5.6%; defense, 12.7%; nonprofit, 4.2%; cooperative (kibbutz), 12.8%; not currently employed, 3.3%. The sample had an over representation of kibbutz members, and their scores were consequently weighted (reduced by 50%) to reflect their proportion in the population.

Instruments

The instruments used were those developed by Bartol and Butterfield (1976) and reported in their study. They consisted of four short stories, each reflecting a different leadership style: initiating structure, production emphasis, consideration, and tolerance for freedom. Each story had four versions; one described a female manager with an Israeli Hebrew name, one a male manager with an Israeli Hebrew name. The other two versions reflected the interests of the larger study in ethnic as well as sex effects and portrayed males with "ethnic names." The four

Table 1
 Summary: Multivariate Analyses of Variance of Leadership Styles

Leadership style	Bartol and Butterfield		Izraeli and Izraeli	
	F^a	p	F^a	p
Initiating structure				
Manager sex	1.91	<.06	1.58	.16
Respondent sex	1.81	.07	0.70	.62
Interaction	0.72	.68	0.54	.75
Production emphasis				
Manager sex	0.57	.80	1.31	.26
Respondent sex	2.21	<.03	0.27	.93
Interaction	0.62	.76	0.64	.67
Consideration				
Manager sex	1.92	<.06	1.05	.39
Respondent sex	1.21	.29	0.61	.70
Interaction	0.76	.64	1.03	.40
Tolerance for freedom				
Manager sex	1.57	.13	1.07	.38
Respondent sex	1.90	<.06	0.61	.69
Interaction	1.04	.40	0.80	.55

Note. F_s are based on Wilks's lambda criterion.

Degrees of freedom for each $F = 8.301$ in the first study and 5.732 in the second study.

versions were otherwise identical. Only results for the versions with the male and female with the Israeli Hebrew name are reported in this study. It should be added that the grammatical structure of the Hebrew language, which has no neuter form, emphasizes gender differences more than does English. Demographic and organizational questions concluded the questionnaire. Level of education was measured by whether respondent had an academic degree. Respondents rated themselves as junior, middle, or senior manager.

Dependent Variables

Each story was followed by a set of five questions used by Bartol and Butterfield to evaluate the manager on a 7-point scale: 1. How productive is this organization now? 2. How satisfied do you think the employees are under this manager? 3. How do you think this manager's boss would evaluate his or her behavior? 4. How would you like to work for this manager? 5. All in all how effective do you think this manager is? For the analysis of the effects of education and organizational rank, an index of performance evaluation was constructed from the combined average scores on the five questions (α ranged from .89 to .90).

Procedure

Each respondent received four stories, one for each leadership style. The name of the manager in each of the four stories reflected one of the four sex/ethnic versions described above. The order of the stories (styles) and order of the versions (sex/ethnic identity) were randomized so that there were 16 versions of the full questionnaire. These were randomly distributed within each course. The researchers explained that the purpose of the study

was to examine preferences of Israeli managers for various leadership styles. No mention was made of our interest in the sex (or ethnic) variable.

Results

The data were analyzed separately for each of the four leadership styles. Results of the multivariate analyses of variance (ANOVA) for the two studies are summarized in Table 1. In this study, neither sex of manager nor sex of respondent had an overall main effect on performance evaluation under any of the four leadership styles. No interaction between sex of respondent and sex of manager was found.

Analyses of variance (ANOVA) were used to examine the effects of education and managerial level or rank on performance evaluation. Sex of manager was included as an independent variable. The results are as follows (see Table 2): For the initiation of structure leadership style the female is rated higher than the male (M for female = 5.13; M for

¹ Three of the questions used in the original study were not included. These were as follows: (a) What will be the future productivity of this branch under this manager? (b) What will be the future satisfaction of this branch under this manager? (c) To what extent do you think this manager should be considered for a raise or a promotion?

Table 2

Analysis of Variance in Performance Evaluation by Sex of Manager, Education, and Managerial Rank of Subject for Each Leadership Style

Source of variance	SS	df	MS	F	eta ²
Initiating structure					
A. Male versus female manager	6.42	1	6.42	3.19*	.008
B. Subjects with academic degree versus those without	0.12	1	0.12	0.07	.000
C. Subject's managerial rank: lower versus middle versus senior	1.59	2	0.80	0.47	.002
Interaction A × B	0.19	1	0.19	0.11	
Interaction A × C	17.13	2	8.57	5.09**	
Interaction B × C	0.53	2	0.27	0.16	
Residual	839.63	499	1.68	—	
Total	867.40	510	1.70	—	
Production emphasis					
A. Male versus female manager	6.83	1	6.83	4.73*	.010
B. Subjects with academic degree versus those without	0.00	1	0.00	0.00	
C. Subject's managerial rank: lower versus middle versus senior	5.11	2	2.56	1.77	
Interaction A × B	0.09	1	.09	0.06	
Interaction A × C	1.61	2	0.81	0.56	
Interaction B × C	0.01	2	0.00	0.00	
Residual	559.08	387	1.44	—	
Total	573.39	398	1.44	—	
Consideration					
A. Male versus female manager	4.31	1	4.31	2.25	.008
B. Subjects with academic degree versus those without	22.23	1	22.23	11.62***	.026
C. Subject's managerial rank: lower versus middle versus senior	28.37	2	14.18	7.41***	.036
Interaction A × B	0.05	1	0.05	0.025	
Interaction A × C	3.07	2	1.54	0.803	
Interaction B × C	16.61	2	8.31	4.34**	
Residual	748.07	391	1.91	—	
Total	838.02	402	2.08	—	
Tolerance of freedom					
A. Male versus female manager	0.14	1	.14	0.11	.001
B. Subjects with academic degree versus those without	1.79	1	1.79	0.23	.004
C. Subject's managerial rank: lower versus middle versus senior	1.25	2	0.56	0.44	.003
Interaction A × B	3.32	1	3.32	2.59	
Interaction A × C	0.20	2	0.10	.08	
Interaction B × C	1.24	2	0.62	.48	
Residual	454.63	355	1.28	—	
Total	463.46	366	1.27	—	

male = 4.88). An interaction effect was found between sex of manager and subject rank: For the female, the higher subject managerial rank, the more favorable the evaluation (junior $M = 4.46$; middle $M = 4.98$; senior $M = 5.35$), whereas for the male the opposite is the case (junior $M = 5.29$; middle $M = 4.98$; senior $M = 4.74$). For the production emphasis style, the female manager is rated lower than the male ($M = 4.78$ and $M = 5.03$, respectively). Although the interaction between sex of manager and managerial rank of subject is not significant, the pattern for the female manager is the same as for the previous style—the higher the rank the more positive the evaluation (junior $M = 4.38$; middle $M = 4.64$; senior $M = 4.91$). For the consideration style, subjects with academic degree rated the female lower than those without academic degree—an unexpected finding. The higher rank the more positive the evaluation of both the male and the female. This pattern, however, is not repeated in other styles. The negative effect of degree on evaluation (Interaction $B \times C$) is evident among junior and middle managers but not among senior managers. For the tolerance of freedom style, there were no main or interaction effects. In summary, subject education was significant for only one leadership style, and the effect was opposite from what one might have expected. Subject managerial rank either as a main effect or in interaction with sex of manager, was significant in three styles—in every case, the higher the rank of the subject the more favorable the rating given to the female manager.

Discussion

No sex of leader effects were found in this study; this finding contrasts with the results reported by Bartol and Butterfield (1976). They found that different criteria were used to evaluate the performance of a male and a female manager in two of the four leadership styles examined: The female was rated higher than the male when she used a consideration style but lower than the male when she used an initiating structure style. Although intended as a replication of the Bartol and Butterfield (1976) study, this research differs from theirs in four variables: respondents,

culture, timing, and instrument; each of these may in part account for the dissimilarities in outcome.

1. Respondents in the original study were students; we studied primarily middle and upper level managers. Evidence suggests that, compared to students and nonsupervisory personnel, managers are more responsive to behavioral cues relevant to the occupational role than to sex role stereotypes (Brown, 1979). In his review of the literature, Brown (1979, p. 605) reported that in studies where a significant difference in the leadership style variable was found, "most of the perceived style differences were perceived by students," and that "practicing managers overwhelmingly feel that there is no difference between male and female leadership styles; whereas students generally hold the opposite to be true" (p. 607). In most other studies where sex of leader was found to have a moderating effect on desired leadership style or behavior, respondents were not managers but were nonsupervisory personnel or subordinates asked to evaluate their superiors (Haccoun, Haccoun, & Sallay, 1978; Hansen, 1974; Petty & Lee, 1976; Petty & Miles, 1975; Terborg & Shingledecker, 1983).

In addition, we found that among managers, rank had a significant effect on evaluation, particularly of the female manager—the higher the rank the more favorable the rating. This finding cannot be explained by differences in education, because the latter was not associated with performance evaluation scores, except for the consideration style where the more educated rated the female manager lower than did the less educated. It may be suggested that women, competing with men for entry into management or newly arrived in managerial positions and congregated at the lower levels of the hierarchy, pose a greater threat to the status and/or position of nonsupervisory personnel and first-line managers than to the status of those higher up, with whom they are not (yet) in competition.

2. The absence of sex differences in this study may indicate that stereotypical biases are changing as a function of time or that greater sensitivity to the issue elicits more socially desirable responses. This argument is supported by the fact that a later replication of their own original study (Butterfield &

Bartol, 1977), based on a sample of both students and managers, did not reproduce the bias in favor of sex-congruent behavior. It is unlikely, however, that the time gap between this and our study (about 6 years) is a significant factor. The feminist revival arrived in Israel approximately a decade after its emergence in the United States (Rein, 1979), and there is no affirmative action legislation in Israel, two factors which in the United States reduced sex bias and increased sensitivity to expressions of prejudice.

3. The dependent variable measure used by Bartol and Butterfield (1976) was shortened to reduce the response time required and thus increase our access to managers. Because the multivariate analysis of variance takes into account the effect of each measure as well as the interrelationships among them, without our rerunning the original data deleting the same three items, it is not possible to determine the effects of the change in the measure.

4. The difference in findings between the United States and Israeli study may be affected by differences in the two cultures. Papenek (1980) has argued that cultures differ in the depth to which gender differences are incorporated and in the extent to which people are supposed to or are permitted to be context sensitive and the degree to which they adjust their behavior in response to the demands of the social situation. Bartol and Butterfield (1976) suggest that to be successful in the United States, women must be sex-role congruent and that women are expected to perform to sex-type even in a managerial position.² There is some evidence that Israelis may have fewer sex stereotypes than Americans. In an Israeli replication of the Broverman et al. (1972) study, Ziv (1972) found only nine of the 19 attributes reported by Broverman, Vogel, Broverman, Clarkson, and Rosenkrantz (1972) to be male stereotypes and 13 of the 19 attributes reported to be female stereotypes. In a cross-cultural replication of the Goldberg (1968) study, which included a sample of kibbutz members and of city (Jerusalem) dwellers, Mischel (1974) found that Israelis showed the same occupational sex role stereotypes as the U.S. subjects regarding the sex association of four fields (law, city planning, dietetics, and primary

education) but no sex bias in evaluating professional journal articles by authors in those fields. Cultural norms of Israeli society "permit" women to adopt patterns of behavior more typically associated with men—such as taking initiative in work situations, being task oriented, and being assertive where required. Although socialization may not foster these behaviors, social norms condone them particularly in achievement situations.

Unfortunately, the dissimilarities of a number of variables in the original study and this one make it difficult to evaluate the effects of any one of them on the differences in results. The findings of this study, however, add to the growing body of research suggesting that when the sample is composed of managers, similar behavioral criteria are used for evaluating both male and female managers. Therefore, a major implication of this research is that a context-appropriate leadership style may be more effective for women than an invariant attachment to traditional female-associated patterns. A second implication is that if criteria for judging women vary with the level of subject in the organization, and those in the lower ranks expect or prefer stereotypical behavior, whereas those higher up do not, women may find themselves under pressure to adjust their style contingent on the setting of their interaction. This requires a juggling act that puts additional strains on women, not faced by men who apparently have substantially greater freedom to engage in a variety of leadership styles (Jago & Vroom, 1982; Inderlied & Powell, 1979).

Previous research has underlined the importance of sponsorship for women (Epstein, 1970; Kanter, 1977). Epstein (1970; 170) notes that the sponsor-protégé relationship "may be more important for her than for a man." The fact that lower level managers rate the female manager less favorably endorses Epstein's statement—whereas the more favorable ratings by senior managers suggest that it may be easier for a woman to get the

² Other studies, however, suggest the opposite to be true. Because the idealized manager is essentially a masculine prototype (Schein, 1973; Schein, 1975; Powell & Butterfield, 1979), women should engage in stereotyped masculine behaviors (Rosen & Jerdee, 1975).

support of those higher up than of her immediate superior.

References

- Bartol, K. M., & Butterfield, D. A. (1976). Sex effects in evaluating leaders. *Journal of Applied Psychology, 61*, 446-454.
- Birsdall, P. (1980). A comparative analysis of male and female managerial in communication style in two organizations. *Journal of Vocational Behavior, 16*, 183-197.
- Brown, S. M. (1979). Male versus female leaders: A comparison of empirical studies. *Sex Roles, 5*, 595-611.
- Butterfield, D. A., & Bartol, K. M. (1977). Evaluators of leader behavior: A missing element in leadership theory. In J. G. Hunt & L. L. Larson (Eds.), *Leadership: The cutting edge*. Carbondale, IL: Southern Illinois University.
- Broverman, I. K., Vogel, S. R., Broverman, D. M., Clarkson, F. E., & Rosenkrantz, P. S. (1972). Sex-role stereotypes: A current appraisal. *Journal of Social Issues, 28*, 59-78.
- Chapman, J. B. (1975). Comparison of male and female leadership styles. *Academy of Management Journal, 18*, 645-650.
- Dipboye, R. L., & Wiley, J. W. (1977). Reactions of college recruiters as a function of interviewee sex and self presentation style. *Journal of Vocational Behavior, 10*, 1-12.
- Donnell, S. M., & Hall, J. (1980). Men and women as managers: A significant case of no significant difference. *Organizational Dynamics, Spring*, 60-77.
- Epstein, C. (1970). *Woman's place*. Los Angeles: University of California Press.
- Goldberg, P. A. (1968). Are women prejudiced against women? *Transaction, April* 28-30.
- Haccoun, D. M., Haccoun, R. R., & Sallay, G. (1978). Sex differences in the appropriateness of supervisory styles: A non-management view. *Journal of Applied Psychology, 63*, 124-127.
- Hansen, P. (1979). *Sex differences in supervision*. Paper presented at the meeting of the American Psychological Association, New Orleans.
- Inderlied, S. D., & Powell, G. (1979). Sex-role identity and leadership style: Different labels for the same concept? *Sex Roles, 5*, 613-625.
- Izraeli, D. N., & Izraeli, D. (1980, September). *Sex and ethnic effects on evaluating leaders*. Research Report No. 37/80, The Israel Institute of Business Research, Tel Aviv University.
- 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.
- Kanter, R. M. (1977). *Men and women of the corporation*. N.Y.: Basic Books.
- Mischel, H. N. (1974). Sex bias in the evaluation of professional achievements. *Journal of Educational Psychology, 66*, 157-166.
- Nieva, V. F., & Gutek, B. A. (1980). Sex effects on evaluation. *Academy of Management Review, 5*, 267-276.
- Papanek, H. (1980, August). *Family status production: A new perspective on the "work" and "non work" of women*. Paper presented at the 75th annual meeting of the American Sociological Association, New York.
- Petty, M. M., & Lee, G. K. (1975). Moderating effects of sex of supervisor and subordinate on relationship between supervisor 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.
- Powell, G. P., & Butterfield, D. A. (1979). The "good manager": Masculine or androgynous? *Academy of Management Journal, 22*, 395-403.
- Rein, N. (1979). *Daughters of Rachel: Women in Israel*. New York: Penquin Books.
- Rosen, B., & Jerdee, T. H. (1975). Effects of employee's sex and threatening versus pleading appeals on managerial evaluation of grievances. *Journal of Applied Psychology, 60*, 442-445.
- Schein, V. E. (1973). The relationship between sex role stereotypes and requisite management characteristics. *Journal of Applied Psychology, 57*, 95-100.
- Schein, V. E. (1975). Relationships between sex role stereotypes and requisite management characteristics among female managers. *Journal of Applied Psychology, 60*, 340-344.
- Stitt, C., Schmidt, S., Price, K., & Kipnis, D. (1983). Sex of leader, leader behavior and subordinate satisfaction. *Sex Roles, 9*, 31-42.
- Terborg, J. R., Peters, H., Ilgen, R., & Smith, F. (1977). Organizational and personal correlates of attitudes toward women as managers. *Academy of Management Journal, 20*, 89-100.
- Terborg, J. R., & Shingledecker, P. (1983). Employee reaction to supervision and work evaluation as a function of subordinate and manager sex. *Sex Roles, 9*, 813-824.
- Ziv, A. (1972). Sex differences in male and female stereotypes among adolescents in Israel, *Hahinoch, 35*, 64-69 (Hebrew).

Received April 23, 1984

Revision received December 19, 1984 ■