

## SOCIAL CONTEXT OF PERFORMANCE EVALUATION DECISIONS

TIMOTHY A. JUDGE

Cornell University

GERALD R. FERRIS

University of Illinois at Urbana-Champaign

**The role of social and situational influences in the performance-rating process has received relatively little research attention yet merits increased attention. Although there has been acknowledgment of the role of social and situational factors in shaping rater cognition and evaluation, research has typically proceeded in a piecemeal fashion, isolating a single variable at a time. Such an approach fails to recognize that performance rating is a process with multiple social and situational facets that need to be considered simultaneously. In the present study, we tested a model of the performance-rating process, employing several social and situational variables that have been infrequently investigated and typically not in conjunction with one another. Results indicated support for the overall model and specific links within it. Implications of the results for performance-rating research are discussed.**

There is perhaps not a more important human resources system in organizations than performance evaluation. Supervisors' ratings of subordinates' performance represent critical decisions that are key influences on a variety of subsequent human resources actions and outcomes. Indeed, this pivotal role of performance evaluation has promoted systematic efforts to develop a more informed understanding of the performance-rating process.

Landy and Farr (1980) issued a call for research investigating the cognitive processes underlying performance appraisal decisions. Although the process focus has generated considerable research concerning various components of performance-rating decisions, more comprehensive investigations incorporating several of those components has been lacking. Furthermore, process-oriented research has been limited by its reliance on laboratory studies (DeNisi & Williams, 1988). Whereas the cognitive processes involved in performance-rating decisions can be well illuminated in laboratory studies, the "quiet" nature of laboratory studies often does not match the "noisy" context in which performance-rating decisions are actually embedded (Lord & Maher, 1989).

---

We thank Robert D. Bretz, Jr., and two anonymous reviewers for comments on earlier versions of this article.

Other researchers in the performance-rating area, although acknowledging that cognitive issues are important, have argued that social and situational factors have been neglected (Dipboye, 1985; Ferris & Judge, 1991; Ilgen & Favero, 1985; Mitchell, 1983; Wexley & Klimoski, 1984). As Mitchell pointed out, given that employees often work in groups, that some of their work is unobserved, and that evaluators often have various motives in evaluating performance, traditional approaches to performance appraisal may be inadequate. Thus, the social context would appear to be important in the investigation of performance-rating decisions.

The purpose of the present study was to propose and test a model of social influence in the performance-rating process. Implicit in the development of the model was recognition that the performance-rating process has multiple social and situational facets that should be simultaneously considered. This approach moves beyond the fragmentary manner in which past research has generally investigated social and situational variables. The proposed model is not intended to be a comprehensive test of all social and situational elements that may affect performance ratings. Rather, we employed a set of key social and situational variables, including some not previously tested, to investigate the overall relationship between social processes and performance ratings, as well as the specific effects of the variables in the model. The results should provide useful information regarding the importance of social context in the performance-rating process.

### **ELEMENTS OF SOCIAL CONTEXT IN PERFORMANCE-RATING DECISIONS**

In the last decade, a great deal of research has investigated the cognitive processes underlying performance appraisal decisions (Borman, 1978; Cooper, 1981; DeNisi, Cafferty, & Meglino, 1984; DeNisi & Williams, 1988; Feldman, 1981; Ilgen & Feldman, 1983; Landy & Farr, 1983; Motowidlo, 1986; Nathan & Alexander, 1985; Nathan & Lord, 1983). Although researchers have learned much about cognitive processes, DeNisi and Williams noted that very little has been new in the models proposed in the last decade and in the elements of those models that have been tested. Furthermore, some performance-rating researchers have warned that an exclusive focus on cognitive processes is likely to miss an important element of performance evaluation, namely the contextual influences within which rating decisions are embedded (Dipboye, 1985; Ilgen & Favero, 1985; Nathan, Mohrman, & Milliman, 1991; Wexley & Klimoski, 1984). Social psychology researchers have echoed this concern. For example, Schneider (1991) suggested that past research on performance evaluation has focused on intrapsychic processes but overlooked the social and situational contexts. He noted that increased research attention regarding the influence of social context on cognitive processes is sorely needed.

Research on cognitive processes can be seen as complementing, rather than opposing, investigation of the relationship between contextual factors

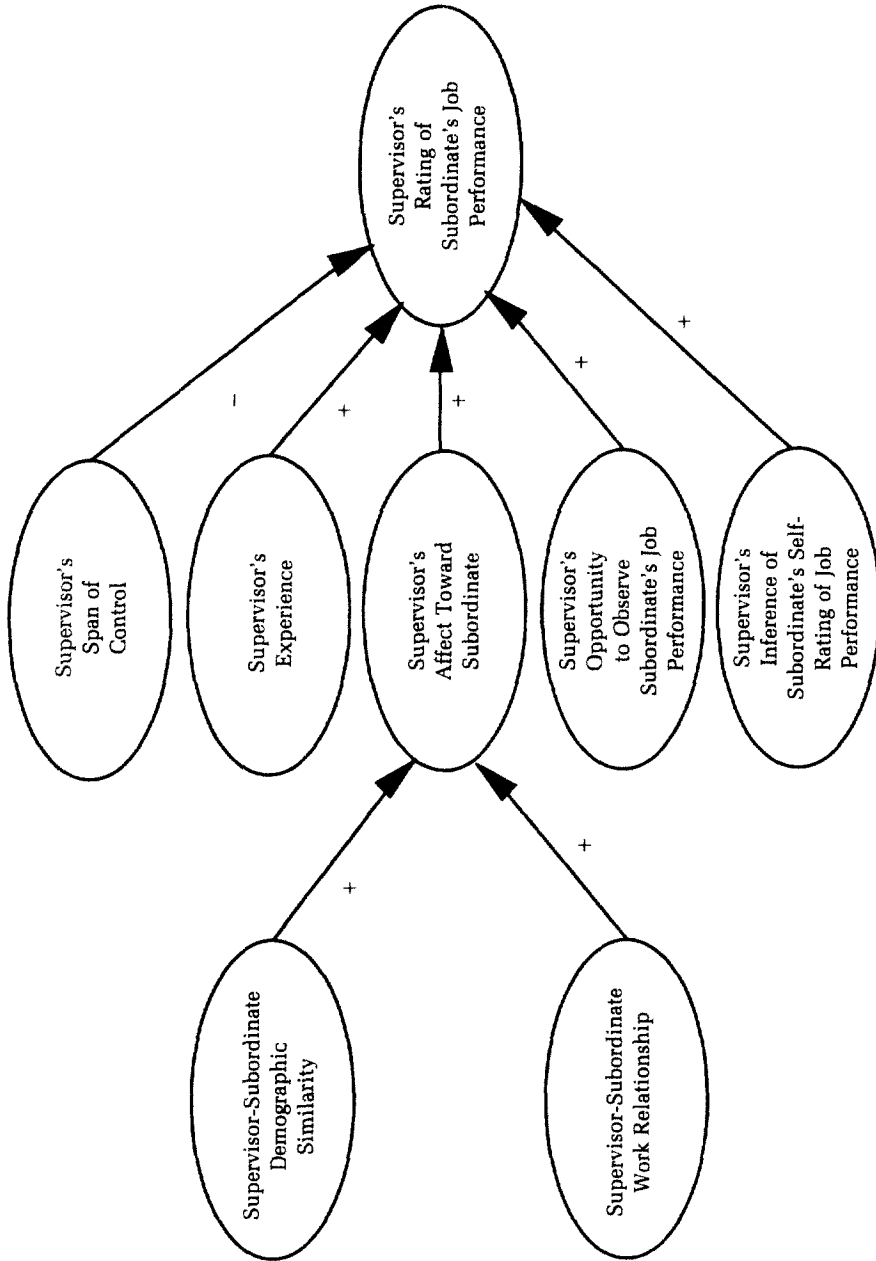
and performance-rating decisions. Research focused on cognition has clearly demonstrated that the rating process is subjective and that extraneous influences may enter the process at each of the many cognitive steps raters use to process information; these steps include attention, encoding, storage, retrieval, integration, and rating (DeNisi & Williams, 1988). Thus, advocates of the cognitive approach to performance evaluation assume that errors in rating result from how individuals input, process, and recall information. Inaccuracy in ratings, or deviations in ratings from "true" performance, are due to specific cognitive processes. If social and situational elements are salient to a rater, they are likely to influence the rater's cognitions of performance information. Thus, although more attention may need to be focused on the context of rating decisions, the cognitive focus is still relevant. In fact, research regarding the subjectivity of the rating process suggests a need to consider contextual elements that may relate to the cognitive processes involved in performance rating.

#### **MODEL OF SOCIAL AND SITUATIONAL ELEMENTS IN THE PERFORMANCE-RATING PROCESS**

The foregoing review indicates that it may be important to consider social and situational influences if researchers hope to better understand the performance-rating process. Whereas some previous research has posited the effects of specific social and situational variables on performance ratings, the effects of those variables either have not been extensively investigated, or have been tested in isolation. In reality, social and situational variables probably do not affect performance ratings in isolation, but rather need to be considered in conjunction with one another. Therefore, in order to understand the overall effect of the social context in performance ratings, we tested a model that allows consideration of a number of relevant social and situational factors simultaneously. The model was based primarily on the conceptualization proposed by Mitchell (1983) and also on specific influences hypothesized by others (Ferris & Judge, 1991; Ferris & Mitchell, 1987; Klimoski & Inks, 1990; Wexley & Klimoski, 1984). Thus, the proposed conceptualization, displayed in Figure 1, is a hybrid model intended to effectively represent the key social context variables affecting the performance-rating process.

The model depicts a number of social and situational influences on supervisors' ratings of subordinates' performance, two of which operate through supervisory affect toward subordinates. Mitchell's (1983) model of social and task influences, which consists of three classes of influences, served as the basis for four links in the model. A key environmental influence on performance ratings hypothesized by Mitchell (1983) was ease of observation. Ease of observation may improve the quality of the observation of subordinate performance and thus increase the accuracy of performance ratings. However, the fact that ease of observation may improve the accuracy of ratings does not inform us about its direction of influence.

**FIGURE 1**  
**Hypothesized Model of Social Influence in the Performance Evaluation Process**



It could be argued that increased opportunity to observe leads to negative performance ratings because more information and justification for appropriately harsh ratings becomes available. If raters tend to rate leniently because they do not possess the evidence to justify or substantiate more negative ratings, such a hypothesis makes sense.<sup>1</sup>

However, the more compelling hypothesis is that ease of observation positively affects performance ratings. The evidence strongly suggests that in general, supervisors prefer to issue positive performance ratings to subordinates (Klimoski & Inks, 1990; Stone, 1973). Furthermore, there is an upward bias in performance ratings, with subordinates generally given more positive ratings than their true performance would indicate (Bass, 1956; DeCotiis & Petit, 1978). This bias probably results from the fact that more negative events, such as subordinate resentment, complaints, and lawsuits, derive from issuing negative ratings than from issuing positive ratings (Bernardin & Beatty, 1984). Given the inclination of supervisors, the performance-rating process is arguably a search for positive information. Following "hypothesis confirmation strategies" (Snyder & Swan, 1978), supervisors seek confirmation of their impressions of individuals (Feldman, 1981). Increased liking resulting from familiarity (Zajonc, 1980) may augment this process. Supervisors, because they are motivated to issue positive ratings, in fact may search for positive data to support their motivations. We expected, then, that the more time they spend observing subordinates, the more positive information they will uncover.

This prediction contrasts sharply with research on employment interviews, which have been characterized as searches for negative information (Rowe, 1989). The difference in the two processes can be understood by examining their outcomes. If an interviewer makes a mistake in hiring, it is best that it be on the side of failing to hire a qualified candidate (overweighting negative information) than on the side of hiring an unqualified candidate (overweighting positive information). Conversely, there are more negative than positive consequences when a supervisor issues an overly negative performance rating. Thus, it is best to err on the positive side in rating performance. If the performance-rating process is a search for positive information in most organizations, the more opportunity a supervisor has to observe a subordinate's performance, the more positive information will be encoded and negative information discounted (Feldman, 1981).

Some prior research also supports the idea that extensive opportunity to observe enhances performance ratings. McFillen and New (1979) found a significant interaction between closeness of supervision and subordinate performance in predicting supervisors' ratings of subordinates' performance. Supervisors judged successful subordinates to have exhibited better overall performance when the supervisors closely supervised the subordinates than when they only generally supervised them. The opposite was the case for subordinates who were performing poorly. In this context, close supervision

---

<sup>1</sup> We thank an anonymous reviewer for suggesting this possibility.

may have raised supervisors' ratings of subordinates' performance because the supervisors attributed the subordinates' performance to their supervisory efforts. Furthermore, these results suggest that greater opportunity to observe tends to produce more variable evaluations of subordinates, which, as McFillen and New (1979) suggested, may be at least partially the result of the greater information available. This finding is relevant to the positive relationship hypothesized between opportunity to observe and performance ratings because of the nature of typical subordinate performance distributions in organizations. In most organizations, the employee performance distribution is neither normal nor reflective of the full range of possible values. Instead, considerable restriction of range on performance ratings occurs for a number of reasons, including self-selection, poor performers' leaving, terminations, and so forth. Most employees who are rated are performing at least satisfactorily; the lower end of the performance distribution has been deleted. Increased opportunity to observe should lead to more favorable evaluations of subordinate performance in such a context, because the more supervisors observe subordinates, the more the former will realize that poor performers are nonexistent or substantially reduced due to the above factors.

Another area related to opportunity to observe that has been considered relevant in the performance evaluation process is supervisor-subordinate distance (Ferris & Judge, 1991; Ilgen & Feldman, 1983; Landy & Farr, 1983; Mitchell, 1983; Wexley & Klimoski, 1984). Distance between supervisors and subordinates can manifest itself both physically and psychologically. High spatial distance between the areas supervisors and subordinates occupy in the regular operation of job-related tasks and duties is likely to reduce the supervisors' opportunity to observe. Research on employee shirking is informative here; it would predict that ease of observation motivates high performance because subordinates realize their efforts are being monitored (Jones, 1984). Furthermore, Ferris and Judge (1991) indicated that low spatial distance provides additional information for supervisors through interpersonal influence dynamics, which favorably affect performance. This premise would suggest a negative relationship between spatial distance and supervisors' ratings of subordinates' performance. Ferris, Judge, Rowland, and Fitzgibbons (in press) confirmed this prediction, reporting that spatial distance and supervisors' ratings of subordinates' performance were significantly, negatively related ( $r = -.29, p < .01$ ). Thus, considerable conceptual and theoretical evidence, as well as the results of empirical research, suggests that ease of observation will positively influence performance ratings.

*Hypothesis 1: A supervisor's opportunity to observe a subordinate's job performance will positively influence the performance rating of the subordinate.*

Ease of observation may be manifested in part by span of control. The more employees a particular supervisor must oversee, the less opportunity

he or she will have to observe each subordinate's performance. However, span of control is by itself an imperfect measure of ease of observation. Because supervisors differ in the degree of nonsupervisory responsibilities they have, supervisors with identical spans of control may not have the same ease of observation. Furthermore, some supervisors with large spans of control may become highly efficient under a heavy work load and actually maintain greater direct control over each subordinate than other supervisors with smaller spans of control. Thus, it is important to establish and maintain the conceptual and empirical distinctions between span of control and direct ease of observation. We expected span of control to have a negative effect on performance ratings, although we have hypothesized that opportunity to observe will have a positive effect on performance ratings because it captures relevancy rather than mere frequency of contact (Landy & Farr, 1983).

Further theoretical support for the effect of span of control on performance ratings can be found in research on employee shirking. Jones (1984) hypothesized that shirking would be more likely to occur in large groups because supervisory monitoring of performance is difficult in such a setting. Support for this proposition comes from Latane, Williams, and Harkins (1979), who found that productivity decreased as group size increased. Research from labor economics has also supported the relationship between monitoring and employee shirking (Groshen & Krueger, 1990; Kahn & Sherer, 1990). This theory and research is relevant because it suggests that increased spans of control make monitoring more difficult, which leads to lower subordinate performance, and thus lower performance ratings (Judge & Chandler, 1990).

*Hypothesis 2: A large supervisory span of control will negatively influence the rating of a subordinate's performance.*

Again drawing from Mitchell (1983), we expected supervisors' experience to influence ratings of subordinates' performance positively. Most past research suggests that supervisory experience positively influences performance ratings (Landy & Farr, 1980), and we would expect this relationship for several theoretical and conceptual reasons. First, it may be that less experienced supervisors rate harshly to demonstrate their capabilities to handle the job of supervisor and make tough decisions. As supervisors gain experience and self-confidence and become established in their jobs, they may see less need to demonstrate their toughness and, in fact, they may well adopt more lenient rating tendencies. Another explanation is that experienced supervisors may simply have a better understanding of, and appreciation for, the complexities of task performance. An experienced supervisor may have a better grasp of the uncertainties inherent in the behavior-performance relationship. Finally, it also may be that supervisors experience the costs of giving unfavorable ratings only over time, through subordinate complaints, appeals, and hostility. As supervisors gain experience, it is likely they also gain some degree of sagacity. More experienced supervisors

may have learned that unfavorable ratings simply are not worth the trouble they cause, and also that their judgments are not error-free, which increases their willingness to give subordinates the benefit of the doubt when issuing performance ratings.

*Hypothesis 3: The extent of supervisory experience will positively influence the supervisor's performance ratings of subordinates.*

The final class of social and task influences on performance ratings Mitchell (1983) discussed were elements of the social context, in this case similarity between supervisor and subordinate. Interpersonal similarity has been an area of considerable investigation in performance evaluation research (Miles, 1964; Pulakos & Wexley, 1983; Senger, 1971; Turban & Jones, 1988; Wexley, Alexander, Greenawalt, & Couch, 1980). More recently, Ducheon, Green, and Taber (1986) found that similarity on several demographic characteristics was positively associated with the mutual affect felt between a supervisor and subordinate. Interpersonal similarity may lead to attraction and compatibility, according to the well-established similarity-attraction theory (Byrne, 1969).

*Hypothesis 4: Demographic similarity between a supervisor and subordinate will positively influence the supervisor's affect toward the subordinate.*

Researchers have also suggested that the nature of the supervisor-subordinate work relationship influences both the attitudes and behaviors of supervisors toward subordinates. Research on leader-member interactions has demonstrated that leaders establish much closer and more frequent interaction patterns with subordinates they identify as members of their in-group (i.e., subordinates with whom they have a high-quality work relationship) than with subordinates they classify as out-group members (i.e., those with whom they have a low-quality work relationship) (Dienesch & Liden, 1986; Graen, 1976; Wayne & Ferris, 1990).

Researchers in the area of cognitive psychology have emphasized the importance of interactions in the formation of people's impressions of others. For example, when individuals are asked to evaluate someone, they retrieve information based on their interactions with that person (Hastie & Park, 1986). As Srull and Wyer (1989) noted, a central component of such evaluations is often affective. This idea suggests that the nature of their interactions at work is likely to be critical to the degree to which a supervisor reports liking a particular subordinate. Furthermore, the "mere presence hypothesis" from social psychology indicates that closeness of interaction leads to liking (Byrne, 1961; Saegert, Swap, & Zajonc, 1973). If that is the case, supervisors who have a close working relationship with certain subordinates in terms of the frequency and quality of their day-to-day interactions will be more likely to like those subordinates than others.



*Hypothesis 5: A close supervisor-subordinate work relationship will positively influence a supervisor's affect toward a subordinate.*

Supervisors' affect is in turn believed to influence their ratings of subordinates' performance. Following prior research (Cardy & Dobbins, 1986; Kingstrom & Mainstone, 1985; Tsui & Barry, 1986; Wayne & Ferris, 1990), we hypothesized that a supervisor's positive affect toward a subordinate has a positive effect on the supervisor's ratings of the subordinate's performance.

*Hypothesis 6: A supervisor's positive affect toward a subordinate has a positive effect on the supervisor's rating of the subordinate's performance.*

Finally, following Klimoski and Inks (1990), we expected a supervisor's inference of a subordinate's self-rating of job performance to positively influence the supervisor's rating. If supervisors think that particular subordinates believe they have performed well, the supervisors experience "expectational pressures," for example, pressure to conform to a subordinate's wishes, that make them feel more accountable for negative ratings and thus pressure them to issue higher ratings (Klimoski & Inks, 1990).

*Hypothesis 7: A supervisor's inference that a subordinate believes he or she has performed well will positively influence the evaluation of the subordinate's performance.*

## METHODS

### Respondents

Participants in this research were 81 registered nurses and their supervisors chosen from all major subunits in the nursing service departments of a 283-bed hospital located in central Illinois. All the staff nurses and supervisors were Caucasian women, and their mean age was 33.72 years, with a range of ages of 21–61 years. The average job tenure for the staff nurses was 12 years, and the range was 1 month to 32.25 years. The mean tenure working for a current supervisor was 12.57 months, with a range of 1 month to 4 years. The cumulative average difference in age and tenure between supervisors and subordinates was 66.41 months (s.d. = 76.00).

As is the case in many organizations, the structure of the nursing service departments did not allow a single supervisor for each subordinate. Thus, for the 81 supervisor-subordinate dyads, there were 81 subordinates and 27 supervisors. We tried to keep the number of subordinates included for any single supervisor to a minimum to circumvent potential response biases on the part of supervisors. The number of subordinates for any given supervisor in this study ranged from one to four. The nursing administration office assisted with the support and coordination of the study and in the selection

of subordinates for each supervisor. Whereas the selection process was not completely random, there was no reason to suspect that it was seriously biased. In fact, we tried to ensure variability on subordinate age and tenure, a goal that appears to have been accomplished, as the figures above suggest.

### **Procedures**

A roster was developed a week before the actual data collection listing 100 nursing supervisor-subordinate dyads from all the major departments in the hospital. On the day data were collected, all day shift supervisors completed questionnaires in a supervisory meeting room about one hour before the end of their shift. They responded to the same set of questions about each of their staff nurses.

The supervisors' staff nurses were let off work about one hour before their shift ended and reported to a conference room to fill out their questionnaires. Pretesting indicated that approximately 30 minutes were required for completion of the staff nurses' questionnaires and approximately 40 minutes were required for completion of the supervisors' questionnaires. On the evening shift, which was sparsely staffed, as is typical in hospitals, all the supervisors reported to a designated conference room near the end of their shift to complete questionnaires, and staff nurses reported to the same room at the end of their shift.

Staff nurses may have more than one supervisor. To circumvent the problem of a nurse's not knowing which supervisor to focus on in a questionnaire, each staff nurse was instructed as to the person she should regard as her supervisor in this study. Because of vacations, absences, and people's choosing not to participate on the day data were collected, the final study group consisted of 81 supervisor-subordinate dyads. We informed all staff nurses and supervisors that they would receive a complete report of the purpose, results, and implications of this research when the study was completed and that the confidentiality of their responses would be maintained. Participation was completely voluntary.

### **Measures**

**Performance rating.** Supervisors rated each subordinate's overall work performance on a five-point Likert scale (1 = very poor, 5 = very good). This measure was similar to that recently used by Nathan and colleagues (1991).

**Supervisors' affect toward subordinates.** Supervisors responded to one item assessing their degree of liking for each subordinate. A five-point scale (1 = I don't like this subordinate at all, 5 = I like this subordinate very much) was used for responses to the question, "How much do you like this subordinate?"

**Supervisors' opportunity to observe subordinates' performance.** We asked each subordinate to respond to the following item: "Sometimes a supervisor's job is such that he/she does not have a good opportunity to observe the work performance of his/her employees (e.g., due to being overloaded with work or due to having so many employees working for him/her

there is just not the opportunity to regularly interact with all of them). How much do you think your supervisor regularly has the opportunity to observe your job performance and thus knows how you are doing?" A five-point scale (1 = almost never; my supervisor almost never has the opportunity to observe my job performance, 5 = very much; my supervisor regularly has the opportunity to observe my job performance) was used for responses.

**Supervisor-subordinate work relationship.** With the focus of this research on understanding the performance evaluation process through the on-going, day-to-day interactions of supervisors and subordinates, it was necessary to use a measure that assessed the relationship between each subordinate and supervisor. We asked the nurses five questions developed by Graen and his associates (e.g., Dansereau, Graen, & Haga, 1975; Graen & Scheimann, 1978) for their measure of leader-member exchange: (1) "How close of a relationship do you have with your supervisor?", (2) "How flexible do you believe your supervisor is in bringing about change in your job?", (3) "What are the chances your supervisor will use his/her power to help you solve problems in your job?", (4) "How much can you count on your supervisor to 'bail you out,' at his/her expense, when you really need him/her?", and (5) "How often do you take your suggestions regarding your job to your supervisor?" We summed the responses, all on five-point scales, to form a work relationship score ( $\alpha = .83$ ).

**Demographic similarity.** A composite measure of demographic similarity between supervisors and subordinates was created by standardizing and summing the absolute differences between the groups on age and job tenure. To transform this figure into a measure of similarity, we divided it into 1. Because the supervisors and subordinates were of the same gender and race and had similar levels of education, we saw age and tenure as the most relevant measures of demographic similarity for this study (Turban & Jones, 1988).

**Supervisors' inference of subordinates' self-ratings of performance.** Supervisors were asked to indicate how they thought each of their subordinates would evaluate their own job performance (1 = very poor, 5 = very good).

**Span of control.** Supervisors gave the number of employees who reported directly to them; the nursing administration office verified this information.

**Supervisory experience.** Self-reported tenure in a current position was the measure of a supervisor's experience.

### Reliability of the Measures

Fundamental to the integrity of a model's estimation results are the psychometric properties of the measures in that model. Whereas one-item measures are not inherently defective (Scarpello & Campbell, 1983), they are of concern because of their unknown reliability. Thus, before proceeding with model testing, we sought evidence for the stability across time of the single-item measures in our model by assessing their test-retest reliability.

The test-retest method of assessing reliability involves administering measures to the same group of people on two or more occasions, with a specified time interval between administrations. In fact, a critical determination in the use of the test-retest method is the precise time interval between testing occasions, because too short an interval can introduce memory effects and too long an interval can increase the likelihood of true rating changes (Ghiselli, Campbell, & Zedeck, 1981). Given the types of measures used in this study, we chose a one-week time interval between administrations to minimize those limitations.

Test-retest data were gathered from 57 graduate and undergraduate students in two personnel management courses. The students were asked to assume the role of a manager who supervises several employees. They received a three-page scenario describing the performance of a subordinate in a particular incident and background and performance data on that employee. We asked them to respond to questions regarding their opportunity to observe the subordinate's performance, their degree of liking for the subordinate, their performance rating of the subordinate, and their inference about the subordinate's self-rating of performance.

One week later, the identical procedure was repeated. Before completing the survey, however, the students indicated the degree to which they remembered the material from the previous week (1 = I remember very little of last week's exercise; 4 = I remember almost all of last week's exercise). We computed test-retest reliability coefficients as the correlations between the responses to the same measures at the two times, controlling for the amount each subject remembered from the previous week's exercise. The test-retest coefficients were as follows: opportunity to observe subordinate's performance .72; degree of liking toward the subordinate, .81; performance rating, .67; and inference about the subordinate's self-rating of performance, .64. The results suggest the single-item measures possess adequate reliabilities, although it is possible that a similar analysis using nonstudent subjects and different procedures might reveal different results.

## RESULTS

### Covariance Structure Model

To test the hypothetical model presented in Figure 1, we estimated a covariance structure model. Such models have a number of advantages. First, covariance structure modeling algorithms, such as LISREL, allow the joint specification and estimation of the measurement and structural models argued to account for observed data (Jöreskog & Sörbom, 1989; Long, 1983), which allows the correction of structural estimates for measurement error. Second, LISREL provides fit indexes that, when examined cumulatively, can provide useful information regarding the overall fit of a model to the given data. Finally, because LISREL incorporates structural equation-modeling techniques into its algorithm, alternatives to an original model can be tested. Such tests may provide some information about the direction of influence,

but when the assumption of weak causal order is violated, alternative model tests should be interpreted cautiously.

A number of assumptions underlie the interpretation of causal effects based on results from covariance structure analysis (James, Mulaik, & Brett, 1982). Rarely are all of these assumptions met in practice, although the consequences of violating them vary in importance (Bollen, 1989). A critical assumption of causal analysis is the necessity of ruling out reciprocal relationships between variables (James et al., 1982). Thus, although in theory causal assumptions are perfectly consistent with covariance structure analysis, practical realities suggest caution in drawing causal inferences from LISREL results.

### Results

Table 1 presents the means, standard deviations, and intercorrelations of the variables used in the analysis. The table reveals that supervisory experience is strongly and positively correlated with demographic similarity between supervisor and subordinate. This correlation occurs because the supervisors studied had less experience in their positions than did their subordinates. Thus, the more experienced supervisors were closer to the experience level of their subordinates than were the less experienced supervisors. Also, supervisors' opportunity to observe subordinates' performance is highly correlated with supervisor-subordinate work relationship, perhaps

**TABLE 1**  
**Descriptive Statistics and Correlations**

Variables	Means	s.d.	1	2	3	4	5	6	7
1. Demographic similarity	0.02	0.01							
2. Supervisor's rating of subordinate's performance	3.98	0.67	.13						
3. Supervisor's affect toward subordinate	4.26	0.72	.21	.30					
4. Supervisor-subordinate work relationship	18.77	4.25	.00	.27	.23				
5. Supervisor's span of control	20.38	10.43	.00	.00	.08	-.05			
6. Supervisor's opportunity to observe subordinate's job performance	3.64	1.15	.03	.21	.22	.56	.05		
7. Supervisor's inference regarding subordinate's self-rating of performance	3.83	0.65	-.06	.34	.10	-.04	-.10	-.08	
8. Supervisor's experience	26.26	28.84	.55	.10	.28	-.03	-.01	.20	-.01

**TABLE 2**  
**Results of Structural Model Estimate**

Fit Statistics	N = 81	N = 27
$\chi^2$	13.67	4.44
df	24	24
$\chi^2/df$	0.57	0.19
Goodness-of-fit index	.960	.960
Adjusted goodness-of-fit index	.940	.940
Root-mean-square residual	.060	.060
Overall $R^2$	.244	.244
$R^2$ , performance	.210	.211

indicating that supervisors tended to spend time with, and thus have greater opportunity to observe, the subordinates with whom they worked best.

Concern might arise about the size of the data set for the covariance structure analysis. Bentler (1985) suggested that a sample-size-to-parameter-ratio of 5 or more is sufficient to achieve reliable estimates in maximum likelihood estimation. Since that ratio in the present study was 6.8:1, we considered the size adequate for the analyses (Brooke, Russell, & Price, 1988).

The data set used for analyses was constructed by matching the appropriate supervisor's responses to each subordinate's responses. There are thus 81 independent observations of subordinates' responses but only 27 independent observations of supervisors' responses. Observations of the supervisor-provided variables are therefore not independent of one another, generating two problems.<sup>2</sup> First, it was not clear upon what number of observations the fit statistics should be calculated. One conservative approach was to calculate fit statistics based on both  $N = 81$  and  $N = 27$ . The fit statistics displayed in Table 2 demonstrate that the hypothesized model represents the data well if either 81 or 27 observations are assumed.<sup>3</sup>

A second problem when all observations are not independent is that there will likely be a positive correlation between error terms, or autocorrelation. The consequences of this violation are that although maximum likelihood will still be an unbiased estimator of the structural parameters, it is no longer the most efficient estimator, nor is it an unbiased estimator of the variance of structural parameters (standard errors). Thus, standard statistical tests of structural coefficients may be biased.

In situations of potential autocorrelation, generalized least squares (GLS) is the preferred estimator. GLS produces unbiased estimates of regression parameters and error terms and thus is well suited to dealing with

<sup>2</sup> The authors thank an anonymous reviewer for pointing this out.

<sup>3</sup> Most researchers consider a nonsignificant chi-square, or a ratio between chi-square and degrees of freedom of 2 or less, to indicate a good fit. Goodness-of-fit indexes and adjusted goodness-of-fit indexes above .90 and .80, respectively, usually imply an adequate fit. When a correlation matrix is used as input, root-mean-square residuals at or below .10 imply a reasonable fit. It is important to keep in mind, however, that the levels judged acceptable for fit statistics are simply experiential rules of thumb since the distributions of most of these statistics are unknown.

autocorrelated errors (Bollen, 1989; Hanushek & Jackson, 1977). LISREL allows estimation of a structural equation model with GLS rather than maximum likelihood. The fit statistics resulting from a GLS estimation changed very little from those in the maximum likelihood estimation ( $\chi^2/df = 0.54$ , goodness-of-fit index [GFI] = .959, adjusted GFI = .939, root-mean-square residual = .058). Furthermore, although the standard errors increased slightly over those in the maximum likelihood estimation, all parameters that were significant in the maximum likelihood estimation (reported below) remained so in the GLS estimation.

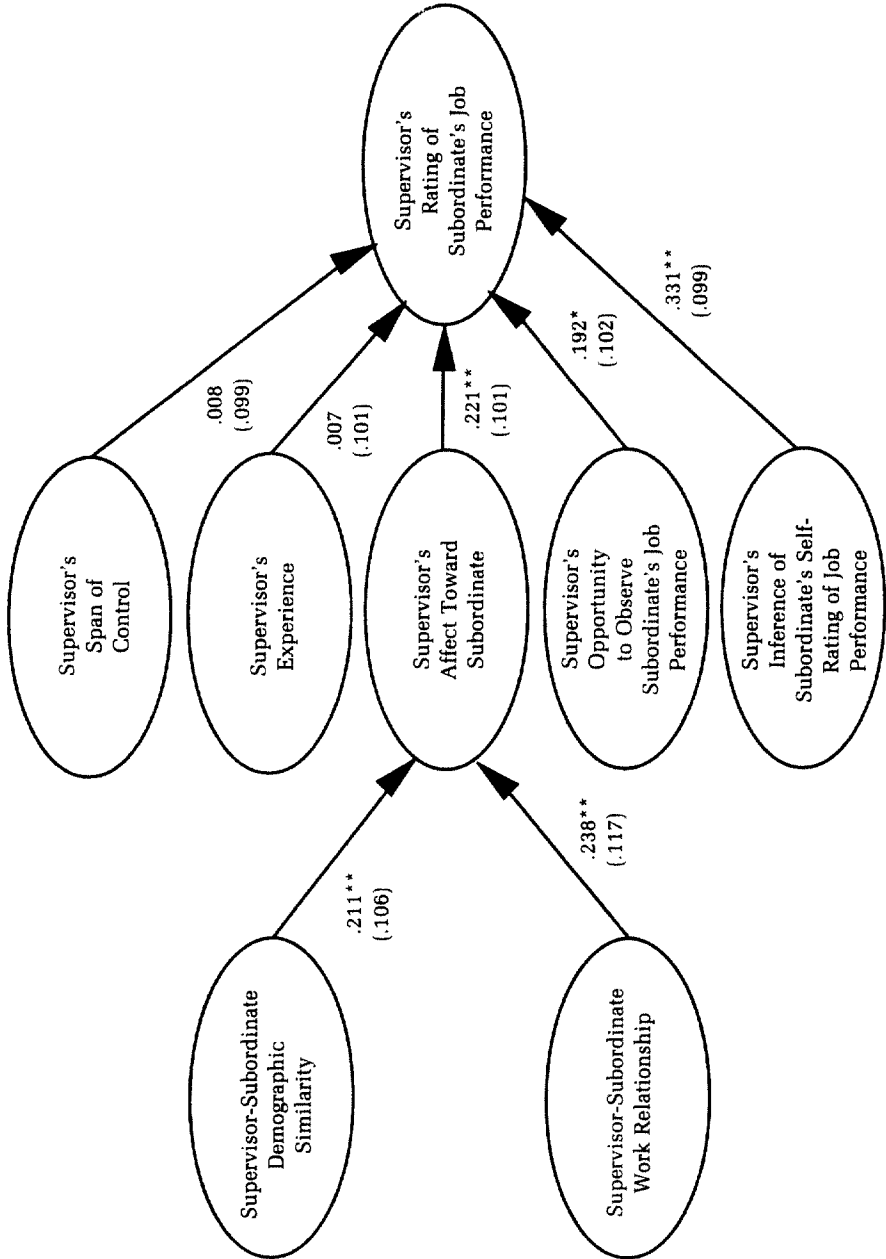
Figure 2 provides the maximum likelihood parameter estimates of the hypothesized model. The figure indicates that most of the hypothesized links were supported. Specifically, findings supported Hypothesis 1: the extent of supervisors' opportunity to observe subordinates' performance positively influenced performance ratings. Supervisors who had greater opportunity to observe subordinates' performance issued more positive ratings than those who had less opportunity to observe. Demographic similarity positively affected supervisors' affect toward subordinates, supporting Hypothesis 4. The more similar a supervisor and subordinate were with respect to age and experience, the more the supervisor reported liking the subordinate. Results also supported Hypothesis 5: the closeness of a supervisor-subordinate work relationship positively influenced the supervisor's affect toward the subordinate. The better the supervisor and subordinate worked together (as reported by the subordinate), the more the supervisor reported liking the subordinate. A supervisor's affect toward a subordinate in turn positively affected the supervisor's rating of the subordinate's job performance (Hypothesis 6). A supervisor who liked a particular subordinate was more likely to issue a positive performance rating than a supervisor who did not like a particular subordinate. Finally, results also supported Hypothesis 7: a supervisor's inference that a subordinate's self-rating of job performance was high positively influenced the supervisor's evaluation of the subordinate's job performance.

Only two of the hypothesized links in the model were not significant. A supervisor's span of control did not significantly influence performance ratings (Hypothesis 2), nor did the extent of supervisory experience (Hypothesis 3). Finally, both demographic similarity and supervisor-subordinate work relationship exerted significant, indirect effects on performance ratings as mediated through a supervisor's affect toward a subordinate, although the magnitudes of the effects (.047 and .053, respectively) were quite modest.

### Alternative Model Testing

Even though the hypothesized model fit the data very well, it was possible that other models might fit the data at least as well. Hayduk (1987) encouraged the testing of alternative, particularly nested, models. Nested models address the issue of whether the decrease in chi-square between a hypothesized model and a model with an additional causal link is significant. If it is, the hypothesis that the original model best represents the data is rejected. The proper model should include the added link.

**FIGURE 2**  
**Structural Estimates of Hypothesized Model<sup>a</sup>**



<sup>a</sup> Statistics are standardized path coefficients; standard errors are in parentheses.

\*  $p < .05$ , one-tailed test

\*\*  $p < .01$ , one-tailed test



Several nonhypothesized links seemed reasonable to investigate. One alternative is that supervisors like high performers, rather than (or in addition to) issuing favorable ratings to those they like. Finding such a link would call into question the validity of the present model, because a supervisor's affect toward a subordinate might be more a result than a cause of performance ratings. In response to this possibility, we estimated the hypothesized model, adding a link from supervisors' evaluation of subordinates' performance to supervisors' affect toward subordinates. Estimation of this model yielded a decrease in chi-square of only 0.11, with 1 less degree of freedom, an insignificant result indicating that adding the link from performance to affect does not significantly add to the explanatory power of the model. Thus, the results do not indicate that high performers are better liked, but rather that supervisors who like their subordinates appraise them more favorably, as hypothesized. It is important to note the limitation that we did not have data on true subordinate performance.

It is possible that the relationship between supervisors' inference regarding subordinates' self-ratings of performance and supervisors' ratings of subordinates' performance was observed because subordinates who rated themselves highly were truly high performers. In such a case, the supervisors' inferences would be confounded with actual performance. Although measures of true performance are difficult, if not impossible, to obtain (Bernardin & Beatty, 1984), especially for nurses, we did collect data on subordinates' self-ratings of performance. This controls for how the subordinate rates his or her own performance, which should bear some relationship to true performance (Vance, MacCallum, Coover, & Hedge, 1988). Adding a link from subordinates' self-ratings of performance to supervisors' ratings of subordinate performance did not significantly improve the fit of the model; the decrease in chi-square was 0.34, and the decrease in degrees of freedom was 1. Furthermore, adding the link did not change the significance of any coefficient in the model, including supervisors' inferences regarding subordinates' self-ratings. Thus, it appears that subordinates' self-ratings do not confound the effect of supervisors' inferences on performance ratings, although again, having measures of true performance would have been useful.

Controlling for subordinates' self-ratings of job performance also reduced the possibility that subordinates were biased in evaluating their supervisors' opportunity to observe their performance. For example, subordinates who received poor performance ratings might have responded that their supervisors rarely had the opportunity to observe their job performance as a way of rationalizing ratings that were too low in their judgment. However, because adding subordinates' self-ratings of performance to the model did not change the result, this possibility seems unlikely. It was a supervisor's belief about what a subordinate believed regarding her own performance, not what the subordinate actually believed, that influenced supervisors' ratings of subordinates' performance.

It was argued that demographic similarity and the supervisor-subordinate work relationship affected performance rating only through af-

fect. In fact, adding a direct link from each of those variables to performance rating did not significantly improve model fit; with respect to span of control, the decrease in chi-square was 1.33 (n.s.); with respect to direct opportunity to observe, the decrease in chi-square was 2.47 (n.s.).

Finally, since the average tenure in a current job was lower for the supervisors studied than for the subordinates, supervisors who were more experienced were more similar to their subordinates than the less experienced supervisors. This situation opened the possibility that supervisors' experience would be positively related to their affect toward subordinates. However, an added link from supervisors' experience to supervisors' affect toward subordinates approached but did not reach significance; the decrease in chi-square was 3.83 and the decrease in degrees of freedom was 1 (n.s.). We may have observed this effect because we had already controlled for overall demographic similarity, which included supervisory experience, in the model.

## DISCUSSION

Traditional conceptualizations of the performance-rating process imply that performance is a knowable and observable objective reality and that performance ratings are reasonable reflections of that reality. In the last decade, performance appraisal research has shifted away from a focus on instrumentation to a focus on psychological variables that underlie the appraisal process. Considerable research in the past ten years has investigated the role of cognitive processes in performance evaluation. However, researchers have called for more research on social and situational influences on the performance-rating process (Dipboye, 1985; Ilgen & Favero, 1985; Mitchell, 1983; Nathan et al., 1991; Wexley & Klimoski, 1984). Little empirical work has addressed those concerns, and the research that has been conducted has tended to investigate individual elements in isolation.

Only quite recently has performance-rating theory recognized the importance of the social contexts of ratings (Mitchell, 1983) and the notion that supervisors' ratings of subordinates' performance may be a socially constructed reality (Ferris & Judge, 1991). Because social contexts are admittedly multidimensional (Ferris & Mitchell, 1987), and their various dimensions or components are not necessarily orthogonal, efforts to represent social context effects must simultaneously incorporate multiple social and situational factors in order to capture the dynamics and totality of their impact. The present study tried to address those issues by proposing and testing a model of social influence in the performance-rating process that employed a number of social and situational influences.

Several specific social and situational influences on performance ratings were identified. Demographic similarity significantly influenced supervisors' affect toward subordinates, a result supporting the similarity-attraction paradigm (Byrne, 1969). Pfeffer (1983) and Tsui and O'Reilly (1989) have contended that the similarity-attraction paradigm generalizes to demographic characteristics. Our results support that proposition, although it would have been desirable to measure more demographic characteristics.

The effect of demographic similarity on supervisors' affect toward subordinates poses implications for the performance-rating process, because demographic similarity exerted a significant, albeit modest, indirect effect on performance rating as mediated through supervisor affect.

The supervisor-subordinate work relationship also exerted a significant effect on supervisors' affect toward subordinates, an effect consistent with arguments advanced by Nathan and colleagues (1991) and Wexley and Klimoski (1984). We hypothesized that the work relationship influenced performance rating only through supervisor affect toward a subordinate and found such an effect. Future research investigating the effects of supervisor-subordinate work relationships on performance ratings and supervisors' affect toward subordinates is warranted, given these results.

Previous researchers have demonstrated that supervisors' affect toward subordinates influences their performance ratings (Cardy & Dobbins, 1986; Kingstrom & Mainstone, 1985; Tsui & Barry, 1986; Wayne & Ferris, 1990). The links between supervisors' affect toward subordinates and performance ratings may have cognitive, information-processing implications. DeNisi and Williams (1988) suggested that affect influences the processing of performance information, and Isen and Baron (1991) shed light on this suggestion in their conceptualization of positive affect and its role in organizational settings. They argued that positive affect facilitates the recall of information stored in memory that possesses a positive affective tone. Thus, positive affect toward a subordinate should result in a supervisor's recalling more positive performance-related behaviors and evaluative impressions, which should lead to the supervisor's rating the subordinate's performance highly.

Supervisory experience, contrary to our hypotheses, did not exert a significant effect on performance ratings. Some research has suggested that experience positively influences performance ratings (Landy & Farr, 1980), perhaps reflecting a progression from toughness to leniency that people move through as they gain supervisory experience. However, our results do not support that hypothesis. One possible explanation for this result is that supervisory experience and demographic similarity were highly correlated in our data (see Table 1). It would be useful for future research to consider directly the psychological process that may cause experienced supervisors to issue more favorable ratings than inexperienced supervisors (several possibilities were suggested in this article's introduction), and under what conditions this process operates (Landy & Farr, 1983). It should be noted that our measurement of supervisory experience was tenure in the current position as supervisor, which could have understated amounts of total supervisory experience for some individuals.

A supervisor's opportunity to observe a subordinate's job performance significantly influenced performance ratings. Consistent with our hypothesis, greater opportunity to observe resulted in higher performance ratings. Ostensibly, opportunity to observe increases the ability to confirm the hypothesis of high performance (Snyder & Swan, 1978), given the general motivation to issue positive ratings reviewed earlier. This result and explana-

tion need to be directly substantiated by future research, however. If confirmed, they may call into question the hypothesis that greater opportunity to observe results in greater accuracy; instead, it may accentuate the positive bias already extant in performance ratings. Obviously, the ideal correction would not rest in supervisors' observing subordinates less, but in eliminating their motivation to search for positive information. Other researchers have made similar suggestions (e.g., Bass, 1956).

Supervisory span of control did not significantly influence performance ratings. The difference between span of control and opportunity to observe as reflections of ease of observation may explain this result. As noted earlier, a low span of control does not imply perfect opportunity to observe, perhaps explaining its weak effect on performance ratings.

Finally, a supervisor's inference regarding a subordinate's self-rating of performance significantly influenced the supervisor's rating of the subordinate's performance. Klimoski and Inks (1990) suggested that the susceptibility of supervisors to the wishes of subordinates is manifested by their yielding to the subordinates' self-ratings of performance. This accountability pressure may distort performance ratings. Such an effect might arise, according to Klimoski and Inks, because supervisors wish to avoid conflict. Practically and scientifically, it is important to understand what causes supervisors to make inferences regarding subordinates' self-ratings of performance, particularly in light of our results suggesting that supervisors' inferences exert a significant effect on performance ratings, even when the actual self-assessments made by subordinates are controlled for.

### **Limitations and Contributions**

Several limitations to this study need to be noted. Ideally, we would have employed a multidimensional measure of performance. Such a measure probably would have yielded a more accurate measure of performance. However, several points should be kept in mind about single-item measures of performance. First, many organizations use single-item, global measures of performance in personnel decisions (Bretz, Milkovich, & Read, 1992). Thus, our measure should be externally valid, resembling the actual evaluation process in organizations. Alternatively, since a number of specific evaluations may precede global ratings, the procedure used in many organizations may facilitate a more accurate assessment than that used in the present study. Second, the test-retest reliability estimates of the single-item measures revealed that the measures possessed acceptable stability, making their use less of a concern.

Results did not support all the hypothesized relationships. Supervisory span of control and experience did not significantly relate to subordinates' performance ratings. Thus, although results supported most of the relationships in the model, which had an acceptable overall fit to the data and was robust to several alternative models, the model itself, or our testing of it, possessed some inadequacies.

Another limitation of the present study is that the data were cross-

sectional. Although alternative model testing partially mitigated that limitation, such tests are always imperfect because it is not possible to test all alternative models. For example, it was not possible to test the reciprocal relationship between supervisor-subordinate work relationship and supervisors' affect toward subordinates because the reciprocal links could not be empirically identified (Hayduk, 1987; Jöreskog & Sörbom, 1989; Long, 1983).<sup>4</sup> Similarly, there was no unique indicator of either supervisors' ratings of subordinates' performance or subordinates' self-ratings, which prohibited valid tests of reciprocal causality. Ideally, we would have measured the variables over time to ensure that our causal ordering was correct. Future research investigating those relationships over time would make a contribution. Replication of these results among different workers would also be of use. Whereas we have no reason to believe that the relationships observed are unique to the group studied, this obviously is an empirical question that could be answered with more occupationally heterogeneous data.

Furthermore, this study did not meet all necessary assumptions for causal inference. James and colleagues (1982), for example, identified ten assumptions necessary for causal inference. As Campbell (1990), Meehl (1978, 1990), and others have pointed out, rarely is it feasible for all the assumptions necessary for causal inference to be met. To be sure, it is an important goal, but in many cases it is not met, particularly when researchers are dealing with soft data, such as relations between psychological variables (Meehl, 1978). Although the results of the present study largely corroborate our model, the failure to satisfy all the assumptions necessary for causal inference suggests that the results should be viewed as suggestive rather than a direct demonstration of causality. Because the model is easily falsifiable (Popper, 1959), future research substantiating the results is needed before anything but tentative causal inferences can be drawn. Although we have discussed influences in the present study, given the equivocal satisfaction of all causal assumptions, it may be best to think of the influences as relationships until corroborative evidence is forthcoming. As Bobko (1990) pointed out, although structural equation modeling aids in interpreting theoretical models, replication and cross-validation of results is often essential. Indeed, this point is relevant with respect to the present study.

Relatedly, because of the failure to meet the assumption of weak causal order, the directional arrows indicated in the model and the alternative model tests should be viewed as suggestive only. The results are clear with respect to the sign and magnitude of the relationships in the model, but they

---

<sup>4</sup> Although recursive, or unidirectional, models are always identified, identification of nonrecursive (reciprocal) models requires that exogenous variables be included in any estimation of nonrecursive links. Thus, it is impossible to identify all nonrecursive relationships in a model because there would be no exogenous variables. Because supervisors' affect toward subordinates and supervisors' rating of subordinates' performance were the only endogenous variables in the model influenced by exogenous variables, this was the only nonrecursive link in the model that could be tested.

can only suggest rather than establish causality. This presents a clear, although difficult, opportunity for future research to undertake rigorous investigations of causality concerns.

Finally, although both supervisors and subordinates gave responses for different variables in this study, precluding method bias for most of the relationships in the model, the variables involved in two of the relationships were assessed from data from supervisors only. Common method variance could have inflated the relationship between supervisors' affect toward subordinates and supervisors' ratings of subordinates' performance and the relationship between supervisors' inference of subordinates' self-ratings of performance and supervisors' ratings of subordinates' performance.

Overall, a strength of the present study was that we obtained measures from two sources, supervisors and subordinates, reducing the possibility that the relations observed were the results of single-source covariance. Furthermore, considering a number of social and situational influences simultaneously reduced the possibility that the results were biased by the omission of variables (James et al., 1982). An even more rigorous procedure would have employed both supervisors' and subordinates' reports on all variables to control all possible response artifacts. Not surprisingly, doing so was not feasible, given the data collection constraints imposed by the organization under study. Nonetheless, our use of multiple-source information should allow more confidence to be placed in the validity of our interpretations of results.

The present study contributed to understanding social elements in the performance-rating process in several ways. First, we found overall support for the importance of several social and situational elements to performance evaluation decisions. Mitchell's (1983) trichotomy of social and task variables was also generally supported. Social influence is more than a single isolated effect; it is manifested by an array of contextual and interpersonal elements. Although the present study is certainly not a comprehensive survey of social and situational elements, the number of variables considered improves upon past research by reflecting the multidimensional and dynamic nature of social contexts. Cumulatively, the findings suggest that the social context does affect the performance-rating process. Several of the specific relationships in the model tested are also of interest. There is a scant literature on supervisory opportunity to observe subordinates' performance and supervisory inference regarding subordinates' self-ratings of performance. The significance of those variables is interesting in its own right and should stimulate future research on them.

In sum, the present study investigated the relationship of social context to the performance-rating process by testing a model of social influence that employed several underinvestigated variables. By concurrently considering the effects of several key aspects of social contexts, we found support for the efficacy of social and situational processes in the performance-rating process. We hope future researchers will continue along these lines by expanding the variables studied and providing a deeper assessment of the causal

relationships among those variables, and thus a more informed understanding of the performance-rating process.

## REFERENCES

- Bass, B. M. 1956. Reducing leniency in merit ratings. *Personnel Psychology* 9: 359–369.
- Bentler, P. M. 1985. *Theory and implementation of EQS: A structural equations program*. Los Angeles: BMDP Statistical Software.
- Bernardin, H. J., & Beatty, R. W. 1984. *Performance appraisal: Assessing human behavior at work*. Boston: Kent.
- Bobko, P. 1990. Multivariate correlational analysis. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2d ed.), vol. 1: 637–686. Palo Alto, CA: Consulting Psychologist Press.
- Bollen, K. A. 1989. *Structural equations with latent variables*. New York: Wiley.
- Borman, W. C. 1978. Exploring upper limits of reliability and validity in job performance ratings. *Journal of Applied Psychology*, 63: 135–144.
- Bretz, R. D., Jr., Milkovich, G. T., & Read, W. 1992. The current state of performance appraisal research and practice: Concerns, directions, and implications. *Journal of Management*, 18: 321–352.
- Brooke, P. P., Russell, D. W., & Price, J. L. 1988. Discriminant validation of measures of job satisfaction, job involvement, and organizational commitment. *Journal of Applied Psychology*, 73: 139–145.
- Byrne, D. 1961. The influence of propinquity and opportunities for interaction on classroom relationships. *Human Relations*, 14: 63–69.
- Byrne, D. 1969. Attitudes and attraction. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, vol. 4: 35–90. New York: Academic Press.
- Campbell, J. P. 1990. The role of theory in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), *Handbook of industrial and organizational psychology* (2d ed.), vol. 1: 39–73. Palo Alto, CA: Consulting Psychologist Press.
- Cardy, R. L., & Dobbins, G. H. 1986. Affect and appraisal accuracy: Liking as an integral dimension in evaluating performance. *Journal of Applied Psychology*, 71: 672–678.
- Cooper, W. H. 1981. Ubiquitous halo. *Psychological Bulletin*, 90: 218–244.
- Dansereau, F., Jr., Graen, G., & Haga, W. J. 1975. A vertical dyad linkage approach to leadership within formal organizations: A longitudinal investigation of the role making process. *Organizational Behavior and Human Performance*, 13: 46–78.
- DeCotiis, T., & Petit, A. 1978. The performance appraisal process: A model and some testable propositions. *Academy of Management Review*, 3: 635–646.
- DeNisi, A. S., Cafferty, T. P., & Meglino, B. M. 1984. A cognitive view of the performance appraisal process: A model and research propositions. *Organizational Behavior and Human Performance*, 33: 360–396.
- DeNisi, A. S., & Williams, K. J. 1988. Cognitive approaches to performance appraisal. In G. R. Ferris & K. M. Rowland (Eds.), *Research in personnel and human resources management*, vol. 6: 109–155. Greenwich, CT: JAI Press.
- Dienesch, R. M., & Liden, R. C. 1986. Leader-member exchange model of leadership: A critique and further development. *Academy of Management Review*, 11: 618–634.
- Dipboye, R. L. 1985. Some neglected variables in research on discrimination in appraisals. *Academy of Management Review*, 10: 116–127.

- Duchon, D., Green, S. G., & Taber, T. D. 1986. Vertical dyad linkage: A longitudinal assessment of antecedents, measures, and consequences. *Journal of Applied Psychology*, 71: 56–60.
- Feldman, J. M. 1981. Beyond attribution theory: Cognitive processes in performance appraisal. *Journal of Applied Psychology*, 66: 127–148.
- Ferris, G. R., & Judge, T. A. 1991. Personnel/human resources management: A political influence perspective. *Journal of Management*, 17: 447–488.
- Ferris, G. R., & Mitchell, T. R. 1987. The components of social influence and their importance for human resources research. In K. M. Rowland & G. R. Ferris (Eds.), *Research in personnel and human resources management*, vol. 5: 103–128. Greenwich, CT: JAI Press.
- Ferris, G. R., Judge, T. A., Rowland, K. M., & Fitzgibbons, D. E. In press. Subordinate influence and the performance evaluation process: Test of a model. *Organizational Behavior and Human Decision Processes*.
- Ghiselli, E. E., Campbell, J. P., & Zedeck, S. 1981. *Measurement theory for the behavioral sciences*. San Francisco: Freeman.
- Graen, G. 1976. Role-making processes within complex organizations. In M. D. Dunnette (Ed.), *Handbook of industrial and organizational psychology*: 1201–1245. Chicago: Rand McNally.
- Graen, G., & Scheimann, W. 1978. Leader-member agreement: A vertical dyad linkage approach. *Journal of Applied Psychology*, 63: 206–212.
- Groshen, E. L., & Krueger, A. B. 1990. The structure of supervision and pay in hospitals. In R. G. Ehrenberg (Ed.), *Do compensation policies matter?*: 134–146. Ithaca, NY: ILR Press.
- Hanushek, E. A., & Jackson, J. E. 1977. *Statistical methods for social scientists*. Orlando, FL: Academic Press.
- Hastie, R., & Park, B. 1986. The relationship between memory and judgments depends on whether the judgment task is memory-based or on-line. *Psychological Review*, 93: 258–268.
- Hayduk, L. A. 1987. *Structural equation modeling with LISREL: Essentials and advances*. Baltimore: Johns Hopkins.
- Ilgen, D. R., & Favero, J. L. 1985. Limits in generalization from psychological research to performance appraisal processes. *Academy of Management Review*, 10: 311–321.
- Ilgen, D. R., & Feldman, J. M. 1983. Performance appraisal: A process focus. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior*, vol. 5: 141–197. Greenwich, CT: JAI Press.
- Isen, A. M., & Baron, R. A. 1991. Positive affect as a factor in organizational behavior. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior*, vol. 13: 1–53. Greenwich, CT: JAI Press.
- James, L. R., Mulaik, S. A., & Brett, J. M. 1982. *Causal analysis: Assumptions, models, and data*. Beverly Hills, CA: Sage.
- Jones, G. 1984. Task visibility, free riding, and shirking: Explaining the effect of structure and technology on employee behavior. *Academy of Management Review*, 9: 684–695.
- Jöreskog, K. G., & Sörbom, D. 1989. *LISREL 7: A guide to the program and applications*. Chicago: SPSS.
- Judge, T. A., & Chandler, T. D. 1990. *Individual-level determinants of the propensity to shirk*. Working paper #90-26, Center for Advanced Human Resource Studies, Cornell University, Ithaca, NY.
- Kahn, L. M., & Sherer, P. D. 1990. Contingent pay and managerial performance. In R. G. Ehrenberg (Ed.), *Do compensation policies matter?*: 107–120. Ithaca, NY: ILR Press.



- Kingstrom, P. O., & Mainstone, L. E. 1985. An investigation of the rater-ratee acquaintance and rater bias. *Academy of Management Journal*, 28: 641–653.
- Klimoski, R., & Inks, L. 1990. Accountability forces in performance appraisal. *Organizational Behavior and Human Decision Processes*, 45: 194–208.
- Landy, F. J., & Farr, J. L. 1980. Performance rating. *Psychological Bulletin*, 87: 72–107.
- Landy, F. J., & Farr, J. L. 1983. *The measurement of work performance: Methods, theory, and applications*. New York: Academic Press.
- Latane, B., Williams, K., & Harkins, S. 1979. Many hands may lighten work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology*, 37: 822–832.
- Long, J. S. 1983. *Covariance structure models: An introduction to LISREL*. Beverly Hills, CA: Sage.
- Lord, R. G., & Maher, K. J. 1989. Cognitive processes in industrial and organizational psychology. In C. L. Cooper & I. Robertson (Eds.), *International review of industrial and organizational psychology*, vol. 4: 49–91. London: Wiley.
- McFillen, J. M., & New, J. R. 1979. Situational determinants of supervisor attributions of behavior. *Academy of Management Journal*, 22: 793–809.
- Meehl, P. E. 1978. Theoretical risks and tabular risks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology*, 4: 806–834.
- Meehl, P. E. 1990. Why summaries of research on psychological theories are often uninterpretable. *Psychological Reports*, 66: 195–244.
- Miles, R. E. 1964. Attitudes toward management theory as a factor in managers' relationships with their superiors. *Academy of Management Journal*, 7: 308–314.
- Mitchell, T. R. 1983. The effects of social, task, and situational factors on motivation, performance, and appraisal. In F. Landy, S. Zedeck, & J. Cleveland (Eds.), *Performance measurement and theory*: 39–59. Hillsdale, NJ: Erlbaum.
- Motowidlo, S. J. 1986. Information processing in personnel decisions. In K. M. Rowland & G. R. Ferris (Eds.), *Research in personnel and human resources management*, vol. 4: 1–44. Greenwich, CT: JAI Press.
- Nathan, B. R., & Alexander, R. A. 1985. The role of inferential accuracy in performance rating. *Academy of Management Review*, 10: 109–115.
- Nathan, B. R., & Lord, R. G. 1983. Cognitive categorization and dimensional schemata: A process approach to the study of halo in performance ratings. *Journal of Applied Psychology*, 68: 102–114.
- Nathan, B. R., Mohrman, A. M., Jr., & Milliman, J. 1991. Interpersonal relations as a context for the effects of appraisal interviews on performance and satisfaction: A longitudinal study. *Academy of Management Journal*, 34: 352–369.
- Pfeffer, J. 1983. Organizational demography. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior*, vol. 5: 299–357. Greenwich, CT: JAI Press.
- Popper, K. R. 1959. *The logic of scientific discovery*. New York: Basic Books.
- 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.
- Rowe, P. M. 1989. Unfavorable information and interview decisions. In R. W. Eder & G. R. Ferris (Eds.), *The employment interview: Theory, research, and practice*: 77–89. Newbury Park, CA: Sage.
- Saegert, S. C., Swap, W., & Zajonc, R. B. 1973. Exposure, context, and interpersonal attraction. *Journal of Personality and Social Psychology*, 25: 234–242.

- Scarpello, V., & Campbell, J. P. 1983. Job satisfaction: Are all the parts there? *Personnel Psychology*, 36: 577–600.
- Schneider, D. J. 1991. Social cognition. In M. R. Rozenzweig & L. W. Porter (Eds.), *Annual review of psychology*, vol. 42: 527–561. Palo Alto, CA: Annual Reviews.
- Senger, J. 1971. Managers' perceptions of subordinates' competence as a function of personal value orientation. *Academy of Management Journal*, 14: 415–423.
- Snyder, M., & Swan, W. B. 1978. Hypothesis-testing processes in social interaction. *Journal of Personality and Social Psychology*, 36: 1202–1212.
- Strull, T. K., & Wyer, R. S. 1989. Person memory and judgment. *Psychological Review*, 96: 58–83.
- Stone, T. H. 1973. An examination of six prevalent assumptions concerning performance appraisal. *Public Personnel Management*, 5: 408–414.
- Tsui, A. S., & Barry, B. 1986. Interpersonal affect and rating errors. *Academy of Management Journal*, 29: 586–599.
- Tsui, A. S., & O'Reilly, C. A., III. 1989. Beyond simply demographic effects: The importance of relational demography in superior-subordinate dyads. *Academy of Management Journal*, 32: 402–423.
- Turban, D. B., & Jones, A. P. 1988. Supervisor-subordinate similarity: Types, effects, and mechanisms. *Journal of Applied Psychology*, 73: 228–234.
- Vance, R. J., MacCallum, R. C., Coover, M. D., & Hedge, J. W. 1988. Construct validity of multiple job performance measures using confirmatory factor analysis. *Journal of Applied Psychology*, 73: 74–80.
- Wayne, S. J., & Ferris, G. R. 1990. Influence tactics, affect, and exchange quality in supervisor-subordinate interactions: A laboratory experiment and a field study. *Journal of Applied Psychology*, 75: 487–499.
- Wexley, K. N., & Klimoski, R. 1984. Performance appraisal: An update. In K. M. Rowland & G. R. Ferris (Eds.), *Research in personnel and human resources management*, vol. 2: 35–79. Greenwich, CT: JAI Press.
- Wexley, K. N., Alexander, R. A., Greenawalt, S. P., & Couch, M. A. 1980. Attitudinal congruence and similarity as related to interpersonal evaluations in manager-subordinate dyads. *Academy of Management Journal*, 23: 320–330.
- Zajonc, R. B. 1980. Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35: 151–175.

**Timothy A. Judge** is an assistant professor of personnel and human resource studies at the Center for Advanced Human Resource Studies, School of Industrial and Labor Relations, Cornell University. He received his Ph.D. degree from the Institute of Labor and Industrial Relations, University of Illinois at Urbana-Champaign. His research interests include dispositional influences, job attitudes and employee withdrawal behaviors, staffing, and social and political influences in human resources management.

**Gerald R. Ferris** is a professor of labor and industrial relations and of business administration at the University of Illinois at Urbana-Champaign. He received his Ph.D. degree in business administration from the University of Illinois at Urbana-Champaign. His research interests include performance evaluation and social and political influences in human resource systems.