Managerial Discretion in the Use Of Self-Ratings in an Appraisal System: The Antecedents And Consequences

Edward Inderrieden
Marquette University, edward.inderrieden@marquette.edu

Robert E. Allen

Timothy Keaveny
Marquette University, timothy.keaveny@marquette.edu

Managerial Discretion in the Use of Self-ratings in an Appraisal System: The Antecedents and Consequences

Edward J. Inderrieden
Associate Professor of Management
Marquette University

Robert E. Allen
Allen and Associates

Timothy J. Keaveny
Professor of Management
Marquette University

The strong position taken by Campbell and Lee (1988) regarding the limited usefulness of self-appraisals as an evaluation tool (Folger and Lewis, 1995) has elicited doubts about the use of self-ratings in the performance appraisal process. The leniency effect associated with self-ratings (Farh and Dobbins, 1989; Harris and Schaubroeck, 1988) raises concern about the legitimacy of using self-ratings, and sets the stage for possible disagreement between superiors and subordinates. However, studies of performance appraisal systems consistently indicate the importance of ratee involvement in the appraisal process (Landy et al., 1978; Dipboye and de Pontbriand, 1981; Greenberg, 1986; Folger and Konovsky, 1989; Cawley et al., 1998), providing indirect support for the use of self-appraisals. Furthermore, the traditional top-down appraisal system is clearly inconsistent with the trends toward employee involvement, suggesting that a performance appraisal interview based on an employee’s self-rating is more appropriate than a discussion based only on the performance evaluation conducted by a manager (Meyer, 1991; Baruch, 1996).

Early work by Bassett and Meyer (1968) reported that the subordinates who self-rated were generally positive about the self-appraisal approach and were more likely to see their manager’s criticism of their performance as warranted when compared to those subordinates who were not given an opportunity to self-rate. More recently, Roberson et al. (1993) conducted a field experiment to in-
vestigate the impact of a “both-rate” self-appraisal (superior and subordinate independently completed ratings) where subordinates were randomly assigned to a “self-rate” or “no self-rate” condition. Roberson et al. (1993) hypothesized that individuals assigned to the self-appraisal condition would report greater satisfaction with the appraisal interview, greater understanding of their manager’s expectations and perceive the procedure to be fairer than those individuals who did not self-evaluate. These hypotheses were not supported. However, an unexpected finding emerged. Roberson et al. (1993) found that individuals in the “no self-rate” group who performed an informal self-rating were more satisfied with the appraisal, and felt their ratings were fairer than individuals who did not self-rate.

The results of the Roberson et al. (1993) study are not consistent with the oft-cited work of Bassett and Meyer (1968). There is a significant difference in the research design of the two studies. In the Bassett and Meyer study, individuals who self-evaluated did not receive a performance evaluation from their superiors; performance appraisal discussions were based solely on the subordinate’s self-evaluation. In the Roberson et al. (1993) study, appraisal discussions were centered on the self-appraisals and the formal evaluation of their superiors. As noted by Campbell and Lee (1988), this situation may create the possibility for conflict. Roberson et al. (1993) surmise that formal self-ratings tend to raise employee expectations, while informal self-ratings may not have the same effect. In addition, formal self-ratings may be instituted in an environment where little employee participation has been allowed, while informal self-ratings are more likely to occur when individuals feel comfortable providing input. Given the concerns related to leniency effects and increased expectations of self-ratings, along with the mixed findings regarding individuals’ reactions to self-appraisals, the question is raised as to under what conditions might self-ratings be useful.

Our study differs from previous investigations of self-ratings in that superiors in the organization we studied were given full control over whether or not subordinates performed a self-rating of their performance. All superiors had received training related to the performance appraisal process. Superiors were required to conduct an appraisal interview following their formal review and rating of subordinates’ performance. Superiors were informed about the benefits of self-ratings but were not required to have subordinates evaluate their own performance. A superior could request a subordinate to self-rate and include this rating in the appraisal discussions; this was up to the discretion of each superior. To our knowledge this is the first study that examines self-ratings in a field setting where superiors are given the discretion whether or not to incorporate self-ratings in the performance appraisal process. We chose to use information from the subordinate’s perspective because we are specifically interested in the interactions between the manager-subordinate dyad. Research on the Leader-Member Exchange (LMX) model shows that managers do treat subordinates differently (Wayne et al., 2002; Liden et al., 1993); however, Larson et al. (1986) noted that managers are likely to report the average feedback they
Inderrieden, Allen and Keaveny give to all subordinates rather than to focus on the feedback given to any particular subordinate.

The nature of this field setting allowed us to examine three research questions. The first question concerned the antecedents of self-ratings, that is, the conditions under which superiors asked subordinates to provide self-ratings of their performance. The second research question examined the relationship between self-ratings and the type of performance appraisal interview that was conducted by the superior. The third research question addressed the outcomes associated with the use of self-ratings. By examining these questions, it is possible to develop a better understanding of the contextual factors associated with the use of self-ratings, the implications of the use of self-ratings for other aspects of the appraisals, and the outcomes associated with the use of self-ratings. We first review the literature and develop a model of the antecedents and consequences of self-ratings in a performance appraisal system. Next, information is provided on the research setting, sample, procedures and measures. Finally, the results and a discussion of the research findings, including limitations and possible future research, are presented.

A study by Lawler et al. (1984) reported that 90% of appraisers and 86% of subordinates indicated self-appraisals should be an important part of performance appraisal; however, less than 50% of the superiors actually employed them. This disparity raises a number of very important questions concerning the factors that encourage or discourage superiors to solicit self-appraisals from their subordinates. Given the importance of the superior in the appraisal process, it is interesting that research on self-ratings has essentially ignored the role of the superior in the appraisal process. As noted by Murphy and Cleveland (1991), in most organizations superiors are assigned the direct responsibility of conducting performance evaluations and, with the exception of the evaluation form, enjoy considerable leeway in performing their duties.

While superiors in general might believe that self-appraisals are useful, as noted by Lawler (1984), we have little information about characteristics of the situation that encourage or discourage the use of self-ratings. Important questions remain as to when do superiors ask subordinates to evaluate their own performance and what is the subsequent impact of these self-ratings on the appraisal process? Our model of the antecedents and consequences of manager-initiated self-ratings is presented in Figure I. We propose that work context and task characteristics of the subordinate’s job, along with the preferred management style of the manager, will influence whether or not an individual is given the opportunity to self-evaluate his or her performance. In turn, individuals who are given an opportunity to self-rate will be more likely to report receiving an appraisal interview that focuses on past performance and ways to improve future performance, resulting in greater job satisfaction and perceptions that the performance appraisal is fair.

Antecedents of Self-Ratings

One of the acknowledged problems associated with the use of self-ratings is the potential for conflict or disagreement (Campbell and Lee, 1988; Harris and Schaubroeck,
At the same time Folger and Lewis (1995) have proposed that self-appraisals could be used to enhance due process in performance appraisal systems, thus increasing employees’ perceptions that appraisals are just and fair. This perceived contradiction might arise because self-appraisals are often viewed as an isolated event rather than as an integral component of the appraisal process. Managers may feel they are being forced to ask subordinates to self-evaluate in situations where they perceive the evaluations are unnecessary or incompatible with their management style. Under what circumstances are managers more likely to feel comfortable asking subordinates to self-evaluate?

The importance of valid and reliable measurement in performance appraisal, especially when used to allocate rewards, is well documented (Milkovich and Newman, 1999). In order to accurately assess performance, superiors must have personal knowledge about the work their subordinates are performing. Several studies have reported the importance of rater’s familiarity with ratee’s work as a determinant of fair appraisals (Greenberg, 1986; Landy et al., 1978). Opportunity to observe a subordinate’s performance should help to improve the accuracy of one’s observation. Previous research has reported more agreement in self-ratings when dimensions of work are observable (Warr and Bourne, 1999; Wohlers and London, 1989). While this reduces the possibility for conflict, it also reduces the need to obtain information from another source. Thus, we expect that superiors who are not able to observe the performance of their subordinates, because of the nature of the task or the size of the work group, will be more likely to ask them to self-rate.

The characteristics of the task must be taken into consideration when assessing performance (Kane and Lawler, 1979; Folger et al., 1992). Complex jobs, where it is harder to identify a means-end relationship, make it more difficult to establish clear, reliable and valid measurements of performance (Folger et al., 1992). Superiors have less information about the activities of subordinates, thus increasing the opportunity for conflict. Highly routine work, where subordinates experience little autonomy, should permit more valid and reliable measurement of performance, thus reducing the likelihood for disagreement concerning performance. Thus, we expect that superiors will be more likely to invite subordinates to self-rate when subordinates’ experience a higher degree of autonomy and variety in their tasks.

Given that performance evaluations are often used for the purpose of distributing raises and making promotion decisions, the appraisal process can be viewed as threatening to both superiors (raters) and subordinates (ratees). While the performance appraisal literature clearly documents that subordinates desire to offer information concerning their performance (Dipboye and de Pontbriand, 1981; Folger and Konovsky, 1989; Greenberg, 1986; Lawler et al., 1984), superiors might not encourage subordinates to self-evaluate if they perceive the potential for conflict or disagreement. Furthermore, it is unlikely that superiors who are directive in their management style would freely provide subordinates the opportunity to self-evaluate (Wexley and Klimoski, 1984). In general, su-
Self-ratings in an Appraisal System

Superiors that have a preferred style of interacting would be uncomfortable deviating from that style (Carroll and Schneier, 1982). A superior who has a good relationship with a subordinate and is characterized as participative may be more likely to see self-evaluations as useful.

Folger et al. (1992) indicate that prior to the actual evaluation of performance it is important that objectives and standards be clearly stated, subordinates have the opportunity to voice their opinions about the standards, and performance feedback be provided on a consistent basis throughout the appraisal period. Greenberg (1986) and Folger and Konovsky (1989) reported that familiarity with job performance and job knowledge is an important procedural aspect of the appraisal process. Effective performance appraisal systems are dependent upon subordinates being knowledgeable of and in agreement with performance criteria associated with their jobs. Ensuring that individuals are informed about job standards may impact a subordinate’s opportunity to self-evaluate in two ways. First, superiors who involve subordinates in setting job standards are exhibiting a participative management style. Asking an individual to self-rate would be consistent with that style. Second, by involving individuals in the setting of job standards a superior would ensure that subordinates would better understand what is expected of them, thus reducing the potential for disagreement. Subordinates’ perceived system knowledge was associated with a reduction in the discrepancy between self-and supervisory ratings (Williams and Levy, 1992). Thus, we expect that superiors who establish clear standards of desired performance and provide performance feedback throughout the appraisal period will be more likely to have their subordinates evaluate their own performance.

As shown in Figure I, we propose that work context and characteristics of the subordinate’s job, along with the managerial style of the rater, influence whether or not subordinates are given the opportunity to self-rate.

Hypothesis 1a: Subordinates who were given the opportunity to self-rate will report their superiors had less opportunity to observe their performance than subordinates who did not self-rate.

Hypothesis 1b: Subordinates who were given the opportunity to self-rate will be from larger work groups than subordinates who did not self-rate.

Hypothesis 1c: Subordinates who were given the opportunity to self-rate will report a higher degree of task variety in their jobs than subordinates who did not self-rate.

Hypothesis 1d: Subordinates who were given the opportunity to self-rate will report a higher degree of autonomy in their job than subordinates who did not self-rate.

Hypothesis 1e: Subordinates who were given the opportunity to self-rate will express a greater knowledge of performance standards than subordinates who were not given the opportunity to self-evaluate.

Hypothesis 1f: Subordinates who were given the opportunity to self-rate received more performance feedback throughout the appraisal period than subordinates who were not given the opportunity to self-evaluate.

Effects of Self-Rating on the Appraisal Interview and Reactions to the Appraisal Process

A concern with using self-ratings is that subordinates have increased expectations regarding the impact they might have on their performance rating (Folger and Greenberg, 1985; Roberson et al., 1993). In the present study, self-ratings were used as a basis for discussion, rather than as input into the superior’s formal evaluation of performance. Superiors had al-
ready completed their formal evaluation of the subordinates’ performance. The work of Korsgaard and Roberson (1995) refers to this type of involvement as non-instrumental participation; subordinates may perceive their comments are welcome but they do not expect to actually influence the outcomes (Greller, 1978; Burke et al., 1978).

As noted in our earlier discussion, we propose that managers ask subordinates to self-rate under two general conditions: (1) when managers perceive a need for additional information and (2) when self-ratings would be consistent with their general management style. We would expect that self-ratings would then serve as a basis for discussing current performance and ways to improve performance. Basset and Meyer (1968) maintained that appraisal interviews are more productive and constructive when based on self-ratings. Nathan et al. (1991) suggested that participation in the appraisal process gives subordinates control over their work performance and signals to subordinates that their supervisors are fair. In an approach similar to that employed by Nathan et al. (1991), this study argues that the social context, operationalized as the opportunity to self-rate, affects the content of the performance appraisal interview. Employees who self-rated will report during their appraisal interview they were given a greater opportunity to discuss issues regarding past performance and opportunity to improve than employees who did not self-rate.

Focusing on the outcomes of the appraisal system, Folger and Lewis (1995) have proposed that self-appraisals could be used to enhance due process in performance appraisal systems, thus increasing employees’ perceptions that appraisals are just and fair. Consistent with this perspective, Farh et al. (1988) reported that self-raters found the appraisal process more acceptable, accurate, fair, and, in general, more effective than traditional approaches to performance appraisal. Similarly, Bassett and Meyer (1968) found that appraisal interviews based on self-ratings were reported to be more productive and constructive than when based solely on superior-prepared reviews. Participation in the appraisal interview has been shown to be a function of the person conducting the interview rather than reason for the review (Greller, 1998). One might expect that managers who encourage subordinates to self-rate would be more likely to encourage involvement in the appraisal interview. While not directly concerned with self-ratings, researchers have demonstrated that the content of performance appraisal reviews could influence important organizational outcomes such as future job performance and satisfaction (Nathan et al., 1990; Black and Gregerson, 1997).

However, researchers have not looked at the relationship between self-ratings, appraisal interviews and satisfaction measures in the same study. While Farh et al. (1988) found self-ratings to be positively related to accurate fair appraisals, we propose that the appraisal interview discussion mediates the relationship between self-rating and the outcome variables. One of the primary concerns about self-ratings has been the presence of leniency effects (Farh and Dobbins, 1989; Harris and Schaubroeck, 1988). Self-ratings increase the opportunity to discuss performance issues in the appraisal interview, thus increasing perceived
fairness of the ratings and job satisfaction. The mere opportunity to self-rate does not impact fairness; rather self-ratings are used as a basis for discussing performance.

Hypothesis 2: Self-raters will indicate they were given a greater opportunity during the appraisal interview to discuss issues related to past performance and ways to improve future performance than individuals who did not self-rate.

Hypothesis 3a: The performance appraisal interview discussion will mediate the relationship between opportunity to self-evaluate and perceived fairness of performance ratings.

Hypothesis 3b: The performance appraisal interview discussion will mediate the relationship between opportunity to self-evaluate and job satisfaction.

Alternative Models

In order to effectively evaluate a hypothesized model, researchers have suggested the testing of alternative, nested models (Hayduk, 1987; Kelley, 1996). Our model as shown in Figure I tests fully mediated pathways from the contextual and managerial style variables to the appraisal interview via self-ratings, and from self-rating to job satisfaction and fairness of ratings through the appraisal interview session. Prior research evidence suggests that it is appropriate to investigate the existence of partial-rather than full-mediation.

Self-ratings may serve as a partial mediator between the contextual and managerial style variables and the appraisal interview discussion rather than a full mediator as hypothesized. Previous research has indicated that managers want to give employees positive evaluations (Klimoski and Inks, 1990); however, if managers are unfamiliar with subordinate’s work they may not have adequate information. The appraisal process becomes a search for more information (Judge and Ferris, 1993) in order to justify the ratings. Managers may choose to use self-evaluations and the appraisal interview as a way of obtaining additional information. Thus, the existence of a partially-mediated relationship between variables measuring contextual/job characteristics and the appraisal interview discussion should be assessed.

Work by Ilgen et al. (1981) examined the perceptions of subordinates concerning their formal appraisal session. Indices of feedback received before the appraisal sessions were strongly related to the subordinates’ perceptions of the atmosphere of the interview (e.g., extent to which the person felt relaxed, not tense), and the helpfulness and specificity of supervisor’s remarks. Thus, these findings suggest that interactions with the supervisor prior to the appraisal interview will have a direct, positive effect on opportunity to self-rate and on the nature of the interview that an individual receives.

In Figure I we hypothesized that the effect of self-ratings on job satisfaction and fairness of ratings will be fully mediated by the appraisal interview discussion. An argument for partial mediation can be made. Korsgaard and Roberson (1995) indicated that self-appraisals have a greater impact on perceived fairness when they are conducted prior to the supervisor’s evaluation. In the current situation, employees were asked to self-evaluate after the manager had already completed the formal evaluation. According to Korsgaard and Roberson (1995) this would be
classified as non-instrumental voice, where employees value the opportunity to express their opinion regardless of impact. Thus, we would expect that self-ratings would be have an impact on the appraisal discussion and a separate, positive effect on the outcome measures of job satisfaction and perceived fairness of ratings. Alternatively, from an instrumental voice perspective, employees who are given the opportunity to self-rate may become frustrated over their inability to influence the formal appraisal ratings. In this situation, we would expect a positive relationship between self-ratings and the performance appraisal discussion but a negative relationship between self-ratings and the outcome variables of perceived fairness of ratings and job satisfaction.

METHODS

Data were collected at a state university in the western United States. Superiors were required to complete a formal evaluation of their subordinates' performance during the last two weeks of November and the first two weeks of December. The performance appraisal form focused on eight categories: position knowledge, quality of work, quantity of work, work planning, initiative, adaptability, dependability, and relationships with co-workers. Superiors rated subordinates on a one-to-five scale on each performance dimension. Scores ranged from below standard to outstanding. An overall evaluation was also recorded for each individual on a five-point scale, ranging from unacceptable to exceptional. This overall evaluation was to be based on the eight individual performance categories and was used as input for making decisions on pay increases. Pay increases had two components: cost-of-living and merit. The cost-of-living component was tied to the rate of inflation. All non-faculty university employees received the same percentage cost-of-living adjustment. The merit portion of the salary increase was based on the overall performance rating. University administrators requested a pool of money for merit increases with final approval made by the state legislature.

Superiors were strongly encouraged to hold formal appraisal interviews between late January and early February for all subordinates. Superiors were instructed on how to use the university's rating form, and received training on how to conduct a successful appraisal interview. The superior's formal evaluation of an employee's performance was presented to the employee either prior to the interview or at the beginning of the interview session. As previously mentioned, superiors could elect to ask subordinates to independently prepare self-ratings prior to being informed of the superior's assessment of their performance. Subordinates evaluated themselves on the same rating form that superiors used to conduct the formal evaluations. The superiors did not receive any specific training on the use of self-ratings. The self-evaluation did not affect the formal performance rating, but could be used as a point of discussion during the appraisal interview.

Sample

All on-campus administrative and staff employees were surveyed. Faculty members were excluded from the study because they operated under a separate pay policy and per-
Self-ratings in an Appraisal System

In order to minimize the common method variance problem, two separate questionnaires were designed and administered—one before the formal evaluation occurred and another after appraisal interviews were conducted. Fourteen hundred forty-six questionnaires were distributed each time. The first questionnaire was distributed in late October, approximately two weeks before the performance evaluations were to be conducted. In response to the first mailing, a total of 789 questionnaires (55% response rate) were returned. The second wave of questionnaires was distributed during the third week of February. The response rate for the second questionnaire was 47 percent (679 completed questionnaires were returned). Responses to both surveys were obtained from 458 employees. Three criteria were necessary to be included in the study: (1) respondents had to be in their current job for a year, (2) their immediate supervisor must have conducted a performance evaluation, and (3) an appraisal interview must have been conducted. This reduced the final sample to 326 individuals. Respondents represented five general employment categories: clerical, professional, technical/paraprofessional, service, and skilled crafts. In this sample, 71% were female, their average tenure with the university was 7 years and their mean age was 39.0 years. Comparisons of respondents in terms of age, gender, tenure at the university, time in present position, and type of position indicated no difference between individuals completing one or both questionnaires. Furthermore, no indirect evidence of non-response bias was found for individuals not responding to either questionnaire. Information from the university’s human resource department confirmed that mean performance ratings for respondents was similar to the performance ratings for the overall staff, and the demographic characteristics of survey respondents were similar to the demographics of all university administrative and staff employees.

Procedures

All administrative and staff employees were invited to participate in this project by means of a letter from one of the investigators which accompanied the survey instrument. Information relevant to this study was collected in two phases. Two months prior to the administration of the university’s formal performance appraisal, employees were sent the first questionnaire through the university mail system. This questionnaire addressed issues concerning the appraisal process occurring prior to the actual performance rating. A second questionnaire was sent to employees approximately two weeks after the end of the period during which performance appraisal interviews were to be completed. Participants were assured that their individual responses would be confidential. Demographic information (birth date, gender and occupational classification) was used to match responses obtained from the two questionnaire administrations.

Measures

Items were answered by means of five-point Likert scales ranging from “strongly disagree” to “strongly agree.” Responses to individual items were aggregated to create scales.
Pre-appraisal Questionnaire

The first questionnaire was separated into three sections and included the information necessary to test the first two sets of hypotheses, with the exception of the self-rating measure and the performance rating assigned by one's superior. The first section asked employees to respond to questions related to their specific job, their superior and general work attitudes. The second section consisted of items regarding the procedural aspects of the appraisal process that occurred prior to the formal evaluation being made. The third section asked for demographic information.

Task Variety. This scale consisted of two items, adapted from Hackman and Oldham (1980), which had an alpha coefficient of .79 ("I do the same thing every day." (reverse scored); "There are always new projects and things to do on my job, so it is difficult for me to know what I will be doing one or two months from now."). High scores indicated greater task variety.

Task Autonomy. This scale was comprised of four items, based on Hackman and Oldham's Job Diagnostic Survey (e.g., "I tend to make a lot of decisions about my job duties on a day-to-day basis."); "I have considerable control over my daily work activities."). The alpha coefficient was .89.

Opportunity for Superior to Observe Work. This variable was measured with three items and had an alpha coefficient of .89. Two of the questions were "My supervisor seldom sees my work," and "I have a lot of contact with my supervisor in performing my job." The first question was reverse scored. This scale is similar to the single-item measure used by Judge and Ferris (1993).

Knowledge of Performance Standards. This scale, consisting of six questions, examined the extent to which the respondent was informed of performance standards ("At the start of the evaluation period, my supervisor told me the important parts of my job and what was expected of me during the coming year.") and the degree to which one was permitted to participate in setting performance criteria ("At the start of the evaluation period, I was given the chance to express my views and received a fair hearing from my supervisor with respect to the duties of my job and the standards by which I would be evaluated."). The alpha coefficient was .88.

Feedback Throughout the Evaluation Period. Nine items were averaged to assess the specificity, timing and frequency of feedback. These measures of performance feedback are consistent with those employed in earlier studies (Ilgen et al., 1981; Larson et al., 1986; Pavett, 1983). Three feedback issues were evaluated: interactions indicating good performance ("Regularly during the year my supervisor lets me know about the parts of my job I do well."); need to improve ("During the year, if my supervisor thought I could perform part of my job better, he/she would talk to me about it."); and feedback regarding changes in performance requirements ("As the situations and priorities in our department changed, my supervisor let me know so that I could keep my efforts focused on the important aspects of my job."). The alpha coefficient for the nine-item measure was .93.

Expected Performance Rating. Survey participants were asked to indicate
what overall performance rating they felt they deserved for the current year. The performance categories were the same as the overall formal evaluation they received from their superior.

**Post-appraisal Questionnaire**

The second questionnaire, which was sent out within two weeks after the appraisal interviews were conducted, focused on the nature of the formal appraisal interview, reactions to the appraisal interview as well as the overall evaluation process, and the performance ratings.

**Self-ratings.** A single-item measure was used to assess whether or not a self-rating was performed. Respondents were asked whether or not their supervisor asked them, prior to the appraisal interview, to provide a self-assessment of their job performance. This scale was labeled “yes,” “no,” or “don’t remember.” Since no respondents checked the “don’t remember” box the self-rating measure was treated as a dichotomous variable. Responses were coded as no = 0 and yes = 1.

**Appraisal Interview.** While the importance of the appraisal interview to the evaluation process is clearly recognized (Nathan et al., 1991), no single instrument exists which measures all aspects of it. In a review of the literature, Burke et al. (1978) reported that performance appraisal interviews can be examined along several dimensions: level of subordinate participation, attitude of the superior, mutual goal setting, existence of criticism, and degree of subordinate involvement. To tap these various dimensions, eight items were developed based upon previous research (Burke et al., 1978; Dipboye and de Pontbriand, 1981; Greller, 1975; Nemeroff and Wexley, 1979). Two scales were created—assessing discussions of past performance (judgmental or evaluative component) and future performance (developmental component)—during the appraisal interview. These factors parallel those reported by Dorfman et al. (1986) in their study of superior perceptions of the appraisal interview. The judgmental aspect of the appraisal review process, comprised of five questions, focused on the past performance of the ratee (e.g., “During the appraisal interview there was a free and open discussion of my job performance between my supervisor and me in which I expressed my views and he/she was open to consider what I had to say.”). Alpha coefficient for the judgmental dimension of the appraisal interview was .88. The developmental dimension, consisting of three questions, focused on how the ratee’s job performance could be improved in the future (“During the performance appraisal interview my supervisor asked me what I thought I could do to improve my job performance and then we discussed my ideas to improve my performance.”). The alpha coefficient for this dimension was .87. The two interview measures had a correlation of .54 (Table 1).

**Fairness of Ratings.** A two-item measure was developed to assess an individual’s perception of the fairness of the performance rating (e.g., “The performance ratings I received from my supervisor for the evaluation period which just ended were fair.”). Alpha coefficient for this measure was .84.

**Job Satisfaction.** This variable was measured using Quinn and Staines (1979) five-item measure. The alpha coefficient was .80.
Performance Rating. At the end of the evaluation period all employees with managerial responsibility evaluated the performance of each employee reporting directly to them. All employees received one evaluation. In order to maintain the anonymity of the employees we asked each respondent to report the rating he/she received from his/her superior. As noted earlier, the performance evaluation consisted of eight items measuring a specific aspect of a subordinate’s job, along with a single question that evaluated the subordinate’s overall performance. Only the question evaluating the subordinate’s overall performance was used to determine pay increases. Each subordinate’s overall work performance was rated on a five-point Likert scale (unacceptable to exceptional). This measure is similar to those used by Judge and Ferris (1993) and Nathan et al. (1991). Given the anonymity of the questionnaires, it is impossible to verify the accuracy of the information provided by each respondent. The Office of Human Resources stated that the performance ratings reported by survey respondents were in line with the actual performance ratings reported for all university employees. The correlation between the subordinate’s expected performance rating (questionnaire one) with actual rating received (questionnaire two) was .44. In a meta-analysis, Harris and Schaubroeck (1988) reported a correlation of .42 between self- and supervisor ratings of blue-collar workers. The correlations for our self-reported measures are consistent with the correlations of self- and supervisor ratings in previous studies.

RESULTS

A total of 122 individuals (36.5 percent) reported that their superiors asked them to provide a self-rating of their performance, while 214 respondents indicated they received no such opportunity. Means, standard deviations, and intercorrelations for all study variables are presented in Table 1. The highest intercorrelations appeared between variables measuring relationship with superior, knowledge of appraisal standards, and feedback during appraisal period. Work group size was not significantly correlated with any of the study variables. Thus, hypothesis 1b was not supported. In order to simplify the structural equation model, the measure of group size was dropped from further analyses.

The impact of several control variables was assessed. An analysis of variance was conducted to check for the potential difference in demographic characteristics for the self-rate and no self-rate groups. No significant difference was found between the two groups for the following variables: age, education level, years in position, years in the organization, expected performance rating and current performance rating for the subordinate, and age of direct superior. As an additional control, chi-square analyses were performed to check for potential gender effects. Women, in general, were slightly more likely to ask subordinates to self-rate ($\chi^2 = 4.3, p < .05$). Gender similarity between superior and subordinate had no effect on self-ratings. No self-rating effect was found for type of job that subordinates held. A confirmatory factor analysis, utilizing Lisrel® 8, was conducted to determine the independ-
<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>S.D.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Self-rating</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Months in current job</td>
<td>43.5</td>
<td>47.9</td>
<td>-05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Years in university</td>
<td>6.5</td>
<td>5.7</td>
<td>-07</td>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Work group size</td>
<td>10.7</td>
<td>18.9</td>
<td>07</td>
<td>-01</td>
<td>05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Autonomy</td>
<td>3.9</td>
<td>0.7</td>
<td>12</td>
<td>05</td>
<td>17</td>
<td>02</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Task variety</td>
<td>2.5</td>
<td>1.0</td>
<td>10</td>
<td>00</td>
<td>06</td>
<td>-06</td>
<td>49</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Observe performance</td>
<td>3.4</td>
<td>1.0</td>
<td>08</td>
<td>-19</td>
<td>-12</td>
<td>-07</td>
<td>06</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Knowledge of standards</td>
<td>3.3</td>
<td>0.9</td>
<td>21</td>
<td>-08</td>
<td>-10</td>
<td>03</td>
<td>20</td>
<td>18</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Feedback</td>
<td>3.4</td>
<td>0.9</td>
<td>17</td>
<td>-10</td>
<td>-12</td>
<td>04</td>
<td>27</td>
<td>27</td>
<td>55</td>
<td>68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Interview—Past performance</td>
<td>3.8</td>
<td>0.7</td>
<td>33</td>
<td>-11</td>
<td>-10</td>
<td>00</td>
<td>31</td>
<td>31</td>
<td>49</td>
<td>56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Interview—Future performance</td>
<td>2.9</td>
<td>1.1</td>
<td>41</td>
<td>-13</td>
<td>-10</td>
<td>00</td>
<td>23</td>
<td>26</td>
<td>28</td>
<td>40</td>
<td>44</td>
<td>59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Expected rating</td>
<td>4.3</td>
<td>0.6</td>
<td>-06</td>
<td>06</td>
<td>13</td>
<td>06</td>
<td>19</td>
<td>16</td>
<td>-01</td>
<td>05</td>
<td>05</td>
<td>06</td>
<td>00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Rating from supervisor</td>
<td>4.1</td>
<td>0.7</td>
<td>01</td>
<td>12</td>
<td>14</td>
<td>-02</td>
<td>17</td>
<td>22</td>
<td>06</td>
<td>22</td>
<td>20</td>
<td>33</td>
<td>17</td>
<td>44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Fairness of rating 3.9</td>
<td>1.0</td>
<td>13</td>
<td>-03</td>
<td>00</td>
<td>-03</td>
<td>21</td>
<td>20</td>
<td>23</td>
<td>34</td>
<td>42</td>
<td>63</td>
<td>43</td>
<td>11</td>
<td>52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Job satisfaction</td>
<td>3.6</td>
<td>1.1</td>
<td>12</td>
<td>-02</td>
<td>07</td>
<td>10</td>
<td>35</td>
<td>34</td>
<td>26</td>
<td>30</td>
<td>36</td>
<td>44</td>
<td>30</td>
<td>10</td>
<td>24</td>
<td>39</td>
</tr>
</tbody>
</table>

R > .14, significant at .01 level
ence of the variables used in this study. The initial measurement model indicated that the proposed exogenous and endogenous variables showed a moderately good fit of the model ($\chi^2 = 745$ [df = 229, $p < .001$]), comparative fit index [CFI] = .98, adjusted goodness-of-fit index [AGFI] = .97, standardized root-mean-square residual [SRMSR] = .12, root-mean-square error of approximation [RMSEA] = .07. Generally-accepted guidelines indicate that values of .90 and above on the AGFI are indicative of a good fit, while values above .95 on the CFI show strong practical significance. Values below .05 for the RMSEA show good fit. Inspection of the modification indices indicated the measurement model could be improved in the following ways. Autonomy and task variety were freed to load on the unobserved construct labeled “job characteristics,” and the judgmental and developmental components of the appraisal interview loaded on a single latent variable, labeled interview discussion. Combining judgmental and developmental components of the appraisal interview is consistent with recent research that development and evaluation are interdependent (Boswell and Boudreau, 2002). These changes significantly improved the fit of the measurement model ($\Delta \chi^2 = 475$, $p < .001$).

Lisrel® 8 was used to test the fit of the hypothesized model and the two alternative models. For the hypothesized model, we tested direct and mediated effects hypotheses that were specified as depicted in the structural equation model displayed in Figure 1, with the following changes mentioned earlier: group size was dropped from the model; task autonomy and task variety were combined to form one variable. In order to simplify the presentation, we did not show the indicators for each exogenous and endogenous variable. With the exception of the RMSEA, the chi-square statistic along with a variety of goodness-of-fit measures indicates that the hypothesized model did not fit the model very well ($\chi^2 = 618.94$, [df = 155, $p < .001$], GFI = .87, AGFI = .83, CFI = .89, NNFI = .87, RMSEA = .029).

While the hypothesized model did not fit the model well, other models might explain the data better. Testing alternative nested models as suggested by Hayduk (1987) will show whether added causal paths will result in a significant decrease in chi-square. We tested for the existence of a partially-mediated relationship between the antecedent variables and the appraisal interview, and between opportunity to self-rate and the outcome variables of job satisfaction and fairness of ratings. The new model is shown in Figure II and displays only statistically significant relationships. Opportunity to observe performance was not significantly related to opportunity to self-rate or the measure of the appraisal interview and was dropped from the final model. The overall fit statistics for the new model are ($\chi^2 = 334.81$, [df = 149, $p < .001$], GFI = .98, AGFI = .98, CFI = .98, NNFI = .97, RMSEA = .05). This model explained 47% of the variance in fairness of the ratings and 50% of the variance in job satisfaction and fits the data better than the hypothesized model. Adding paths directly from the antecedent variables to the appraisal discussion significantly improved the fit ($\Delta \chi^2 = 241.55$, df = 4, $p < .001$). Specifically, variables measuring job characteristics ($\beta =$
.17, \( p < .01 \), setting standards (\( \beta = .29, p < .001 \)) and feedback (\( \beta = .22, p < .05 \)) were significantly related to appraisal interview discussion. Self-rating partially mediates the relationship between feedback received during the year and the appraisal discussion.

We also added direct paths between opportunity to self-rate and the outcome variables, fairness of rating and job satisfaction, to test for the existence of partial mediation effects of the appraisal interview. Both of these paths were statistically significant and improved the model fit (\( \Delta \chi^2 = 42.58, \text{df} = 2, p < .001 \)). Opportunity to self-rate was negatively related to fairness of rating (\( \beta = -.25, p < .05 \)) and job satisfaction (\( \beta = -.33, p < .01 \)). The appraisal interview partially mediates the relationship between opportunity to self-rate and the outcome variables.

**Antecedents of Self-ratings**

Tests of the first hypothesis showed that feedback received during the year was positively related to opportunity to self-rate (\( \beta = .26, p < .05 \)), while knowledge of performance standards and the contextual job factors were not significantly related to opportunity to self-rate. However, opportunity to self-rate along with job characteristics, opportunity to set performance standards and feedback received throughout the year were significantly related to discussion of past and future performance in the appraisal interview.

**Appraisal Interview and Outcome Measures**

Given the results from the measurement model, we used our measures of discussion of past performance and future performance as indicators of an overall measure of interview discussion. As shown in Figure II, we assessed the impact of self-ratings on the opportunity to discuss performance issues in the appraisal interview. Opportunity to self-rate was strongly related to the opportunity to discuss performance issues (\( \beta = .49, p < .001 \)), supporting hypothesis 2. Hypotheses 3a and 3b stated that the opportunity to discuss performance issues in the appraisal interview would mediate the relationships between opportunity to self-rate and perceived fairness of performance ratings and job satisfaction. As depicted in Figure II, the appraisal interview only partially mediates the relationship between opportunity to self-rate and the outcome measures of perceived fairness of ratings and job satisfaction. Discussions of past and future performance in the appraisal interview are strongly related to perceived fairness of the ratings and job satisfaction. Opportunity to self-rate is negatively related to perceived fairness of ratings and job satisfaction, the opposite direction of the bivariate relationships shown in Table 1.

**DISCUSSION**

Understanding the factors that influence the opportunity to self-rate is clearly important. The findings of this study demonstrate that activities involved in appraising performance should not be viewed in isolation of each other. Performance evaluation is a process that evolves throughout the year, culminating with the formal appraisal interview, rather than a singular event (Blasingame et al. 1981). Much of the criticism surrounding
Figure II
Antecedents and Consequences of Self-Ratings

- Fairness of Rating
- Job Satisfaction
- Appraisal Interview

Opportunity to Self-Rate

- Job Characteristics
- Setting standards
- Feedback

Parameters are standardized parameter estimates. This is a simplified version of the model. Error terms, factor loadings and correlations between exogenous factors are not shown (n = 326).

* p < .05, ** p < .01
Self-ratings has emerged from studies that view self-ratings as an event, not a component of the process. While experimental studies are clearly suitable for investigating the existence of possible rating errors (e.g., leniency and halo), little is known about the actual use of self-ratings in ongoing organizational appraisal systems.

We found that supervisors were somewhat more likely to ask subordinates to self-rate when they had provided performance feedback throughout the year. As hypothesized, we might expect that managers who keep employees informed of their performance throughout the year will be more likely to involve them in the appraisal process. However, contrary to our initial hypotheses, self-appraisals were not used in this situation to help the supervisor gather more information about performance. Neither opportunity to observe performance or job characteristics were predictors of opportunity to self-evaluate. A possible explanation for this is that since employees self-appraised after the formal performance evaluations were conducted, managers did not use self-appraisals to gather more information on performance; rather, self-ratings were used primarily to provide a basis for discussion during the appraisal interviews.

The results of this study suggest that the relationships associated with the use of self-ratings should be interpreted with care. Previous studies focused on the impact of self-ratings with either the performance appraisal interview or a variety of outcome variables. We included measures of both the appraisal interview and outcome variables in this study. The importance of the appraisal interview to the evaluation process was demonstrated by the results of the structural equation model. The measure assessing the discussion of past and future performance, not the opportunity to self-rate, was the strongest determinant of the individual’s reactions to the performance appraisal process. Correlation analysis indicates that opportunity to self-rate is positively related to both opportunity to discuss performance in the appraisal interview and outcome variables of fairness of ratings and job satisfaction. However, the results of the structural equation model indicate the possible existence of bifurcated effects when controlling for the mediating effects of the appraisal interview on the outcome variables. Self-ratings led to greater opportunity to discuss performance issues in the appraisal interview, but a negative relationship emerged between self-ratings and the outcome variables. Post hoc analysis showed that when the effects of the appraisal interview are controlled, measuring partial correlation, opportunity to self-rate was negatively related with fairness of ratings and job satisfaction.

A possible explanation is that requesting a self-appraisal may raise an individual’s expectation that he or she will have more influence (Folger and Greenberg, 1985). If that expectation is not met, he or she may experience frustration. The negative relationship between self-ratings and the outcome measures of perceived fairness of ratings and job satisfaction indicates the existence of an instrumental effect of voice (opportunity to self-appraise). Korsgaard and Roberson (1995) propose that instrumental voice effects exist when individuals perceive that voice (opportunity to self-evaluate) gives them an opportu-
nity to indirectly impact the decision. In this situation, the formal evaluations were completed and self-ratings had no impact on the formal ratings. It appears that expectations of influence were raised by the opportunity to self-rate. While individuals who self-rated did have the opportunity to provide information about their performance, this discussion had no impact of the formal rating. On a practical note, Waldman et al. (1998) reported that some organizations have stopped using 360-degree feedback because it created negative attitudes and resulted in inflated ratings. The opportunity to self-rate may have bifurcated effects. Self-ratings may be followed by greater discussion during the appraisal interview, resulting in general greater perceived fairness of ratings and increased job satisfaction, but also may increase expectations of some employees.

Our findings indicate that self-ratings have an overall positive impact on the performance appraisal process. Clearly, self-ratings serve as a catalyst for discussion during the appraisal interview. From a procedural justice perspective, in this study self-ratings satisfy the importance of representation as identified by researchers (Leventhal et al., 1980; Greenberg, 1986; Folger and Konovsky, 1989). Since self-ratings were collected after the formal ratings were conducted and recorded, there were no mechanisms in place to insure accuracy or to correct a potential mistake. In line with procedural justice research, our findings indicate that when self-ratings are used opportunities to correct or adjust the performance ratings should be available. This may become even more important in situations, like the current study, where performance-based pay systems exist.

Limitations and Caveats

Several significant limitations of this investigation exist. Our desire to maintain the anonymity of respondents restricted our ability to match subordinates with their direct superiors. Thus, all information about the appraisal process comes from the subordinate. Questions of response bias are clearly raised in this situation. We attempted to address this problem by gathering information on two separate occasions, prior to performance assessments and after the appraisal interviews were completed. Questions concerning the superiors' use of self-ratings in the past, and whether or not superiors ask all subordinates in their work group to self-evaluate, remain unanswered. This study marks the first attempt to understand the conditions under which superiors voluntarily encourage subordinates to self-evaluate. This work focused primarily on the subordinate's perceptions of the supervisor's behavior. Other potentially important variables, such as managerial trust, were not assessed.

An additional note of caution should be made. For the purpose of control, in the pre-appraisal questionnaire we asked all respondents to indicate what performance rating they felt they deserved. Thus, all individuals in the study reported an overall performance score. While we found significant effects for the supervisor-initiated self-appraisal, care should be taken in interpreting the findings.

Future Research

The criticisms of self-ratings have been quite strong, resulting in gen-
eral disenchantment. Interestingly, this viewpoint has arisen despite scant empirical research of self-ratings in actual performance appraisal systems. Our study suggests that self-ratings make a significant contribution to subordinates’ perceptions of appraisal interview effectiveness.

This study shows that in a situation where self-rating occurs at the discretion of the superior, the prior behavior of the superior appears to be important in determining whether a subordinate will be given the opportunity to complete a self-rating. In order to gain more insight into this issue, information also should be obtained from superiors to check for consistency of perceptions about the nature of the dyadic relationships. We attempted to limit problems associated with method variance by administering two questionnaires.

Our findings suggest that the use of self-ratings may be affected by the extent to which a superior prefers a participative style that involves subordinates in the process and keeps them informed. An alternate explanation is that a superior may be selective in determining which subordinates are asked to complete self-ratings. Future research should examine voluntary self-ratings at both the individual and the work-group level to determine if managerial style is an important predictor, or is the interpersonal relationship that exists between a superior and subordinate more important. The use of the LMX Model (Dansereau et al., 1975; Graen and Cashman, 1975), focusing on in-groups and out-groups, would provide a strong conceptual starting point. While this study focused on voluntary ratings, it would be interesting to examine the conditions under which superiors avoid the use of self-ratings where they are mandated.

In a similar vein, Fried et al. (1992), employing the expectancy theory framework, studied the conditions under which superiors refuse to conduct appraisal interviews. Job experience, time under superior, trust and managerial style were significant predictors. This model could be extended to self-ratings to determine if the same reasons apply.

References


