Comparing the Performance Appraisal Practices in Large Firms with the Directions in Research Literature: Learning More and More about Less and Less

Robert D. Bretz, Jr.
George T. Milkovich
Walter Read

Working Paper 89-17

CORNELL ilr School of Industrial and Labor Relations

COMPARING THE PERFORMANCE APPRAISAL PRACTICES IN LARGE FIRMS WITH THE DIRECTIONS IN RESEARCH LITERATURE: LEARNING MORE AND MORE ABOUT LESS AND LESS

Robert D. Bretz, Jr.
George T. Milkovich
Center for Advanced Human Resource Studies

Walter Read
Director of Compensation, IBM Corporation
and
Center for Advanced Human Resource Studies

Working Paper #89-17

Revised: September 1990

COMPARING THE PERFORMANCE APPRAISAL PRACTICES IN LARGE FIRMS WITH THE DIRECTIONS IN RESEARCH LITERATURE: LEARNING MORE AND MORE ABOUT LESS AND LESS

ABSTRACT

Managers responsible for the performance appraisal practices of the Fortune 100 were surveyed to describe their firm's current practices and to solicit their views regarding important issues related to these practices. First the results of the survey are reported. These results are compared to the current research directions and issues discussed in human resource journals over the past five years. Survey results reveal that performance appraisal systems in large firms were designed primarily by human resource specialists, with limited input from either managers or employees. Objective based plans are far and away the most widely used. Little concern over psychometric properties of scales is evident. Managers spend about six hours per employee per year appraising performance, but are not evaluated on how well they conduct them. According to the managers responsible for performance appraisal, employee acceptability and employee sense of fair treatment are the most important issues they face. Over 90 percent of the firms use performance appraisals in their merit pay decisions. The survey findings contrast sharply with the directions of performance appraisal research. The recently published research is dominated by cognitive process and psychometric issues; the vast majority of which are conducted in laboratory settings, using student subjects and paper people or video scenarios. The differences between current practices and concerns of managers in large organizations, and research directions are discussed. Suggestions are offered for future research.

COMPARING THE PERFORMANCE APPRAISAL PRACTICES IN LARGE FIRMS WITH THE DIRECTIONS IN RESEARCH LITERATURE: LEARNING MORE AND MORE ABOUT LESS AND LESS

The gap between performance appraisal as practiced by managers and the issues examined in research journals is a recurring theme in the performance appraisal literature. According to some writers, research has done little to improve performance appraisal's usefulness as a tool to help managers improve their decision making (Thorndike, 1949; Banks & Murphy, 1985; Napier & Latham, 1986). The issues that dominate performance appraisal research (i.e. formats, evaluator training and cognitive processing), and the methodological designs being used in this research seem at odds with organizational realities. For example, Banks & Murphy (1985) warned that if cognitive process research continued along contemporary lines, the gap between performance appraisal research and practice would increase. Napier & Latham (1986) suggested that progress on performance appraisal practice has lagged because the research which might inform practice has ignored Thorndike's (1949) call for practicality in its quest for measurement elegance. Bernardin & Villanova (1986, p. 58) state that "a more systematic and detailed delineation of the modal setting is in order" if we are to improve the degree to which performance appraisal research contributes to performance appraisal practice. They further suggest that surveys providing detailed descriptions of the state of current practice are essential for the "development of more realistic research settings".

There is a growing concern that much organizational research, while methodologically strong, lacks substantive application and is directed toward increasingly selective audiences of researchers to the neglect of other audiences, such as policy makers

and managers. (Bedeian 1989, Whyte 1989). We undertook this study to examine the specific performance appraisal practices of large U.S. industrial organizations and to critically analyze the state of performance appraisal research in light of these practices. While we recognize that guiding the practice of managers is not the sole purpose of scholarly research, it is one of the objectives, particularly since performance appraisal is such an applied research topic.

Performance Appraisal Practices

Sample

Sampling the current state of performance appraisal practice in U.S. industry presents some interesting issues. U.S. industry is highly diverse, varying by technology, size, performance, strategy, management style, culture -- a staggering set of potential factors exist on which to design a study. We decided to focus on the practices of the largest private sector U.S. firms; those in the Fortune 100. Arguably, formalized performance appraisal is most evident in large organizations. Formalized systems, attention to process, formats, trained raters and multiple uses would, we believe, be more likely to receive attention in large organizations. In one sense the Fortune 100 represent the "most successful" firms in terms of sales volume, profits, assets, and equity. In another sense, they may represent the most bureaucratically burdened, least innovative firms. Perhaps more importantly, inclusion in the Fortune 100 signals other firms and managers in those firms that "these companies may be doing something right". Consequently, describing the performance appraisal practices among this group of firms seemed a reasonable place to begin to understand the current state of practice.

Questionnaire

The initial questionnaire was developed around issues identified in the performance appraisal literature. Subsequently, managers from IBM, Corning, Digital, Pfizer, and Xerox served as a review panel to help identify the issues of concern to both managers and researchers. For example, many organizations have different appraisal policies and use different appraisal methods for various employee groups. Input from the panel resulted in redesigning questions to allow for different responses to the questions in regard to the

appraisal of executives, managers, professional, non-exempt, and other hourly employees.

The questionnaire was organized into seven basic dimensions; (1) system design, (2) system characteristics, (3) system management, (4) procedural and distributive justice, (5) uses and current issues, (6) appraisal and pay increases and (7) performance distributions. The final instrument was 20 pages long, soliciting 505 coded responses, plus open-ended questions and requests for documentation. It was designed to be completed by individuals in policy-making positions. However, due to the detail required by some of the questions, respondents were encouraged and expected to seek technical assistance where needed. The most senior person in each Fortune 100 organization listed in the American Compensation Association Directory received the questionnaire. Instructions urged these contacts to forward the questionnaire to the manager responsible for the appraisal system. About one week after the mailing, phone calls were made to insure that the questionnaire had been received, to answer questions about the study, and to encourage participation by the most appropriate person in each organization. In most cases, several people in the organization contributed to completing the questionnaire.

Results

We confirmed that 92 of the Fortune 100 companies received a questionnaire. Sixty-five organizations (70%) responded with two firms reporting that they do not do formal performance appraisals. Thus, our description of employer performance appraisal practices is based on the sample of sixty-three firms that indicated they engaged in formal individual appraisals. These organizations employed an average of 20,816 exempt and 31,407 non-exempt employees. The non-respondents tended to be more decentralized than those that responded. Consequently our results do not contain the results of some highly

decentralized firms. However in several of these firms each of the subunits completed separate questionnaires. We did not include the responses from these subunits in the analyses but subunit results do not differ in any meaningful way from those reported by other firms. No other differences were noted between respondents and non-respondents.

The results are reported on the aforementioned seven basic dimensions developed with the aid of the research panel. We then discuss the current performance appraisal research and compare the survey results with directions evident in the research. We conclude with observations about the limitations of this study and implications for future research.

System Design

The performance appraisal systems used by large U.S. Firms are designed primarily by personnel specialists with only limited input from the managers who use the system and even less from employees whose performance is rated. As Figure 1 indicates, personnel specialists were highly involved (all $m > 4.1$ where 1 = no involvement and 5 = great involvement), for all job groups, compared to line manager involvement (all $m < 3.2$) and employee involvement (all $m < 2.6$) (All $F > 39.4$, $p < .001$).

Considering the recent interest in participation and involvement, we anticipated that more recently implemented systems might rely on more input from line managers and employees. The average age of the systems was 9.5 years old; however, firms with systems five years old or less were no more likely to have involved managers or employees in their design than were firms with systems over five years old (all $T < 1.5$, $p > .13$, n.s.). No industry differences were noted in involvement of employee groups or in whom respondents thought should be involved in the system design. A 1977 report by the

Conference Board on managerial performance appraisal practices stated that 50% of the organizations they surveyed had performance appraisal systems that were less than three years old. It appears that within the Fortune 100, many appraisal systems have not significantly changed since that time. Caution is required here. During our debriefing sessions, many managers told us that the "basic system had not changed but modifications were common." Among these were changes such as the number of levels in the scale, and the detail required in specifying objectives.

Insert Figure 1 Here

System Characteristics

Formats. For exempt employees, the objective based approach is the preferred form (see Figure 2). Seventy-eight percent reported using it for executives, 83% for managers, and 70% for professionals. In a less detailed survey, (Wyatt 1989) also identified objective based appraisals as the most common format. Graphic rating scales were reported as a distant second (10% use them for executives, 20% for managers, and 23% for professionals).

Objective based schemes are used to a lesser extent for non-exempt employees (31% of the firms). An equal percentage use graphic rating scales (31%), fewer use behaviorally anchored rating scales (14%), 7% use forced choice, and the remainder rely on various other techniques or do not conduct formal appraisals for their non-exempt employees. Fifty-two percent reported that formal appraisals are not conducted for hourly employees.

Another 23% reported that appraisals are conducted for hourly employees but that no particular format is used.

Forced distributions and ranking systems are often used in conjunction with other appraisal methods. This is true especially for managerial, professional, and non-exempt employees. Forced distributions are used for executives and hourly employees, by only 10% of organizations, but by 22% for managers, 27% for professionals, and 20% for non-exempt employees. Additionally, ranking procedures are used for executives by 12% of organizations, and by 26% for managers, 28% for professionals, 18% for non-exempt, and 4% for hourly employees.

Insert Figure 2 Here

Raters and Sources of Rating Information. As Figure 3 indicates, the immediate manager is the key evaluator of employee performance (e.g., executives, 42%; managers, 47%; professionals, 67%; nonexempt, 74%; and other hourly, 56%). For managerial and professional employees, the second level manager also has significant input (28% and 18% respectively) but plays a minor role in evaluating executives, nonexempt and hourly employees. In these instances second level managers supply less than 10% of the overall rating input. Higher level managers (third level or above) supply significant input in ratings of executives (26%) and managers (15%) but not for other employees.

Contrary to the popular press (e.g. Keichell, 1989, p. 201) no evidence was found that peer and subordinate ratings are "mushrooming". While some organizations do report conducting subordinate, peer, and self appraisals, less than 3% of the rating information

used in the final appraisal decision comes from these sources. Wyatt (1989) reports that only four percent of their sample use peer review but does not indicate the weight that peer ratings carry in the final appraisal.

Insert Figure 3 Here

Quantitative indices are used to supply some performance information in most organizations. For executives and managers, profits, sales and costs were cited as important measures; for professionals, the acquisition and use of job-specific knowledge, was important and attendance, and quality and quantity of work performed were important for nonexempt and other hourly employees. These measures were fairly constant across all industry classifications.

System Management

Time Spent. As Figure 4 shows, significantly different amounts of time are spent on the appraisal process for various employee groups ($F = 4.58$, $p. = .001$). An average of 8 hours per employee per year is spent in appraisals of executives and managers, 6 hours for professionals, and 3.6 hours for each non-exempt employee. However, these averages are inflated by a handful of firms that report spending between 20 and 40 hours per employee per year on the appraisal process.

Insert Figure 4 Here

Decision Making. Performance appraisal policy decisions (e.g. whether to conduct formal appraisals, whether to link pay to performance, etc.) are made at the corporate level in the majority of firms (68%) but a significant proportion (32%) make them at the business unit level. An exception appears to be computing and electronics firms in which policy decisions are about as likely to be made at the business unit level (43%) as they are at the corporate level (57%). Decisions regarding appraisal practices (e.g. type of format to use, rater training issues, etc.), however, tend to be made at the business unit level (52%) rather than corporate level (43%). Very few organizations allow decisions about performance appraisal policies or practice to be made at the facility level (5%). Recall, however, that one of the reasons for non-response to the survey is decentralization of all performance appraisal decisions. Thus, our sample is biased toward organizations with centralized design and policy. Considering the sample size of 65 firms, a few responses toward corporate or unit level would affect these percentages.

Training. As Figure 5 indicates, most organizations report extensive use of performance appraisal training programs but there are significant differences in the types of training used ($F = 6.24$, $p. < .001$). Training programs are likely to include how to conduct appraisal interviews and provide feedback (90% of the organizations report doing this), how to use the forms (83% do this), setting performance standards (78% do this), how to recognize good performance (66% do this) and how to avoid making rating errors (56% do this). Employees receive virtually no training in how to best use the process to receive feedback or improve performance. Training is focused on the manager; preparing employees for their role in the appraisal process simply does not occur among Fortune 100 firms.

Insert Figure 5 Here

Accountability. It is uncommon for managers to be evaluated on how they manage the appraisal process (see Figure 6). Basic motivation models tells us that people will tend to behave in ways that maximize their expected payoffs or in ways for which they are reinforced (e.g. Vroom, 1964). Given this, if managers are expected to conduct appraisals in a thoughtful and thorough fashion, whether they do so could easily be a performance dimension in their own appraisals. Only 22% of these Fortune 100 firms report rating managers on how well they conduct the performance appraisals for which they are responsible.

Insert Figure 6 Here

Expectations and Standards. Consistent with the objective based approaches, about 70% of our firms reported that employees participate in setting their own performance standards. However, the involvement varied between employee groups ($F = 37.37$, $p. < .001$). Executives, managerial and professional employees are involved in setting their performance standards in over 80% of the firms, non-exempt employees in about 55%, and hourly employees in 16%. In organizations that do not use objective-based systems, it is most common for employees to learn of standards and expectations through discussions with their managers or through written information such as job descriptions supplied by their managers.

Procedural and Distributive Justice

Most organizations report being concerned with procedural justice issues. Sixty-four percent report having an informal dispute resolution system (e.g. open door policies) that employees may use to contest the appraisal outcome. An additional 26% report a formalized process available for this purpose (e.g. binding decisions made by a third party), and 10% report that no appeals process is available (see Figure 7). Computing and electronics firms are unusual since 43% report a formal appeals process and an additional 50% report an informal mechanism. Appeals seem less likely to be entertained in the chemical, pharmaceutical, and petroleum firms; 28% report that they do not have even an informal appeals process for appraisal disputes.

Insert Figure 7 Here

While it may be common to have a mechanism to handle appeals, as Figure 8 indicates, it is far less common to solicit employee opinions about the appraisal process. Only one-third of the organizations conduct attitude surveys to determine either the managers' or the employees' perceptions of fairness of the appraisal process or the results obtained. Similarly, only about 40% of Fortune 100 firms survey employees' satisfaction with the appraisal system.

Insert Figure 8 Here

Important Performance Appraisal Issues and Uses in Practice

Consistent with a concern for procedural justice, respondents identified fairness issues as the most important they faced (see Table 1). The top two issues raised were (1) a performance appraisal system that is accepted by those being rated, and (2) employee sense of being treated fairly by the appraisal process. The next three most important issues focused on the results obtained; (3) whether employees believe that the results are fair, (4) the type of feedback given, and (5) the usefulness of performance appraisal as a tool to help manage performance. Even when prompted by items in the questionnaire, cognitive processes and psychometric issues such as halo and leniency were not considered important by decision makers in these firms.

Insert Table 1 Here

Participants were also asked to indicate what use they actually made of performance appraisal information (see Table 2). The average rank order includes (1) advising employees of work expectations, (2) improving work performance, (3) administering pay on the basis of merit, (4) determining employee potential, (5) counseling and developing employees, and (6) making promotion decisions. The least frequent uses are equally informative. At the bottom of the list were (1) assisting in long-range employment planning, (2) validating selection procedures, (3) making work assignments, and (4) making decisions about layoffs, terminations, or transfers. The last one, aid making layoff decisions, is interesting because one of the alleged uses of performance appraisal is to

improve workforce quality by identifying the top and bottom performers. Debriefings with managers indicated that performance rankings, as contrasted with ratings, are more useful for making layoff decisions. Ratings, according to managers, are not sufficiently "fine tuned" compared to rankings. These results were relatively consistent across industry groups.

Insert Table 2 Here

Performance Appraisal and Pay Delivery

Ninety-two percent of our firms use appraisal results in determining merit pay increases. Again, this practice is widespread across all industries and consistent with other survey results which report over 90% of all companies using performance appraisal information to determine pay increases (Hewitt, 1989; Laud, 1983). Most organizations believe that the performance appraisal systems accomplish merit pay objectives fairly well (2.5 on a five-point scale where 1 = very well, 3 = somewhat, and 5 = not at all), and a majority (62%) claim that the performance pay system provides sufficient differentials between high and low performers (see Figure 9).

Insert Figure 9 Here

Ranking appears to be used to supplement other performance appraisal methods for the sole purpose of allocating limited resources in merit increase decisions. About one-

third of the firms rank managerial and professional employees for the purpose of distributing pay increases.

Performance Distributions

The results of greatest interest to managers during our debriefings, were the number of levels used in other firms' appraisals and their actual distributions of ratings. Our results showed that 5 performance levels is most common but about 20% use more and 20% use less.

Figure 10 shows the average distribution of employees by group, in each of these levels. Level one represents the highest (best) rating available. The label "Far Exceeds Expectations" is representative of the type of label many organizations assign to this level of performance. Similarly, levels two through four represent the second, third, and fourth levels of performance. Level four also contains all employees from lower levels in systems where level five was not considered the worst possible rating. For example, in a system that uses eight performance levels, level four would consist of all employees in levels four through seven. Level five, labeled "Unsatisfactory" represents the percentage of employees receiving the lowest performance rating report by each organization.

Insert Figure 10 here

As the results indicate, even though most firms report systems with five or more levels, generally only three levels are used. Both the expected (desired) and the observed distributions are clearly top heavy. Also, the observed distribution of performance is

generally higher than the distribution desired. Note that even though large percentages of employees are expected in the top levels, for all employee groups the observed entries in the top two levels exceed the expected entries, while the expected entries in the bottom levels generally exceed the observed but only in a few cases are these differences statistically significant. The distribution data for hourly employees was sporadic; few firms appraise hourly employees and fewer yet reported these data.

Figure 11 shows the percentage of employees from each group that were rated in the top two levels. For example, 69% of executives were rated as "Exceeds Objectives" or "Far Exceeds Objectives." As anecdotal evidence and researchers' concern about leniency have suggested, it appears that the norm in U.S. industry is to rate employees at the top end of the scale. Skewed performance distributions clearly exist and appear to be common. Notice that even among organizations that rank and use forced distributions, the proportions of high-end ratings are similar to the distributions in those that do not. This suggests that the perceptions of unfairness and dissatisfaction believed to be associated with forced distributions may not be justified since the distributions observed do not differ from other distributions obtained by processes that are judged to be more fair. Also note that even among those organizations that do not believe they have a skewing problem (23% of the respondents), the percentages are generally greater than the remaining 77% who report they have a problem. Clearly, some managers do not see the higher rating distribution as a problem.

Insert Figure 11 here

Recent Research Issues

Having explored in detail the performance appraisal practices of the Fortune industrial 100, we now turn our attention toward the current state of performance appraisal research. What occupies the time and resources of researchers who study performance appraisal? What performance appraisal questions are being asked and reported in our empirical and theoretical journals? What performance appraisal issues are examined in the professional and practitioner oriented human resource journals? Caution is required here. We recognize that popularity, as judged by frequency of occurrence in research journals, does not necessarily imply importance. Some issues may be easier to research than others, and some data are certainly more accessible than others. However, we believe (or hope) that researchers and journal editors are not likely to expend scarce resources on trivial issues.

A computerized search (ABI/Inform) augmented by reviewing the tables of contents from several academic and professional journals was conducted. While recognizing that our literature review may not be exhaustive, we believe that it is extensive enough to indicate, with some precision, the direction performance appraisal research has taken over the past five years. Eighty research articles about performance appraisal published between 1985 and mid-1989 were identified. An additional 73 articles from the professional literature were considered. Only the last five years of research is reviewed because reviews covering earlier periods are available (e.g. Bernardin & Beatty, 1984; Bernardin & Villanova, 1986; DeNisi, Cafferty & Meglino, 1984; DeNisi & Williams, 1988; Feldman, 1981; Landy & Farr, 1980; Wexley & Klimoski, 1984). Readers should also note that

some of the cognitive processing studies identified below are discussed in some detail in a broad-based review of the cognitive processing literature by Lord & Maher (1989). Our review is intended to show the overall directions and methodologies of current performance appraisal research. Therefore, the partial overlap with Lord & Maher's review is necessary in order to provide the reader with a complete picture of the directions that performance appraisal research has taken.

Overview of Recent Literature

Table 3 serves as a guide to performance appraisal research published during 1985 through mid 1989. Published articles are categorized by primary issue explored and methodology used. Where more than one issue was examined, we tried to determine the primary focus and categorize the article only once to avoid redundancy.

The research published in the past five years is heavily weighted toward cognitive process issues. Twenty-four research-based articles directly addressed information processing themes. The vast majority of these (N=17) were conducted in laboratory settings using student subjects (N=16) and either paper people (N=7) or video-tape (N=10) formats. Rater/ratee characteristics have also received considerable attention (15 published studies identified) but research has not been concentrated on any particular characteristic. Study of psychometric issues remains common (N=12) with more attention on halo (N=7) than on other issues. Feedback issues are the one area where field studies appear to be the rule rather than the exception. Research about sources of ratings, rater training, and format issues is limited, however, these issues continue to be discussed in the practitioner oriented sources. Other issues including fairness, attitudinal and behavioral consequences of

appraisals, and appraisal uses are barely being explored by research but represent major themes in the practitioner literature and among managers in our survey. The literature tends to treat these issues in case-study descriptions or in "how to" articles. In the following sections, each of these major segments of the literature is examined in detail.

Cognitive Processing of Information

Information processing concerns dominate the studies of performance appraisal appearing in the research journals. Laboratory settings and student subjects are very common; only five studies appear to have followed Banks and Murphy's (1985) call for incorporating non-student subjects and field settings in research designs of cognitive process issues (Schmitt, Noe, & Gottschalk, 1986; Hogan, 1987; Mount & Thompson, 1987; Jolly, Reynolds & Slocum, 1988; Huber, Podsakoff & Todor, 1986). Cognitive processing studies conducted over the last five years appear to be concentrated around two issues: (1) how prior expectations or knowledge of prior performance levels affect the appraisal process, and (2) the role of memory in the recall process.

Hogan (1987) examined the effects of prior expectations on performance ratings using 49 banking supervisor-subordinate diads and a longitudinal design. She reported that supervisors' expectations introduced error into the rating process, and that disconfirmation of prior expectations appears to lower ratings. The effect of prior expectations was also examined in a field setting of subordinates' ratings of managers whose behaviors were considered either congruent or incongruent with prior expectations (Mount & Thompson, 1987). They suggest that when behavior is congruent with expectations, appraisal results tend to be more accurate but also contain more leniency and halo.

Several studies of the effects of prior knowledge have also been conducted in laboratory settings. Huber, Neale & Northcraft (1987) used paper people and a managerial sample to study the effects of ratee and rater characteristics and performance standards on performance-related judgments. They found that past rating history tends to anchor the current rating and that rater characteristics moderate the relationship between rated and objective performance. Using student subjects and video-taped presentations of performance, research has shown that knowledge of prior performance caused contrast effects (i.e., bias away from level of prior performance) rather than assimilation effects (Murphy, Balzer, Lockhart & Eiseman, 1985; Smither, Reilly, & Buda, 1988). Additionally, Steiner & Rain (1989) reported that the order in which good and poor performance was observed had an effect on performance ratings and that raters may bias judgment about inconsistent extreme performance (unusually good or poor) toward the general impression already held. Finally, Schmidt, Hunter & Outerbridge (1986) used path analysis on data from four independent samples and concluded that job knowledge has much greater impact on supervisory ratings than did job sample performance.

There are significant implications here, given the way performance appraisal is currently practiced among Fortune 100 firms. Since the vast majority of the performance appraisal systems in use are objective-based, there will virtually always be some prior expectations of performance. The mechanism through which a manager and subordinate arrive at mutually agreed upon goals (i.e. the definition of objective-based systems) requires an assumption from each individual regarding the expected performance level. Furthermore, since performance is measured against established goals, prior knowledge of job performance can also be expected. Therefore, prior expectations and prior knowledge

not only exist, they are salient key features in the appraisal process used in most Fortune 100 organizations. Research has neither considered the impact of issue salience, nor have research designs clearly articulated the importance of this issue in objective-based environments.

Job knowledge and ratee knowledge influence how information is processed. Schmitt, Noe & Gottschalk (1986) used 153 school administrators to test the degree to which raters use similar methods of combining information and whether rater agreement is based on job-relevant inputs or on shared bias. They reported that overall ratings from different sources varied because different rater groups attached higher relative weights to the job-related performance dimensions that were most salient to them. Laboratory research using student subjects and paper people also suggests that job and ratee knowledge have significant effects on conceptual similarity and rating covariance, and on halo (Kozlowski, Kirsch & Chao, 1986).

The role of memory has also been important in cognitive processing research. All of the studies we identified use student subjects and laboratory settings. Under laboratory conditions, ratings made one day after performance was observed were already affected by memory decay (Murphy & Balzer, 1986). Conversely, Smither and Reilly (1987) reported that rater intelligence, not rating delays affect rating accuracy. Some evidence (Kozlowski & Kirsch, 1987) suggested that memory decay may cause job and ratee knowledge to decrease and result in halo error and inaccuracy. In examining whether memory system characteristics or rating task characteristics affect how performance data is recorded, Balzer (1986) suggests that raters may be more likely to record behavior that is inconsistent with first impressions -- a contrast effect. When the rater's memory demands are great, bias in

favor of general impressions or recent performance may be expected (Murphy, Gannett, Herr & Chen, 1986a). Stress has been shown to impact memory by (1) causing less differentiation across dimensions (halo), (2) affecting information processing at the retrieval stage, and (3) possibly doing so at the categorization stage as well (Srinivas & Motowidlo, 1987). On a related topic, DeNisi, Robbins & Cafferty (1989) reported that even the way performance diaries are organized influences how information is processed.

The research on memory characteristics may be useful but is unlikely to alter performance appraisal practice given the conditions under which it occurs among the Fortune 100. The usefulness becomes apparent with the realization that managers spend less than one day per year addressing the performance concerns of each employee. This time includes keeping records, completing forms, preparing for the interview and conducting the feedback interview. This is a relatively small amount of time spread over a long period, particularly when the performance of several employees must be recalled and reported. Understanding how the memory accommodates these demands is useful. However, we now know that among Fortune 100 firms, it is uncommon to hold managers accountable for how they manage the appraisal process. Therefore, assuming rationality, no amount of understanding how to improve recall will result in behavioral changes. Behavioral changes are most likely to result from changes in the reward system that make it beneficial for managers to attend to the process.

The remainder of the cognitive processing studies are very diverse and difficult to classify. Information collected from 22 nursing supervisors was used to construct a cognitive map of their appraisal processes (Jolly, Reynolds & Slocum, 1988). They found that values accounted for significant variation in performance ratings. In laboratory settings

using student subjects, personality theory (traits) has been shown to influence even behaviorally-based ratings (Krzystofiak, Cardy & Newman, 1988), and information acquisition patterns (ranking versus rating) have been shown to influence the processing of information (Williams, DeNisi, Magleno & Cafferty, 1986). Sex-role stereotypes were found to not influence causal attributions of performance and therefore had only small effects on performance ratings (Kinicki & Griffeth, 1985). Williams, DeNisi, Blencoe & Cafferty (1985) report that appraisal purpose and outcome had limited effects on how raters utilize information, and Feldman, Camburn, & Gatti (1986) suggest that illusory correlation is not likely to cause bias in ratings. Nathan & Alexander (1985) suggest a model for inferential accuracy based on the degree of congruence between the rater's implicit theory of performance and the actual occurrence of behavior, and the rater's willingness to make judgments with limited information. Finally, Dipboye (1985) suggests that overemphasis on cognitive determinants of performance ratings has led to neglect of behavioral, social, and affective determinants of bias in the rating process.

Several issues emerge from this brief exploration of the cognitive processing literature. First, there is a heavy reliance on student subjects and laboratory settings. While there is some evidence that laboratory settings may provide equally valid and generalizable results as those obtained in field settings (e.g. Locke, 1986), there is also convincing meta-analytic evidence that in the performance appraisal arena, effect sizes in paper-people studies are significantly larger than in studies involving observation of behavior (Murphy, Herr, Lockhart & Maguire, 1986b). Particularly since performance appraisal is an applied subject, the potential effects of situational and contextual variables must be considered. Sterile environments that dilute the richness and complexity of the

environment necessarily change the phenomenon of interest. The task of rating the performance of someone with whom an on-going relationship exists is both conceptually and operationally different than the rating task presented in laboratory settings. It seems almost paradoxical that the richness and complexity of the rating context makes the performance appraisal process uniquely suited to the application and testing of cognitive processing theory since the research designs implemented tend to neutralize the very factors that make the setting appropriate.

Rater/Ratee Personal Characteristics

As Table 3 reveals, the effects of rater/ratee personal characteristics also received considerable attention in the primary research journals. Fifteen articles addressed individual characteristic issues. Four of these examined sex/gender effects but conflicting results were reported. No sex (or race) effects were reported in field settings where job analysis was used to develop a task-based performance appraisal instrument (Thompson & Thompson, 1985). Likewise, no gender differences were reported when rating familiar tasks in work situations where feedback was available (Shore & Thornton, 1986). Conversely, students tended to give women professors higher ratings (Dobbins, Cardy & Truxillo, 1988). In a different, laboratory setting, using student subjects and paper people, Dobbins, et al. (1988) found that when ratings were made for administrative (versus developmental) purposes, raters with traditional stereotypes of women tended to be less accurate. Finally, Benedict & Levine (1988) used student subjects and a laboratory setting to show that females were more lenient with poor performers and tended to delay performance appraisals and feedback sessions more than males did.

The most notable feature of these articles is the tendency to show significant effects in studies utilizing student samples but the absence of significant effects in field settings. This raises the question of whether the conditions encountered in laboratory or experimental settings sufficiently capture the complexity of cross-gender relationships and sex-role stereotypes that exist in work settings.

Ratee age received limited attention. A field study of nursing supervisors reported that younger subordinates were rated higher than older subordinates performing the same job, and that supervisors' causal attributions appeared to be related to subordinate's age (Ferris, Yates, Gilmore & Rowland, 1985). Also in a field setting, Lawrence (1988) found that deviation from age norms was associated with performance ratings. Managers that were ahead of their age cohort received higher ratings while those behind their age cohort received lower ratings. However, meta analysis results suggest that job performance and age are unrelated (McEvoy & Cascio, 1989). Therefore age effects noted in other studies may be attributable to bias in ratings.

A meta analysis of race effects confirmed prior findings that similarity in race between rater and ratee results in higher ratings (Kraiger & Ford, 1985). In studies of other characteristics, (1) emotional disability was shown to inflate ratings when clear professional standards were not present (Czajka & DeNisi, 1988), (2) attractiveness inflated ratings of non-managerial women, deflated ratings of managerial women, and had no effect on men (Heilman & Stopeck, 1985), and (3) the effects of rater-ratee acquaintance may depend on rating format (Kingstrom & Mainstone, 1985). Rater affect also appears to influence rating behavior (Tsui & Barry, 1986). However, Cardy & Dobbins (1986)

suggest that affect influenced the rating not by increasing leniency but by introducing noise into the process. By doing so, cognitive processing again becomes an issue.

The research on personal characteristics of raters and ratees is relatively balanced between laboratory and field settings and between student and employed samples. Given the labor markets trends expected to continue over the coming decade, research about age, gender, and race effects seems particularly timely. The labor force is expected to continue to age as the baby boom generation moves through. Additionally, the role of women and minorities is expected to increase, particularly among managerial ranks. If personal characteristics change the way ratings are assigned, these trends represent significant issues to be addressed.

Psychometric Issues

The psychometrics of assessing performance continues to attract research resources. Twelve articles targeted psychometric issues. As Table 3 shows, the majority of these dealt with halo error, and particularly with examining and critiquing competing methods of measuring halo (Pulakos, Schmitt & Ostroff, 1986; Lance & Woehr, 1986; Feldman, 1986). The relationship between halo and accuracy remains controversial. Becker & Cardy (1986) argue that this relationship is ambiguous, that variance and correlational forms of halo may yield either similar or divergent results. Conversely, Fisicaro (1988) concludes that a negative relationship exists between halo and accuracy. Seeming to contradict theories of rating (e.g. Wherry, 1983), halo increased as the opportunity for students to observe performance of professors increased (Jacobs & Kozlowski, 1985). Interestingly, in a

laboratory setting, true halo (true correlations between performance dimensions) was shown to have only a small effect on observed halo (Murphy & Reynolds, 1988).

In other articles on psychometric issues, Smither, Barry & Reilly (1989) in an investigation of the validity of expert true score estimates, report that experts were more accurate than non-experts regardless of the true intercorrelations between performance dimensions. Participation in selection also seem to affect ratings. Schoorman (1988) reported that supervisors who had a say in the hiring decision and who viewed the applicant as favorable, subsequently tended to give lenient performance ratings while those that participated in hiring but viewed the applicant as unfavorable, tended to give more severe ratings. Dickinson (1987) focused on a methodology for studying the validity and accuracy of ratings and suggested combining multitrait-multimethod and person perception designs as a way to improve accuracy. Finally, Sulsky & Balzer (1988) argue that accuracy in performance measurement is lacking due to poor definitions of accuracy, methodological and theoretical limitations of true score development, and the absence of a cohesive theory of performance.

This last point is particularly noteworthy since it corresponds with many of the concerns that managers voiced about the current performance appraisal research. What does accuracy in performance appraisal imply? Many researchers would suggest that accurate appraisal are those that are both reliable and valid, and conceptually near the true score level of performance. The anecdotal evidence we have collected suggests that managers would define accurate appraisals as those that are accepted by employees and allow the identification of relative contribution to organizational effectiveness within the context of the organization and the constraints imposed by the regulatory environment in

which it operates. This definition is quite different than one involving deviations from notions of true scores.

Appraisal Sources

Sources of performance information represent the next major segment of appraisal research and is an area in which the research and professional journals have displayed interest (see Table 3). Self appraisals are receiving more attention than other sources. The literature is sending mixed messages about the usefulness of self appraisals. Using student subjects in a laboratory setting, Farh & Werbel (1986) studied the effect of appraisal purpose on self appraisals and found that self appraisals for evaluative purposes suffered a leniency bias but leniency decreased when appraisals were expected to be validated. Supporting results are reported in a field study by Farh, Werbel & Bedeian (1988). They found that self ratings were more accurate when subjects knew their self appraisals would be compared with other performance ratings. Conflicting results have also been reported. In a field study using soldiers as subjects, Fox and Dinur (1988) report low validity of self ratings regardless of the expectation of validation. Meta-analytic results suggest only moderate relationships exist between self-supervisor and self-peer ratings (Harris & Schaubroeck, 1988). Vance, MacCallum, Coovert & Hedge (1988) report that among a sample of jet engine mechanics, peer, self, and supervisory ratings can be equally valid sources but that performance in some situations can be better evaluated by one source than another. Campbell and Lee (1988) suggest that self appraisals are best when used for developmental rather than evaluative purposes and that self appraisals can improve future performance by creating a self-fulfilling prophecy similar to the effects of goal setting.

In other studies of rating sources, McEvoy & Buller (1987) reported that food processing employees believe that peers are more lenient than supervisors when making ratings used for wage determination. Furthermore, long-term employees are less accepting of peer appraisals than newer employees tend to be. Finally, meta analysis results suggest that the relationship between two sources of performance data, supervisory ratings and objective measures, is weak (r = .27) but tends to be better for relative (.66) versus absolute (.21) rating formats and better for composite (.37) versus global (.19) rating methods (Heneman, 1986).

It seems that an important point is frequently overlooked in research on rating sources. Rather than focusing on who should rate the performance of others and examining the psychometric properties of various rating sources, perhaps research should be examining the propriety of various rating sources under various conditions. That is, when should ratings from alternative sources be used and how should they be integrated with ratings from the immediate manager? The notion of multiple perspectives and 360 degree feedback is certainly useful but the significance is lost when, as in our sample, virtually no weight is placed on the information generated from non-traditional sources of performance information.

Appraisal Feedback

As Table 3 reveals, most of the articles addressing feedback were conducted in field settings, distinguishing this area of research from those dominated by laboratory settings and student subjects. Many of these studies focused on the effects of performance feedback. Discussion of pay and advancement during the performance feedback session

was shown to lead to higher employee satisfaction with the process but did not influence future performance (Dorfman, Stephen & Loveland, 1986). In contrast, Prince & Lawler (1986), report that salary discussion during the appraisal interview has either no relationship or a positive relationship with future behavior. However Pearce and Porter (1986), reported that feedback describing an employee as "satisfactory" (as compared to above average or outstanding) leads to reduced organizational commitment and negative attitudes toward the performance appraisal system.

The dimensionality of feedback was explored in two articles. In a longitudinal field study of university employees Dorfman, Stephen & Loveland (1986) found three dimensions to performance appraisal feedback (being supportive, emphasizing improvement, and discussing pay and advancement). When rating good performers, supervisors tended to provide more support and less emphasis on improvement. Furthermore, Russell & Goode (1988) report that satisfaction with feedback may be multi-dimensional. Managers who are satisfied with the performance appraisal in general, may not be satisfied with the feedback it provides. Rather, satisfaction with feedback is a function of satisfaction with supervisor and the rating received.

Examining feedback source and message in a field setting, Earley (1988) reported that self-generated and specific feedback (versus supervisory-generated and general feedback) were directly related to performance. This agrees with Bannister's (1986) results of a laboratory study using student subjects that concluded that source credibility and content influenced recipient response to the feedback. Ilgen & Moore (1987) explored feedback content in a laboratory setting and found that feedback about quantity lead to

higher quantity, feedback about quality lead to higher quality and feedback about both lead to both.

Rater Training

As Table 3 shows, rater training received limited attention. These studies focused on training to control or eliminate rating errors such as halo and leniency. Training focused on rater errors successfully reduced leniency and halo but also reduced accuracy. These results led Hedge & Kavanagh (1988) to suggest that training should emphasize observation and decision making processes rather than simply error reduction. In a laboratory study using student subjects and video-taped lectures, Athey & McIntyre (1987) showed that frame-of-reference training improved retention of information, improved accuracy and decreased halo. A recent literature review of twenty-four rater training studies suggests the training methods best suited for reducing halo, reducing leniency, and improving accuracy (Smith, 1986). Banks and Roberson (1985) argue that rater training programs do not incorporate standard dimensions of test development and the notion of performance appraisals as "tests" violates accepted standards of test construction. Several practitioner-oriented articles discuss "how to" conduct effective rater training programs.

Performance Appraisal Formats

Of the studies examining formats, behaviorally based methods received the most, although limited, attention. Using a sample of mechanics, Hughes & Prien (1986) evaluated alternative scoring methods for mixed standard scales. They found little differences between the methods and suggest choosing one based on ease of application or

explanation. Prien & Hughes (1987), using a state government sample, showed that mixed standard scales can be used to identify and minimize individual rater error and system-wide problems. In 1987, Murphy & Constans using student subjects and video-taped representations of performance, showed that behavioral anchors may lead to biased recall of performance. But two years later, using the same methodology, Murphy suggested that the earlier results were not likely to be observed in organizational settings (Murphy & Pardaffy, 1989).

Other Issues

The professional journals are replete with articles discussing the performance appraisal practices in various organizations, for various employee groups, and under varying conditions. Case studies and "how to" articles are common (see Table 3). The remainder of the research-based articles were the sole examples of studies on particular issues. Barrett & Kernan (1987) review court cases since Brito vs. Zia arising from terminations based on performance appraisal and offer suggestions to avoid litigation. Using a managerial sample, Greenberg (1986) reports that perceived fairness of performance evaluations depends heavily on the presence of procedural characteristics (e.g., communication, appeals process, job knowledge, consistency) and distributive characteristics (e.g., rating based on performance, action based on rating). Data from our sample supports these contentions. Sackett, Zedeck & Fogli (1988) used a sample of supermarket cashiers to explore errors due to differences between typical and maximum job performance. They found a low correlation between the two and commented on the appropriateness of using procedures that tap maximum versus typical performance. But Napier & Latham (1986)

argue that all this research on performance appraisal seems to matter very little. They report that managers perceive no consequences (good or bad) from conducting thorough performance appraisals and saw little practical value in doing so. However, their sample consisted of managers from only two organizations, therefore, these results may be company or situation specific. In any case, they are inconsistent with the sentiments expressed by the managers in this Fortune 100 sample. Turning away from managers to the more friendly environs of Division 14 Members, Cleveland, Murphy & Williams (1989) derived a four factor structure of the uses of performance appraisal data. The factor structure included (1) between individual uses, (2) within individual uses, (3) system maintenance uses, and (4) documentation uses. Finally, Steel & Mento (1986) reported that 10% of the variance in performance ratings can be attributed to situational factors such as job-induced obstacles, interpersonal and social obstacles, environmental constraints, and policy/procedural constraints. However, these factors did not have effects on more objective, non-judgmental measures of performance.

Insert Table 3 Here

Conclusions

Cognitive processing issues clearly dominated the past five years of performance appraisal research. Prior expectations, prior job knowledge and memory decay were all found to affect performance appraisals. The characteristics of raters and ratees, particularly the effects of gender, also received research attention. However no consensus emerged. For example, no gender (or race) effects were reported in fields studies but student subjects

in lab setting did exhibit gender bias. Halo and accuracy of appraisals were the psychometric topics of choice among researchers. As with gender effects, the relationship between halo and accuracy seems unresolved. Methodologically, paper people or video scenarios with student subjects in laboratory settings is the norm. The clear exceptions are the field studies of the consequences and dimensionality of appraisal feedback. It appears that salary discussions during feedback have either no effect or a positive effect on future performance, but labelling someone as satisfactory rather than above average or outstanding reduces commitment and satisfaction with the appraisal system. Perhaps some of the most interesting recent research was found in the "sole example" studies. Examples include Greenberg's (1986) study of perceived fairness of appraisals as a function of procedural and distributive characteristics, Sackett, Zedeck & Fogli's (1988) use of typical versus maximum job performance and Napier & Latham's (1986) finding that managers see no consequences or practical value from conducting appraisals. These studies point to important issues that have barely been addressed.

The conclusion we draw from this and earlier reviews of appraisal research (e.g. Bernardin & Beatty, 1984; Bernardin & Villanova, 1986; DeNisi, Cafferty & Meglino, 1984; DeNisi & Williams, 1988; Feldman, 1981; Landy & Farr, 1980; Wexley & Klimoski, 1984) is that our knowledge of performance appraisal has expanded greatly in recent years but remains limited. First, the existing research is mainly single-issue studies. For example, the effects of prior expectations, prior knowledge of performance and memory decay have been studied separately from the alternative use of appraisals (administrative or developmental), the characteristics of raters/ratees or the types of scales and formats employed.

Next, certain appraisal issues have received attention while others are virtually ignored. The predominance of studies examine information processing, and psychometric issues. Yet virtually no systematic research exists on how employers actually appraise employees, how they use and evaluate these appraisals and what issues they believe are important.

Implications for Appraisal Research and Decision Making Gap

On one level, the last five years of appraisal research and the interests of managers responsible for the design and administration of appraisal systems do not appear to converge. This is not necessarily a problem since relevancy for decision makers is not the end-all for all research efforts. Yet performance evaluation is an applied subject and as such should eventually lead to improvements in practice. Continued reliance on student samples and laboratory settings is not facilitating the transfer of research into application. We do however, need better understanding of the information processing capabilities and limitations of human decision making. We also need to continue developing a more comprehensive theory of the rating process. Since Wherry's work in the 1950s' (see the Appendix to Landy and Farr, 1983), the collection of studies on information processing is the most serious, concentrated attempt to date to better understand the rating process. In that framework, continued research along those lines is useful indeed. However, attention must be paid to the potential effects of situational or contextual variables. Examining appraisal issues in sterile environments not only limits the generalizability of the results, and it also removes the issues from the attention and interests of human resource decision makers. Doing so limits valuable interaction and application.

The issues of interest to managers and researchers may be different but not mutually exclusive. Managers are concerned with fairness and using appraisal systems which help them manage more successfully. Researchers have been engaged in studies to improve understanding of how information is translated into ratings so that bias and error may be removed. Assuming bias and error contribute to suboptimal decisions, limiting these factors may result in better decision making and ultimately fairer appraisals. Therefore, the gap between research and practice on the cognitive processing issues may lie less in the substance of what is studied and more in the fashion in which the issues are framed, and the methodologies used to explore them.

Nevertheless, some very important issues raised by managers are receiving little or no research attention. Most important of these is the need for more research on fairness in appraisal procedures and results. We found only one study which examined procedural and distributive justice in performance evaluation (Greenberg, 1986). While a considerable body of theoretical discussion exists, appraisals offer unique opportunities to examine the determinants of fair procedures under varying conditions (e.g. different occupational groups, across firms, and perhaps most interestingly under conditions when the distributive results such as pay increases or ratings are judged to be unfair) (Greenberg, 1988, 1990).

We know from the survey that among Fortune 100 firms, the appraisal systems in use were designed by human resource specialists with very limited input from the managers who use them or the employees on whom they are used. We also know that rater training is common, but that those being evaluated (i.e. employees) are not trained or empowered to better use the process for their own development or advancement. Finally, the survey

results showed that 78% of the firms do not hold managers responsible for how well they conduct appraisals.

Determinants of Appraisal System Decisions

The survey data, while descriptive in nature, do present another fundamental research issue. What are the determinants of the managerial choices in performance appraisal system design and administration? The survey makes it clear that managers make different choices about performance appraisal in their firms. What factors lead managers to choose to decentralize policy and administrative responsibilities, to select a forced distribution approach, to use more objective performance indices, to place more weight on second level managers as sources of information, to change the number of levels in the appraisal scale or even to determine the desired performance rating distributions?

Performance appraisal systems may be considered to be a series of decisions which are affected by organizational, environmental, and even dispositional factors. The survey reveals some appraisal decisions that vary across industries and occupational groups and others that do not seem to vary. For example, firms in Aerospace and Automotive industries report spending more time on appraisals. Major differences among industries appear to exist in the number of rating levels used; 67% of those in Food and Beverages report using 5 level scales, only 37% of the firms in Chemicals and Pharmaceuticals do. Similar differences appear across employee groups. While MBO is the most common approach, graphic rating scales are relatively common among nonexempt employees and "no system" is common for hourly (and unionized employees). Perhaps what is required is a more comprehensive approach to the performance appraisal concept (Mohrman, Resnick-

West & Lawler, 1989). Research is also needed which examines the variety of situational variables which affect the choices involved in appraisal design and administration.

While other surveys of performance appraisal practices exist, this survey differs in some important ways from those that have come before. First, the review panel focused on how to identify the issues of interest to both managers and researchers and how to design a study that would advance both the state of research and the state of practice. Second, this was an intensive exploration of a very specific group of organizations; the Fortune Industrial 100. We believe that the 70% response to such a detailed questionnaire make the information gathered from this sample a unique contribution to how performance appraisal is handled in large industrial organizations. Much more research is needed to further delineate the modal setting among other kinds of organizations (e.g., decentralized, smaller, public sector, not for profit, international), other more rigorously defined occupations (e.g., clerical, engineers and scientists, production teams) and under various environmental conditions (notably whether upcoming labor shortages affect performance rating distribution and the use of forced distributions).

Effects of Performance Appraisal Decisions

Napier & Latham (1986) found that managers perceived no consequences, positive or negative, of conducting performance appraisals. Conversely, Longenecker, Sims & Gioia (1987) report that because of actual and perceived negative consequences of accurate appraisal, some managers knowingly make ratings that are inaccurate. With these exceptions, no research in the past five years has examined the effects of different appraisal system designs or processes on employee attitudes and behaviors or on organization performance. Potential research topics include: Are the effects of appraisal procedures or results perceived to be fair? Does the use of three, four, or five rating levels make any difference? What are the effects of the time devoted to appraisals, the nature of the measures used, or the use of ranking or forced distributions? Are Napier & Latham's observations correct? Deming argues strongly that performance appraisal does have consequences (Scholtes, 1987). So serious are the consequences of appraisal feedback, that Deming urges firms to cease all individual performance appraisals and to evaluate unit or plant level performance instead. This approach, reportedly widely used in Japanese factories, focuses on assisting those whose performance is "out of the system." Deming's notions have received little attention in practitioner literature and no attention from researchers. The possibility of no feedback seems difficult to attain. Even without formal individual appraisals, informal appraisals by team leaders and peers seems inevitable and perhaps potentially less systematic and more vulnerable to biases. Here again the potential for cognitive information processing in informal appraisal or team settings may be a rich research vein.

The effects of skew in performance rating distributions on managerial pay decisions and employee attitudes needs to be examined. Highly skewed ratings affect the distribution of merit pay increases. In fact, some argue that the size of the merit fund pool affects the ratings skew (Milkovich & Newman, 1987). Smaller funds force managers to give higher ratings which result in smaller average increases for high rated performers. These smaller pay increases coupled with high ratings deliver mixed signals to employers and may affect their attitudes toward merit pay. Our study makes the extent of performance distribution skew in the Fortune 100 explicit for the first time. Some of our results suggest that tracing changes in appraisal system decisions over time may be useful. For example, some firms reported shifting to forced distributions in an effort to correct their ratings skew. Such changes may lower employees' satisfaction with appraisals on the one hand, but improve employees' satisfaction with their pay increases on the other. More generally, it seems desirable to understand why firms make changes in their systems and to examine the effects of these changes.

References

Athey TR, McIntyre RM. (1987). Effect of rater training on rater accuracy: Levels-of-processing theory and social facilitation theory perspectives. Journal of Applied Psychology, 72(4), 567-572.

Balzer WK. (1986). Biases in the recording of performance-related information: The effects of initial impression and centrality of the appraisal task. Organizational Behavior and Human Decision Processes, 37, 329-347.

Banks CG, Murphy KR. (1985). Toward narrowing the research-practice gap in performance appraisal. Personnel Psychology, 38, 335-345.

Banks CG, Roberson L. (1985). Performance appraisers as test developers. Academy of Management Review, 10(1), 128-142.

Bannister BD. (1986). Performance outcome feedback and attributional feedback: Interactive effects on recipient responses. Journal of Applied Psychology, 71(2), 203-210.

Barrett GV, Kernan MC. (1987). Performance appraisal and terminations: A review of court decisions since Brito v. Zia with implications for personnel practices. Personnel Psychology, 40, 489-503.

Becker BE, Cardy RL. (1986). Influence of halo error on appraisal effectiveness: A conceptual and empirical reconsideration. Journal of Applied Psychology, 71(4), 662-671.

Bedeian, AG. (1989). Totems and taboos: Undercurrents in the management discipline. The Academy of Management News, 19(4), 2-6.

Benedict ME, Levine EL. (1988). Delay and distortion: Tacit influences on performance appraisal effectiveness. Journal of Applied Psychology, 73(3), 507-514.

Bernardin HJ. (1986). Subordinate appraisal: A valuable source of information about managers. Human Resource Management, 25(3), 421-439.

Bernardin HJ, Abbott J. (1985). Predicting (and preventing) differences between self and supervisory appraisals. Personnel Administrator, 151-157.

Bernardin HJ, Beatty RW. (1984). Performance appraisal: Assessing human behavior at work. Boston: Kent.

Bernardin HJ, Beatty RW. (1987). Can subordinate appraisals enhance managerial productivity? Sloan Management Review, Summer, 63-73.

Bernardin HJ, Buckley MR. (1981). A consideration of strategies in rater training. Academy of Management Review, 6, 205-212.

Bernardin HJ, Villanova P. (1986). Performance appraisal. In EA Locke (Ed.), Generalizing from the Laboratory to Field Settings, 43-62. Lexington, MA: Lexington.

Brown DG. (1987). Development of performance standards: A practical guide. Public Personnel Management, 16(2), 93-114.

Brumback BG (1988). Some ideas, issues and predictions about performance management. Public Personnel Management, 17, 387-402.

Buford JA, Burkhalter BB, Jacobs GT. (1988). Link job descriptions to performance appraisals. Personnel Journal, 132-140.

Campbell DJ, Lee C. (1988). Self-appraisal in performance evaluation: Development versus evaluation. Academy of Management Journal, 13(2), 302-314.

Cardy RL, Dobbins GH. (1986). Affect and appraisal accuracy: Liking as an integral dimension in evaluating performance. Journal of Applied Psychology, 71(4), 672-678.

Cayer M, DiMattia DJ, Wingrove J. (1988). Conquering evaluation fear. Personnel Administrator, June, 97-197.

Cocheu T. (1986). Performance appraisal: A case in points. Personnel Journal, September, 48-55.

Conference Board. (1977).

Cleveland JN, Murphy KR, Williams RE. (1989). Multiple uses of performance appraisal: Prevalence and correlate. Journal of Applied Psychology, 74, 130-135.

Czajka JM, DeNisi AS. (1988). Effects of emotional disability and clear performance standards on performance ratings. Academy of Management Journal, 31, 394-404.

Dearden J. (1987). Measuring profit center managers. Harvard Business Review, September-October, 84-88.

Deets NR, Taylor DT. (1986). How xerox improved its performance appraisals. Personnel Journal, April, 50-52.

DeNisi AS, Cafferty TP, Meglino BM. (1984). A cognitive model of the performance appraisal process: A model and research propositions. Organizational Behavior and Human Performance, 33, 360-396.

DeNisi AS, Robbins T, Cafferty TP. (1989). Organization of information used for performance appraisals: Role of diary-keeping. Journal of Applied Psychology, 74(1), 124-129.

DiNisi AS, William KJ. (1988). Cognitive approaches to performance appraisal. In KM Rowland & GR Ferris (Eds.), Research in Personnel and Human Resources Management, Vol.6, 109-155. Greenwich, CT: Jai Press.

Dickinson TL. (1987). Designs for evaluating the validity and accuracy of performance ratings. Organizational Behavior and Human Decision Processes, 40, 1-21.

Dipboye RL. (1985). Some neglected variables in research on discrimination in appraisals. Academy of Management Review, 10(1), 116-127.

Dobbins GH, Cardy RL, Truxillo DM. (1988). The effects of purpose of appraisal and individual differences in stereotypes of women on sex differences in performance ratings: A laboratory and field study. Journal of Applied Psychology, 73(3), 551-558.

Dorfman PW, Stephan WG, Loveland J. (1986). Performance appraisal behaviors: Supervisor perceptions and subordinate reactions. Personnel Psychology, 39, 579-597.

Drazin R, Auster ER. (1987). Wage differences between men and women: Performance appraisal ratings vs. salary allocation as the locus of bias. Human Resource Management, 26(2), 157-168.

Earley PC. (1988). Computer-generated performance feedback in the magazine-subscription industry. Organizational Behavior and Human Decision Processes, 41, 50-64.

Edwards MR, Sproull JR. (1985). Making performance appraisals perform: The use of team evaluation. Personnel, 28-32.

Farh JL, Werbel JD. (1986). Effects of purpose of the appraisal and expectation of validation on self-appraisal leniency.

Farh JL, Werbel JD, Bedeian AG. (1988). An empirical investigation of self-appraisal-based performance evaluation. Personnel Psychology, 41, 141-156.

Fedor DR, Buckley MR. (1988). Issues surrounding the need for more frequent monitoring of individual performance in organizations. Public Personnel Management, 17, 435-442.

Feldman JM. (1981). Beyond attribution theory: Cognitive processes in performance appraisal. Journal of Applied Psychology, 66, (2), 127-148.

Feldman JM. (1986). A note on the statistical correction of halo error. Journal of Applied Psychology, 71(1), 173-176.

Feldman JM, Camburn A, Gatti GM. (1986). Shared distinctiveness as a source of illusory correlation in performance appraisal. Organizational Behavior and Human Decision Processes, 37, 34-59.

Ferris GR, Yates VL, Gilmore DC, Rowland KM. (1985). The influence of subordinate age on performance ratings and causal attributions. Personnel Psychology, 38, 545-557.

Fisicaro SA. (1988). A reexamination of the relation between halo error and accuracy. Journal of Applied Psychology, 73(2), 239-244.

Fox S, Dinur Y. (1988). Validity of self-assessment: A field evaluation. Personnel Psychology, 41, 581-592.

Friedman MG. (1986). 10 steps to objective appraisals. Personnel Journal, 66-71.

Fulk J, Brief AP, Barr SH. (1985). Trust-in-supervisor and perceived fairness and accuracy of performance evaluations. Journal of Business Research, 13, 299-313.

Gabor A. (1989). Catch a falling star system. U.S. News and World Report, June 5, 43-44.

Gellerman SW, Hodgson WG. (1988). Cyanamid's new take on performance appraisal. Harvard Business Review, May-June, 36-41.

Gibb P. (1985). Appraisal goals and controls. Personnel Journal, 89-93.

Girard R. (1988). Is there a need for performance appraisals? Personnel Journal, 89-90.

Goodell R. (1988). Room for improvement. Personnel Administrator, June, 132-142.

Gomez-Mejia LR, Page RC, Tornow WW. (1985). Improving the effectiveness of performance appraisal. Personnel Administrator, January, 74-81.

Greenberg J. (1986). Determinants of perceived fairness of performance evaluations. Journal of Applied Psychology, 71, 340-342.

Greenberg J. (1988). Using explanations to manage impressions of performance appraisal fairness. In J. Greenberg & R. Bies (Chairs). Communicating Fairness in Organizations. Symposium presented at the Meeting of the Academy of Management, Anaheim, CA, August 1988.

Greenberg J. (1990). Looking fair vs. being fair: Making impressions of organizational justice. Forthcoming in BM Staw & LL Cummings (Eds.), Research in Organizational Behavior, Vol. 12. Greenwich, CT: Jai Press.

Hall TC. (1987). Starting over. Training and Development Journal, December, 60-62.

Harper SC. (1986). Adding purpose to performance reviews. Training and Development Journal, 53-55.

Harris C. (1988). A comparison of employee attitudes toward two performance appraisal systems. Public Personnel Management, 17, 443-456.

Harris MM, Schaubroeck J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. Personnel Psychology, 41, 43-62.

Hedge JW, Kavanagh MJ. (1988). Improving the accuracy of performance evaluations: Comparison of three methods of performance appraiser training. Journal of Applied Psychology, 73, 68-73.

Heilman ME, Stopeck MH. (1985). Being attractive, advantage or disadvantage? Performance-based evaluations and recommended personnel actions as a function of appearance, sex, and job type. Organizational Behavior and Human Decision Processes, 35, 202-215.

Heneman RL. (1986). The relationship between supervisory ratings and results-oriented measures of performance: A meta-analysis. Personnel Psychology, 39, 811-826.

Heneman RL, Wexley KN, Moore ML. (1987). Performance-rating accuracy: A critical review. Journal of Business Research, 15, 431-448.

Hewitt (1989).

Hogan EA. (1987). Effects of prior expectations on performance ratings: A longitudinal study. Academy of Management Journal, 30, 354-368.

Howard CG. (1987). Out of sight - not out of mind. Personnel Administrator, 82-90.

Huber VL, Neale MA, Northcraft GB. (1987). Judgment by Heuristics: Effects of ratee and rater characteristics and performance standards on performance-related judgments. Organizational Behavior and Human Decision Processes, 40, 149-169.

Huber VL, Podsakoff PM, Todor WD. (1986). An investigation of biasing factors in the attributions of subordinates and their supervisors. Journal of Business Research, 14, 83-98.

Hughes GL, Prien EP. (1986). An evaluation of alternate scoring methods for the mixed standard scale. Personnel Psychology, 39, 839-847.

Hyde AC. (1988). The new environment for compensation and performance evaluation in the public sector. Public Personnel Management, 17, 351-358.

Ilgen DR, Moore CF. (1987). Types and choices of performance feedback. Journal of Applied Psychology, 72(3), 401-406.

Jacobs R, Kozlowski SW. (1985). A closer look at halo error in performance ratings. Academy of Management Journal, 28, 201-212.

Jolly JP, Reynolds TJ, Slocum JW. (1988). Application of the means-end theoretic for understanding the cognitive bases of performance appraisal. Organizational Behavior and Human Decision Processes, 41, 153-179.

Kane JS, Freeman KA. (1986a). MBO and performance appraisal: A mixture that's not a solution, part 1. Personnel, 26-36.

Kane JS, Freeman KA. (1986b). MBO and performance appraisal: A mixture that's not a solution, part 2. Personnel, 26-32.

Kaufman R. (1988). Preparing useful performance indicators. Training and Development Journal, 80-83.

Kerr JL. (1988). Strategic control through performance appraisal and rewards. Human Resource Planning, 2(3), 215-223.

Kiechell (1989). Fortune, June 19, 201.

Kingstrom PO, Mainstone LE. (1985). An investigation of the rater-ratee acquaintance and rater bias. Academy of Management Journal, 28, 641-653.

Kinicki AJ, Griffeth RW. (1985). The impact of sex-role stereotypes on performance ratings and causal attributions of performance. Journal of Vocational Behavior, 27, 155-170.

Kirkpatrick DL. (1986). Performance appraisal: Your questions answered. Training and Development Journal, 68-71.

Kozlowski SW, Kirsch MP. (1987). The systematic distortion hypothesis, halo, and accuracy: An individual-level analysis. Journal of Applied Psychology, 72(2), 252-261.

Kozlowski SW, Kirsch MP, Chao GT. (1986). Job knowledge, ratee familiarity, conceptual similarity and halo error: An exploration. Journal of Applied Psychology, 71(1), 45-49.

Kraiger K, Ford JK. (1985). A meta-analysis of ratee race effects in performance ratings. Journal of Applied Psychology, 70(1), 56-65.

Krzystofiak F, Cardy R, Newman J. (1988). Implicit personality and performance appraisal: The influence of trait inferences on evaluations of behavior. Journal of Applied Psychology, 73(3), 515-521.

Lance CE, Woehr DJ. (1986). Statistical control of halo: Clarification from two cognitive models of the performance appraisal process. Journal of Applied Psychology, 71(4), 679-685.

Landy FS, Farr JL. (1980). Performance rating. Psychological Bulletin, 87, 72-107.

Landy FS, Farr JL. (1983). The measurement of work performance. Methods, theory, and applications. Orlando, Florida: Academic Press.

Lanza P. (1985). Team appraisals. Personnel Journal, 47-51.

Latham GP, Wexley KN, Purcell ED. (1975). Training managers to minimize rating errors in the observation of behavior. Journal of Applied Psychology, 60, 550-555.

Laud (1983).

Laumeyer & Beebe (1988). Employees and their appraisals. Personnel Administrator, December, 76-80.

Lawrence BS. (1988). New wrinkles in the theory of age: Demography, norms, and performance ratings. Academy of Management Journal, 31, 309-337.

Lawrie JW. (1989). Your performance: Appraise it yourself! Personnel, 21-23.

Levy M. (1989). Almost-perfect performance appraisals. Personnel Journal, 76-83.

Locher AH, Teel KS. (1988). Appraisal trends. Personnel Journal, 139-145.

Locke EA. (1986). Generalizing from Laboratory to Field Settings. Lexington, MA: Lexington Books.

Longenecker CO, Gioia DA. (1988). Neglected at the top - executives talk about executive appraisals. Sloan Management Review, Winter, 41-47.

Longenecker CO, Sims HP, Gioia DA. (1987). Behind the mask: The politics of employee appraisal. The Academy of Management EXECUTIVE, 1(3), 183-193.

Lord RG, Maher KJ. (1989). Cognitive processes in industrial and organizational psychology. In CL Cooper & I Robertson (Eds.) International Review of Industrial and Organizational Psychology, 49-91. Jon Wiley & Sons, Ltd.

Lowe TR. (1986). Eight ways to ruin a performance review. Personnel Journal, 60-62.

Martin DC, Bartol KM. (1986). Training the raters: A key to effective performance appraisal. Public Personnel Management, 15(2), 101-109.

McBriarty MA. (1988). Performance appraisal: Some unintended consequences. Public Personnel Management, 17, 421-434.

McEvoy GM. (1988). Evaluating the boss. Personnel Administrator, 115-120.

McEvoy GM, Buller PF. (1987). User acceptance of peer appraisals in an industrial setting. Personnel Psychology, 40, 785-797.

McEvoy GM, Cascio WF. (1989). Cumulative evidence of the relationship between employee age and job performance. Journal of Applied Psychology, 74(1), 11-17.

Metz EJ. (1988). Designing legally defensible performance appraisal systems. Training and Development Journal, 47-51.

Meyer JP, Paunonen SV, Gallatly IR, Goffin RD, Jackson DN. (1989). Organizational commitment and job performance: It's the nature of the commitment that counts. Journal of Applied Psychology, 74(1), 152-156.

Milkovich TG, Newman JM. (1987). Compensation. Plano, TX: Business Publications, Inc.

Mohrman AM, Resnick-West SM, Lawler EE, III. (1989). Designing Performance Appraisal Systems. San Francisco, CA: Jossey-Bass.

Mount MK, Thompson DE. (1987). Cognitive categorization and quality of performance ratings. Journal of Applied Psychology, 72(2), 240-246.

Murphy KR, Balzer WK. (1986). Systematic distortions in memory-based behavior ratings and performance evaluations: Consequences for rating accuracy. Journal of Applied Psychology, 71(1), 39-44.

Murphy KR, Balzer WK, Lockhart MC, Eisenman EJ. (1985). Effects of previous performance on evaluations of present performance. Journal of Applied Psychology, 70(1), 72-84.

Murphy KR, Constans JI. (1987). Behavioral anchors as a source of bias in rating. Journal of Applied Psychology, 72(4), 573-577.

Murphy KR, Gannett BA, Herr BM, Chen JA. (1986a). Effects of subsequent performance on evaluations of previous performance. Journal of Applied Psychology, 71(3), 427-431.

Murphy KR, Herr BM, Lockhart MC, Maguire E. (1986b). Evaluating the performance of paper people. Journal of Applied Psychology, 71(4), 654-661.

Murphy KR, Pardaffy VA. (1989). Bias in behaviorally anchored rating scales: Global or scale-specific? Journal of Applied Psychology, 74(2), 343-346.

Murphy KR, Reynolds DH. (1988). Does true halo affect observed halo? Journal of Applied Psychology, 73(2), 235-238.

Naffziger DW. (1985). BARS, RJPs and recruiting. Personnel Administrator, 85-96.

Napier NK, Latham GP. (1986). Outcome expectancies of people who conduct performance appraisals. Personnel Psychology, 39, 827-837.

Nanry C. (1988). Performance linked training. Public Personnel Management, 17, 457-464.

Nathan BR, Alexander RA. (1985). The role of inferential accuracy in performance rating. Academy of Management Review, 10(1), 109-115.

Oliver JE. (1985). Performance appraisals that fit. Personnel Journal, 66-71.

Pearce JL, Porter LW. (1986). Employee responses to formal performance appraisal feedback. Journal of Applied Psychology, 71(2), 211-218.

Perry J, Petrakis BA. (1988). Part I - Developments in organizational performance. Can pay for performance succeed in government? Public Personnel Management, 17, 359-368.

Prien EP, Hughes GL. (1987). The effect of quality control revisions on mixed standard scale rating errors. Personnel Psychology, 40, 815-823.

Prince JB, Lawler EE. (1986). Does salary discussion hurt the developmental performance appraisal? Organizational Behavior and Human Decision Processes, 37, 357-375.

Pulakos ED, Schmitt N, Ostroff C. (1986). A warning about the use of a standard deviation across dimensions within ratees to measure halo. Journal of Applied Psychology, 71(1), 29-32.

Reed PR, Kroll MJ. (1985). A two-perspective approach to performance appraisal. Personnel, 51-57.

Regel RW, Hollmann RW. (1987). Gauging performance objectively. Personnel Administrator, 74-78.

Reinhardt C. (1985). The state of performance appraisal: A literature review. Human Resource Planning, 105-110.

Romanoff (1989). The ten commandments of performance management. Personnel, January, 24-89.

Russell JS, Goode DL. (1988). An analysis of managers' reactions to their own performance appraisal feedback. Journal of Applied Psychology, 73, 63-67.

Sackett PR, Zedeck S, Fogli L. (1988). Relations between measures of typical and maximum job performance. Journal of Applied Psychology, 73(3), 482-486.

Scherkenbach (1985). Performance appraisal and quality: Ford's new philosophy. Quality Progress. April, 40-46.

Schmidt FL, Hunter JE, Outerbridge AN. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. Journal of Applied Psychology, 71(3), 432-439.

Schmitt N, Noe RA, Gottschalk R. (1986). Using the lens model to magnify raters' consistency, matching, and shared bias. Academy of Management Journal, 29, 130-139.

Schneier CE (1989). Capitalizing on performance management, recognition, and reward systems. Compensation and Benefits Review, 21, 20-30.

Schneier CE, Beatty RW, Baird LS. (1986). Creating a performance management system. Training and Development Journal, 74-49.

Schneier CE, Beatty RW, Baird LS. (1986). How to construct a successful performance appraisal system. Training and Development Journal, 38-42.

Schneier CE, Geis A, Wert JA. (1987). Performance appraisals: No appointment needed. Personnel Journal, 80-87.

Scholtes PR. (1987). An elaboration on Deming's teachings on performance appraisal. Madison,WI: Joiner Associates, Inc.

Schoorman FD. (1988). Escalation bias in performance appraisals: An unintended consequence of supervisor participation in hiring decisions. Journal of Applied Psychology, 73(1), 58-62.

Shapiro GL, Dessler G. (1985). Are self appraisals more realistic among professionals or nonprofessionals in health care? Public Personnel Management, 14(3), 285-290.

Shore LM, Thornton GC. (1986). Effects of gender on self - and supervisory ratings. Academy of Management Journal, 29, 115-129.

Sims RR, Veres JG, Heninger SM. (1987). Training appraisers: An orientation program for improving supervisory performance ratings. Public Personnel Management, 16(1), 37-46.

Slattery PD. (1985). Performance appraisal without stress. Personnel Journal, 49-52.

Smith DE. (1986). Training programs for performance appraisal: A review. Academy of Management Review, 11(1), 22-40.

Smither JW, Barry SR, Reilly RR. (1989). An investigation of the validity of expert true score estimates in appraisal research. Journal of Applied Psychology, 74(1), 143-151.

Smither JW, Reilly RR, Buda R. (1988). Effect of prior performance information on ratings of present performance: Contrast versus assimilation revisited. Journal of Applied Psychology, 73(3), 487-496.

Smither JW, Reilly RR. (1987). True intercorrelation among job components, time delay in rating, and rater intelligence as determinants of accuracy in performance ratings. Organizational Behavior and Human Decision Processes, 40, 3369-391.

Srinivas S, Motowidlo SJ. (1987). Effects of raters' stress on the dispersion and favorability of performance ratings. Journal of Applied Psychology, 72(2), 247-251.

Steel BS. (1985). Participative performance appraisal in Washington: An assessment of post-implementation receptivity. Public Personnel Management, 14(2), 153-171.

Steel RP, Mento AJ. (1986). Impact of situational constraints on subjective and objective criteria of managerial job performance. Organizational Behavior and Human Decision Processes, 37, 254-265.

Steiner DD, Rain JS. (1989). Immediate and delayed primacy and recency effects in performance evaluation. Journal of Applied Psychology, 74(1), 136-142.

Sulsky LM, Balzer WK. (1988). Meaning and measurement of performance rating accuracy: Some methodological and theoretical concerns. Journal of Applied Psychology, 73, 497-506.

Taylor RR, Smith L. (1987). Performance standards: Developing an employee appraisal system to enhance productivity in a county government. PAQ, Summer, 217-238.

Thompson DE, Thompson TA. (1985). Task-based performance appraisal for blue-collar jobs: Evaluation of race and sex effects. Journal of Applied Psychology, 70(4), 747-753.

Thorndike RL. (1949). Personnel selection: Test and measurement techniques. New York: Wiley.

Tsui AS, Barry B. (1986). Interpersonal affect and rating errors. Academy of Management Journal, 29, 586-599.

Vance RJ, MacCallum RC, Coovert MD, Hedge, JW. (1988). Construct validity of multiple job performance measures using confirmatory factor analysis. Journal of Applied Psychology, 73(1), 74-80.

Vines LS. (1988). How's the boss dong? Human Resource Executive, 36-38.

VonGlinow MA. (1985). Reward strategies for attracting, evaluating, and retaining professionals. Human Resource Management, 24(2), 191-206.

Wagel WH. (1987). Performance appraisal with a difference. Personnel, February, 4-6.

Wehrenberg SB. (1988). Train supervisors to measure and evaluate performance. Personnel Journal, 77-79.

Wexley KN, Klimoski R. (1984). Performance appraisal: An update. In K Rowland & G Ferris (Eds.) Research in personnel and human resources management, Vol. 2, 35-79.

Wexley KN, Snell SA. (1987). Managerial power: A neglected aspect of the performance appraisal interview. Journal of Business Research, 15, 45-54.

Wherry RJ. (1983). Wherry's theory of rating. In FL Landy & JL Farr (Eds.), The measurement of work eprformance. Methods, theory, and applications. Orlando, FL: Academic Press.

Wight (1985).

Williams KJ, DeNisi AS, Blencoe AG, Cafferty TP. (1985). The role of appraisal purpose: Effects of purpose on information acquisition and utilization. Organizational Behavior and Human Decision Processes, 35, 314-339.

Williams KJ, DeNisi AS, Meglino BM, Cafferty TP. (1986). Initial decisions and subsequent performance ratings. Journal of Applied Psychology, 71(2), 189-195.

Woods JG, Dillion T. (1985). The performance review approach to improving productivity. Personnel, March, 20-27.

Wyatt (1989).

Table 1.

Most important Performance Appraisal issues.

1.	A Performance Appraisal System accepted by those doing the rating.
2.	Employee sense of being treated fairly by the Performance Appraisal process.
3.	Employee belief that Performance Appraisal results are fair.
4.	Type of feedback given in appraisal interviews.
5.	Managers' judgment about usefulness of Performance Appraisal as tool to help manage.
6.	Performance Appraisal System accepted by those being rated.

Table 2.

Most important uses for Performance Appraisal information

1.	Improving work performance.
2.	Administering pay on the basis of merit.
3.	Advising employees of work expectations.
4.	Counseling employees.
5.	Making promotion decisions.
6.	Motivating employees.
7.	Assessing employee potential.
8.	Identifying training needs.
9.	Better working relations between managers and employees.
10.	Helping employees set career goals.
11.	Assigning work more efficiently.
12.	Making transfer decisions.
13.	Making decisions about layoffs or terminations.
14.	Assisting in long-range planning.
15.	Validating hiring procedures.
16.	Justifying other managerial actions.

Table 3
Summary of Research and Managerial Performance Appraisal Literature

	Methodology						
Primary Issues Explored	**Field Study**	**Lab Study**	**Lit Review**	**Survey**	**Discussion**	**Case Study**	**"How To"**
Information Processing	Hogan (AMJ) 1987 29 Teller Supervisors Jolly, et al. 1988 (OBHDP) 22 Nurse Sup. Mount & Thompson 1987 (JAP) 174 Mgrs. Schmitt, et al. 1986 (AMJ) 153 Sch. Admin. *Huber, et al. 1986* *(JBR) 593* *Govt. Emp.*	Balzer 1986 (OBHDP) 80 Students Video DeNisi, et al. 1989 (JAP) 63 Students Video Feldman, et al. 1986 (OBHDP) 152 Students Paper People Huber et al. 1987 (OBHDP) 229 Mgrs. Paper People Kinicki & Griffeth 1985 (JVB) 143 Students Paper People Kozlowski & Kirsch 1987 (JAP) 227 Students Paper People Kozlowski, et al. 1986 (JAP) 192 Students Paper People Kzystofiak, et al. 1988 (JAP) 200 Students Paper People Murphy & Balzer 1986 (JAP) 46 Students/30 Students Video Murphy, et al. 1985 (JAP) 82 Students/71 Students Video Murphy, et al. 1986 (JAP) 183 Students Video Smither & Reilly 1987 (OBHDP) 90 Students Video Smither, et al. 1988 (JAP) 51, 53, 109 Students Video Srinivas & Motowidlo 1987 (JAP) 120 Students Video Steiner & Rain 1989 (JAP) 541 Students Video Williams et al. 1985 (OBHDP) 120 Students Paper People Williams, et al. 1986 (JAP) 70 Students Video			Dipboye 1985 (AMR) Nathan & Alexander 1985 (AMR) Schmidt, et al. 1986 (JAP)		

Table 3 Continues

Table 3 Continued
Summary of Research and Managerial Performance Appraisal Literature

	Methodology						
Primary Issues Explored	**Field Study**	**Lab Study**	**Lit Review**	**Survey**	**Discussion**	**Case Study**	**"How To"**
Rater/Ratee Characteristics							
Sex/Gender	Dobbins, et al. 1988 (JAP) 810 Students Study 2 Shore & Thorton 1986 (AMJ) 70 Assemblers and Supervisors Thompson & Thompson 1985 (JAP) 206 Industrial Workers 233 Petroleum Workers *Drazin & Auster 1987 (HRM) 2631 Fin. Serv. Employees*	Benedict & Levine 1988 (JAP) 64 Students Video Dobbins, et al. 1988 (JAP) 103 Students Study 1					
Age	Ferris, et al. 1985 (PP) 81 Nurses Lawrence 1988 (AMJ) 390 Mgrs.		McEvoy & Cascio 1989 (JAP) Meta Analysis				
Race			Kraiger & Ford 1985 (JAP) Meta Analysis				
Disability		Czajka & De Nisi 1988 (AMJ) 60 Students Video					
Attractiveness		Heilman & Stopeck 1985 (OBHDP) 34 Students Paper People					
Acquaintance	Kingstrom & Mainstone 1985 (AMJ) 80 Sales Supervisors						
Affect	Tsui & Barry 1986 (AMJ) 349 Mgrs.	Cardy & Dobbins 1986 (JAP) 66 Students Paper People					

Table 3 Continued
Summary of Research and Managerial Performance Appraisal Literature

	Methodology						
Primary Issues Explored	**Field Study**	**Lab Study**	**Lit Review**	**Survey**	**Discussion**	**Case Study**	**"How To"**
Psychometrics							
Halo	Jacobs & Kozlowski 1985 (AMJ) 1031 Students	Becker & Cardy 1986 (JAP) 170 Students Paper People Murphy & Reynolds 1988 (JAP) 128 Students Video			Feldman 1986 (JAP) Fisicaro 1988 (JAP) Lance & Woehr 1986 (JAP) Pulakos, et al. 1986 (JAP)		
Accuracy		Smither, et al. 1989 (JAP) 42 Students Video	*Heneman, et al. 1987 (JBR) Conventional*		Dickenson 1987 (OBHDP)	Sulsky & Balzer 1988 (JAP)	
Leniency	Schoorman 1988 (JAP) 354 Clerical						
Feedback	Dorfman, et al.1986 (PP) 242 Univ. Emp. Earley 1986 (OBHDP) 60 Mag. Subscription Processors Pearce & Porter 1986 (JAP) 348 Fed. Employees Prince & Lawler 1986 (OBHDP) 9 Orgs. Russell & Goode 1988 (JAP) 204 Mgrs. *Wexley & Snell 1987 (JBR) 98 Sales Mgrs.*	Bannister 1986 (JAP) 149 Students VIdeo Ilgen & Moore 1987 (JAP) 222 Students Video				*Nanry 1988 (PPM)*	*Harper 1986 (TDJ)* *Wight 1985 (PA)*
Raters/Appraisal Sources			Heneman, R. 1986 (PP) Meta Analysis				
Self	Farh, et al. 1988 (PP) 88 Faculty Fox & Dinur 1988 (PP) 857 Soldiers Vance, et al. 1988 (JAP) 256 Mechanics *Shapiro & Dessler 1985 (PPM) 146 Health Care Supervisors*	Farh & Werbel 1986 (JAP) Video	Harris & Schaubroeck 1988 (PP) Meta Analysis	*Steel 1985 (PPM) 1000 State Emp.*	Campbell & Lee 1985 (AMJ) *Bernardin & Abbott 1985 (PA)*		*Lawrie 1989 (P)*
Peers	McEvoy & Buller 1987 Employees						*Edward & Sproull 1* *(P)* *Lanza 1985 (PJ)*

Table 3 Continued
Summary of Research and Managerial Performance Appraisal Literature

	Methodology						
Primary Issues Explored	**Field Study**	**Lab Study**	**Lit Review**	**Survey**	**Discussion**	**Case Study**	**"How To"**
Subordinates					*Bernardin 1986 (HRM)* *Kiechel 1989 (F)* *McEvoy 1988 (PA)* *Vines 1988 (HRE)*	*Bernardin & Beatty 1987 (SMR)*	
Rater Training	Hedge & Kavanagh 1988 (JAP) 52 Superv.	Athey & McIntyre 1987 (JAP) 108 Students Video	Smith 1986 (AMR) (24 Studies) Conventional		Banks & Roberson 1985 (AMR) *Wehrenberg 1988 (PJ)*		*Brown 1987 (PPM)* *Kaufman 1988 (TDJ* *Martin & Bartol 1986 (PPM)* *Sims, et al. 1987 (PPM)*
Formats							*Oliver 1985 (PJ)*
BARS		Murphy & Constans 1987 (JAP) 180 Students Video Murphy & Pardaffy 1989 (JAP) 94 Students Video			*Naffziger 1985 (PA)*		
MSS	Hughes & Prein 1986 (PP) 49 Mechanics Prein & Highes 1987 (PP) 2000 Govt. Supervisors					*McBriarty 1988 (PPM)*	
Formats							
MBO				*Kane & Freeman 1986a (P)* *Kane & Freeman 1986b (P)*			*Gibb 1985 (P)*
Graphic Scales	*Harris 1988 (PPM)* *138 Govt Employees* *330 Govt Employees*						

Table 3 Continued
Summary of Research and Managerial Performance Appraisal Literature

	Methodology						
Primary Issues Explored	**Field Study**	**Lab Study**	**Lit Review**	**Survey**	**Discussion**	**Case Study**	**"How To"**
Perf. App. Practices			*Reinhardt 1985 (HRP) Conventional*	*Locher & Teal 1988 (PJ) 324 Small Bus.*	*Longenecker & Gioia 1988 (SMR) Executives*	*Cayer, et al. 1988 (PA)* *Cochen 1986 (PJ)* *Deets & Taylor 1986 (PJ)* *Gellerman & Hodgson 1988 (HBR)* *Gomez-Mejia, et al. 1985 (PA)* *Goodell 1988 (PA)* *Hall 1987 (TDJ)* *Laumeyer & Beebe 1988 (PA)* *Taylor & Smith 1987 (PAQ)* *Wagel 1987 (P)* *Woods & Dillion 1985 (P)*	*Buford et al. 1988 (PJ) J.D. & P.A.* *Dearden 1987 (HBR) Profit Center Mgrs.* *Girard 1988 (PJ) 2 Buckets - "middle big"* *Howard 1987 (PA) Expatriots* *Reed & Kroll 1985 (I* *Regel & Hollman 198 (PA) Accountants*
Performance Mgmt.					*Brumback 1988 (PPM)*	*Perry & Petrakis 1988 (PPM)* *Schneier 1989 (CBR)*	*Kirkpatrick 1986 (TD* *Levy 1989 (PJ)* *Romanoff 1989 (P)* *Schneier, et al. 1987* *Schneier, et al. 1986a (TDJ)* *Schneier, et al. 1986b (TDJ)*
Others	Meyer, et al. 1989 (JAP) 114 Managers Org. Commitment Napier & Lathan 1986 (PP) 32 & 39 Managers Consequences Sackett, et al. 1988 (JAP) 635 & 735 Cashiers Typical vs. Max. Performance Steel & Mento 1986 (OBHDP) 438 Managers Situational Constraints *Fulk, et al. 1985 (JBR) 198 Engineers Fairness* *Kerr 1988 (HRP) 20 Orgs. Congruence* *Longenecker, et al. 1987 (AME) 60 Execs. Uses of PA*		Barrett & Kernan 1987 (PP) Conventional Legal Issues Murphy et al. 1986b (JAP) Meta Analysis Paper People	Cleveland et al. 1989 (JAP) 106 APA, D14 Members. Uses. Greenberg 1986 (JAP) 217 Managers Fairness	Banks & Murphy 1985 (PP) Future Research Directions *Fedor & Buckley 1988 (PPM) Legal* *Gabor 1989 (US News) Deming* *Hyde 1988 (PPM) Link to pay* *Lowe 1986 (PJ) Link to pay* *Scholtes 1987 (Joiner) Deming* *Slattery 1985 (PJ) Stress*	*VonGlinow 1985 (HRM) Link to pay*	*Friedman 1986 (PJ) Rating Errors* *Metz 1988 (TDJ) Legal* *Scherkenbach 1985 (QP) Deming*

Where: AME = Academy of Management Executive, AMJ = Academy of Management Journal, AMR = Academy of Management Review, F = Fortune, HBR = Harvard Business Review, HRE = Human Resour Executive, HRM = Human Resource Management, JAP = Journal of Applied Psychology, JBR = Journal of Business Research, JVB = Journal of Vocational Behavior, OBHDP = Organizational Behavior and Hun Decision Processes, P = Personnel, PA = Personnel Administrator, PJ = Personnel Journal, PP = Personnel Psychology, PPM = Public Personnel Management, QP = Quality Progress, TDJ = Training and Developi Journal.

Figure 1.

Participation in Performance Appraisal System Design.

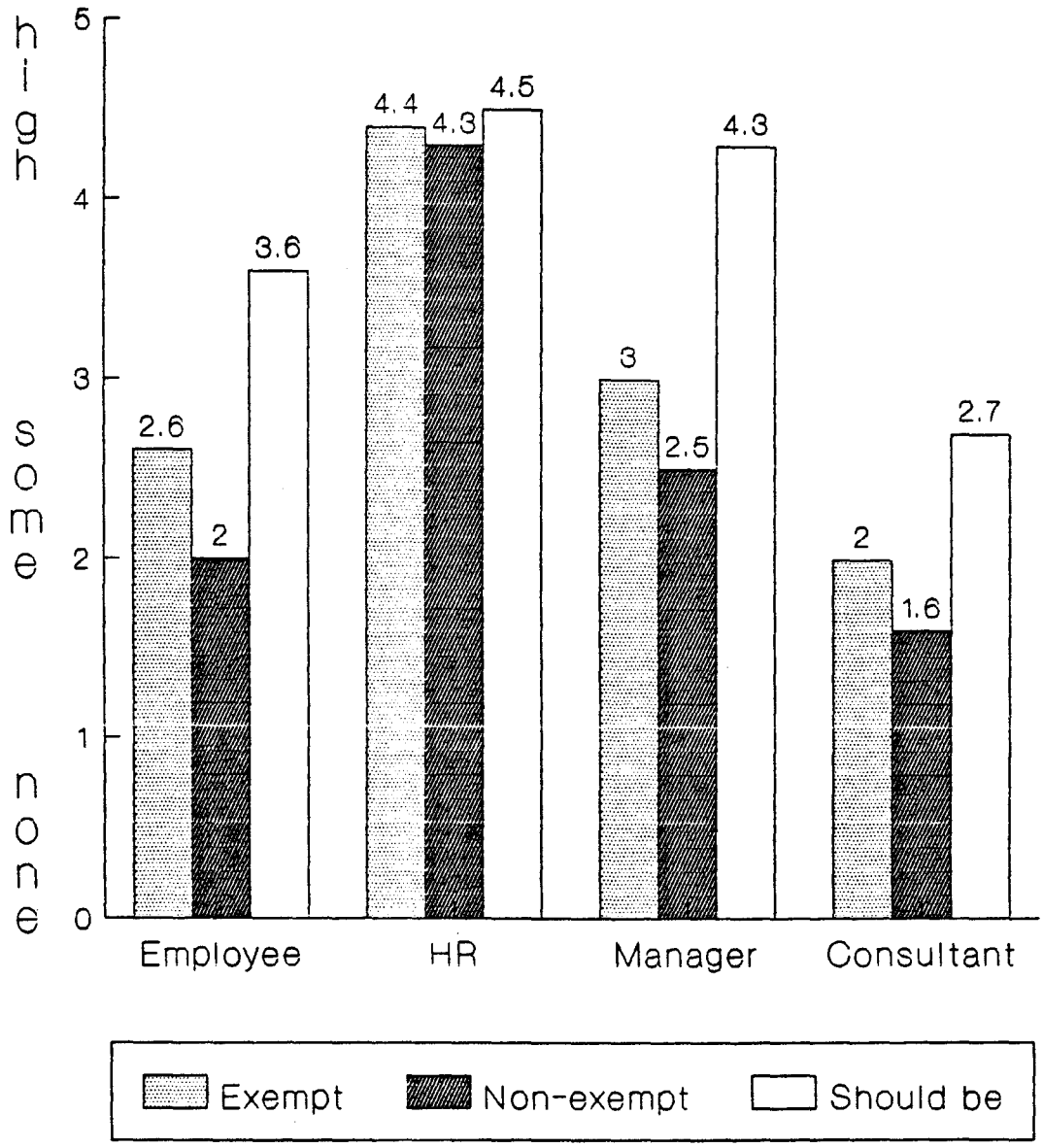


Figure 2.

Performance Appraisal Formats used by F100 Firms.

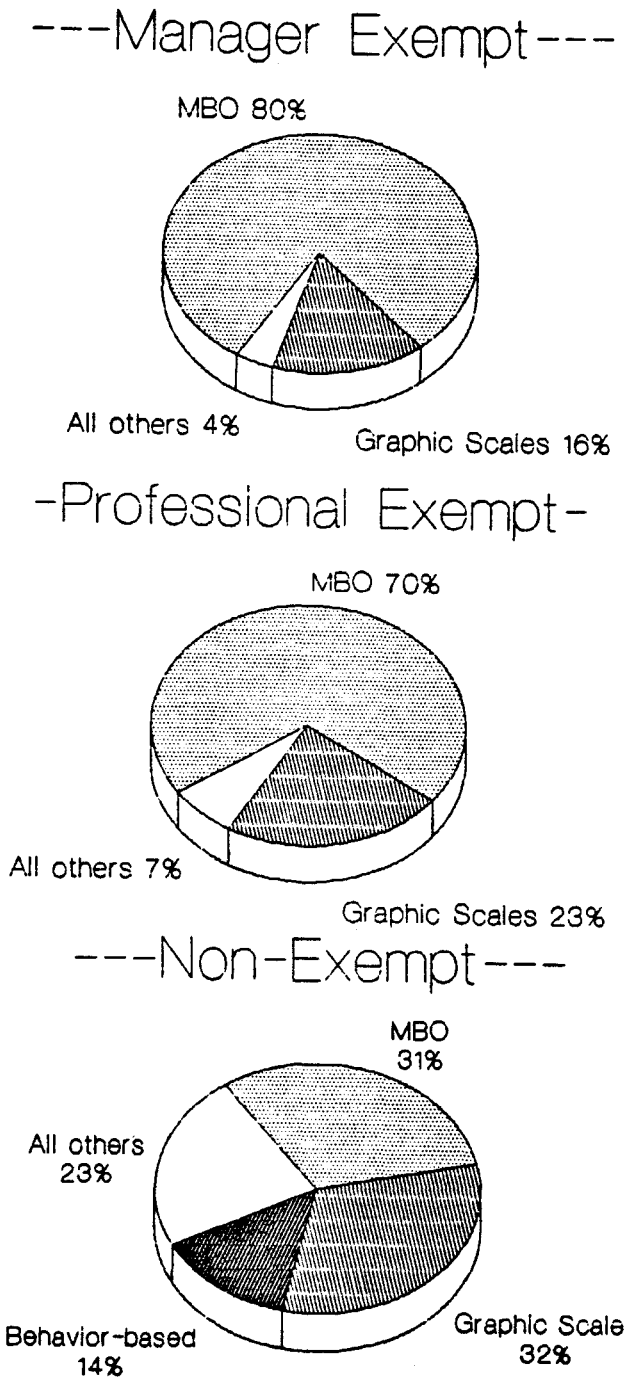


Figure 3.

Importance of various rating sources.

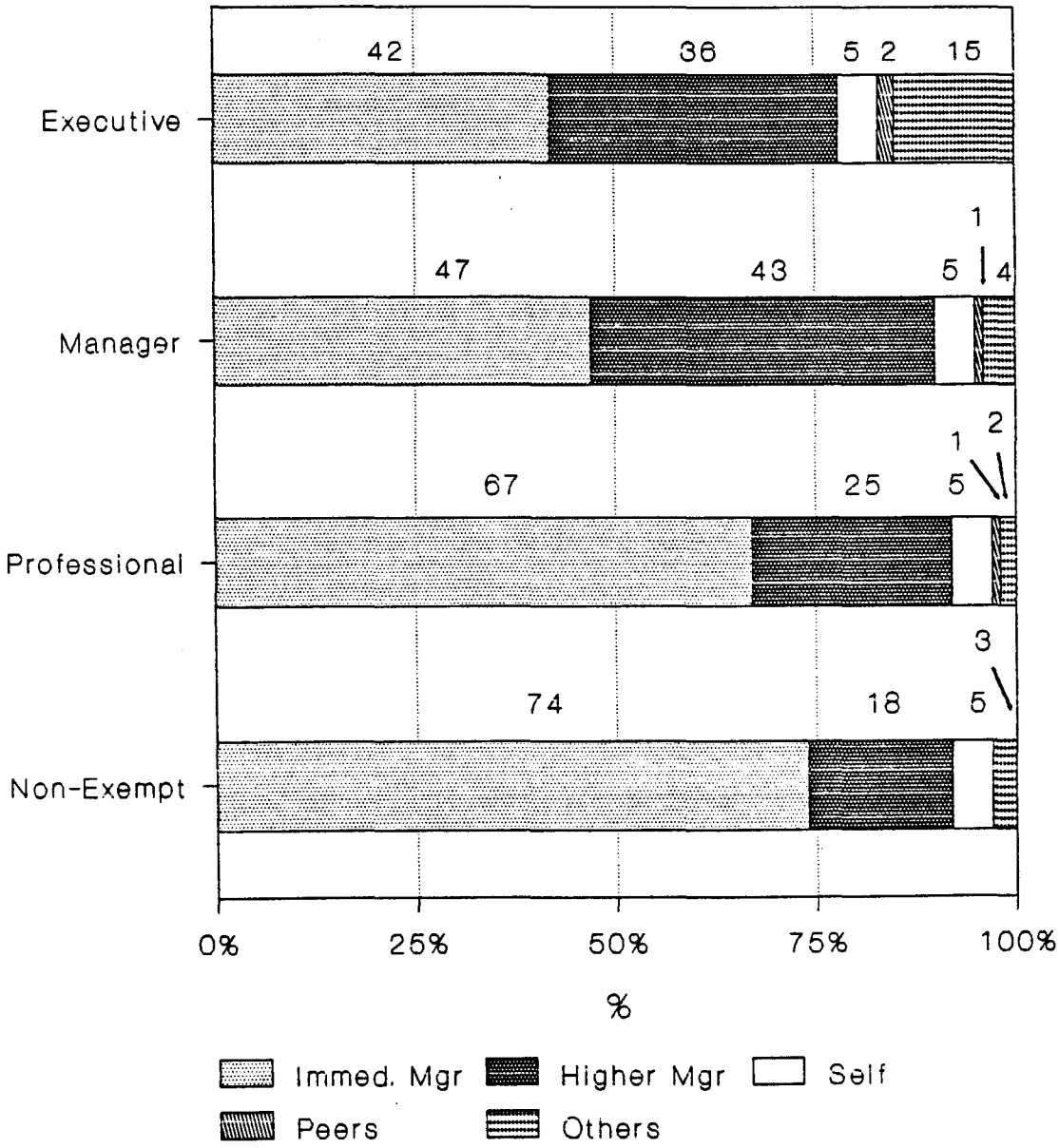
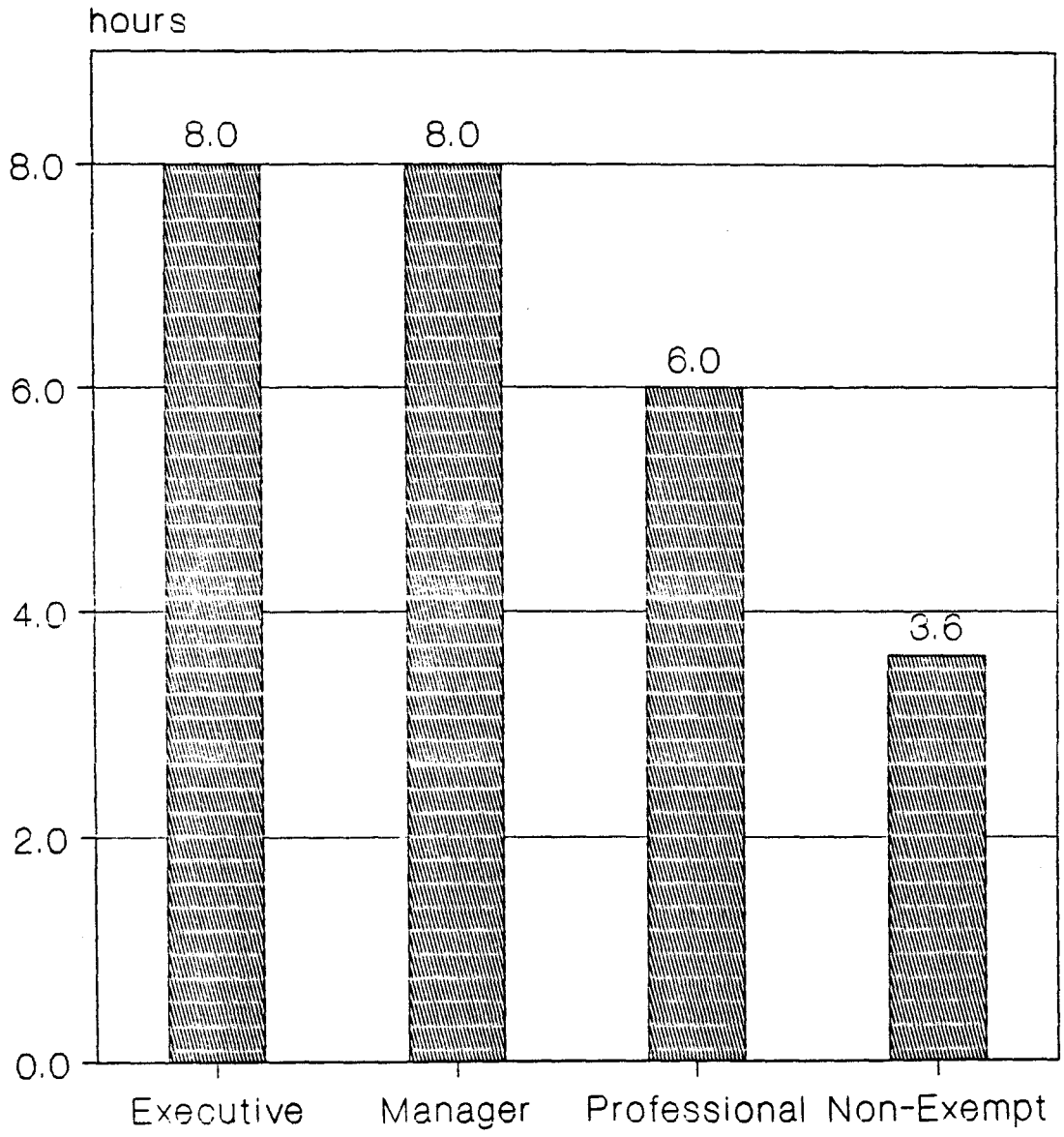


Figure 4.

Time spent on Performance Appraisal per employee per year.



(F=4.58 , p=.001)

Figure 5.

Percent of responding organizations that conduct each type of performance appraisal training.

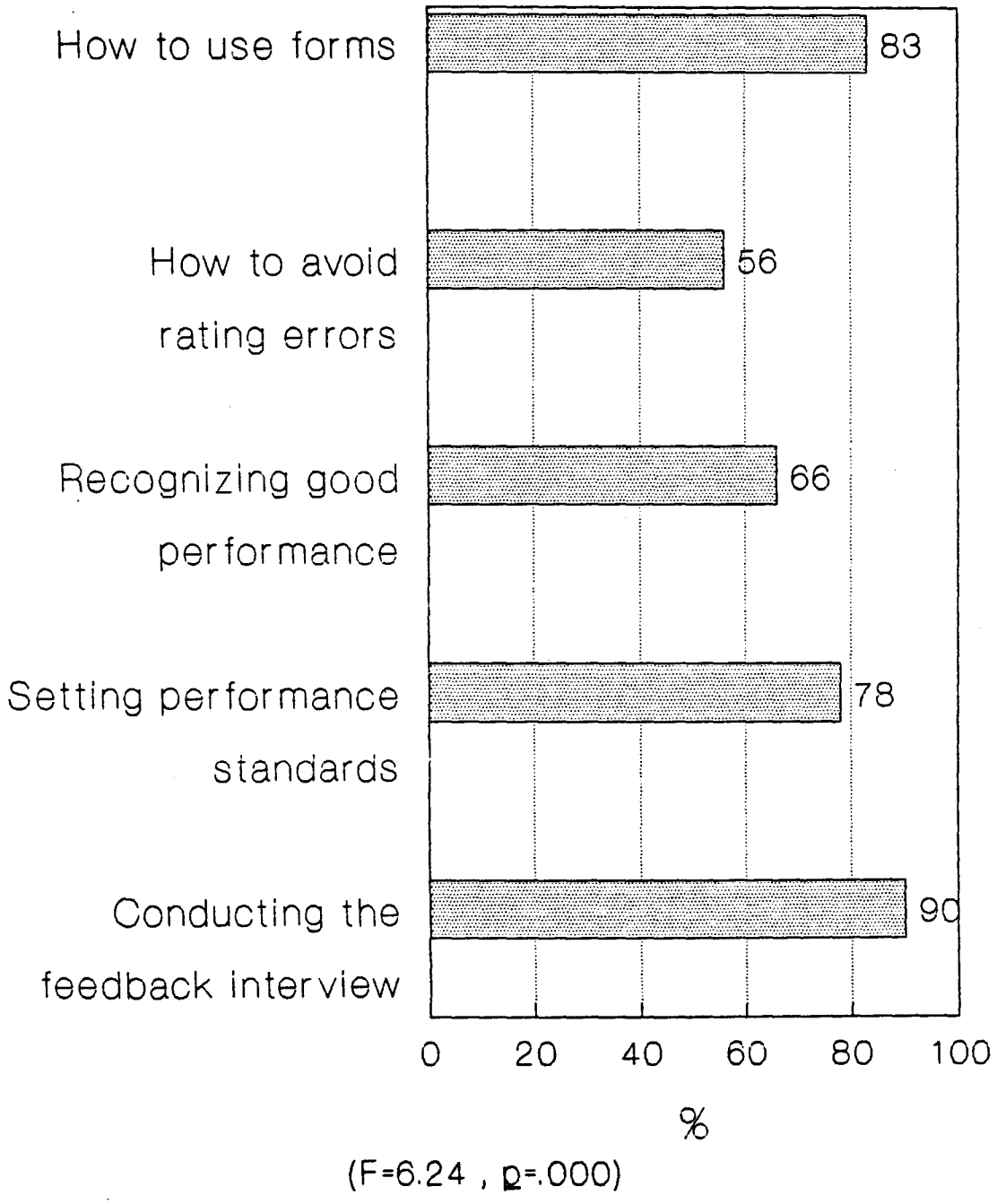


Figure 6.

Are managers rated or held accountable for how they manage the appraisal process of their subordinates?

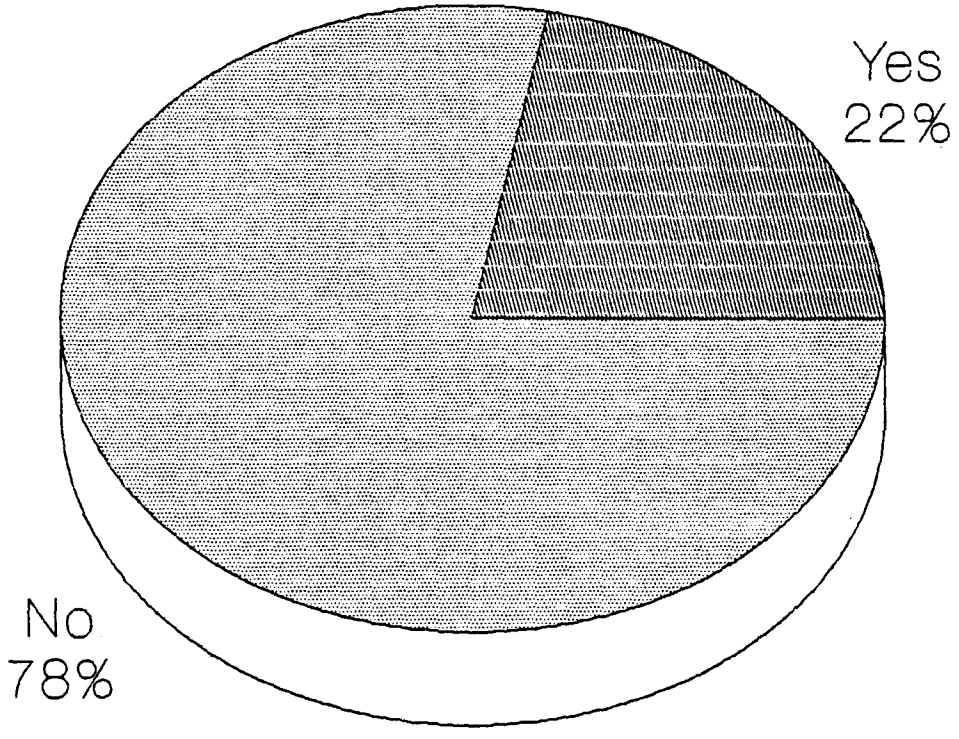


Figure 7.

Can performance appraisal results be appealed?

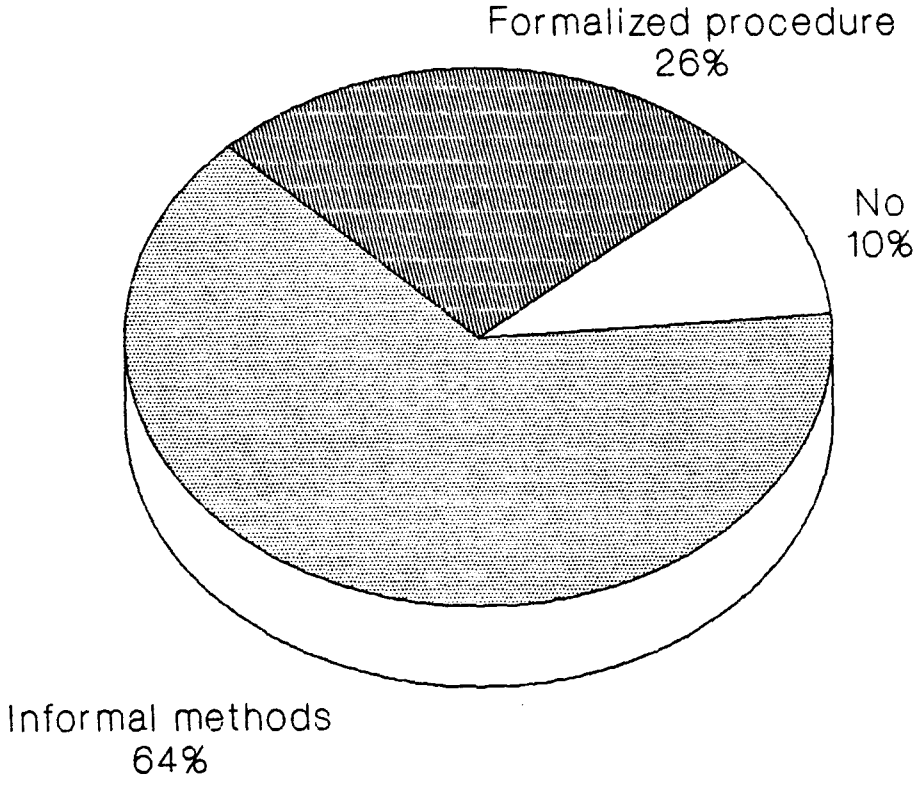


Figure 8.

Percentage of F100 firms that solicit employee reaction via surveys to performance appraisal practices.

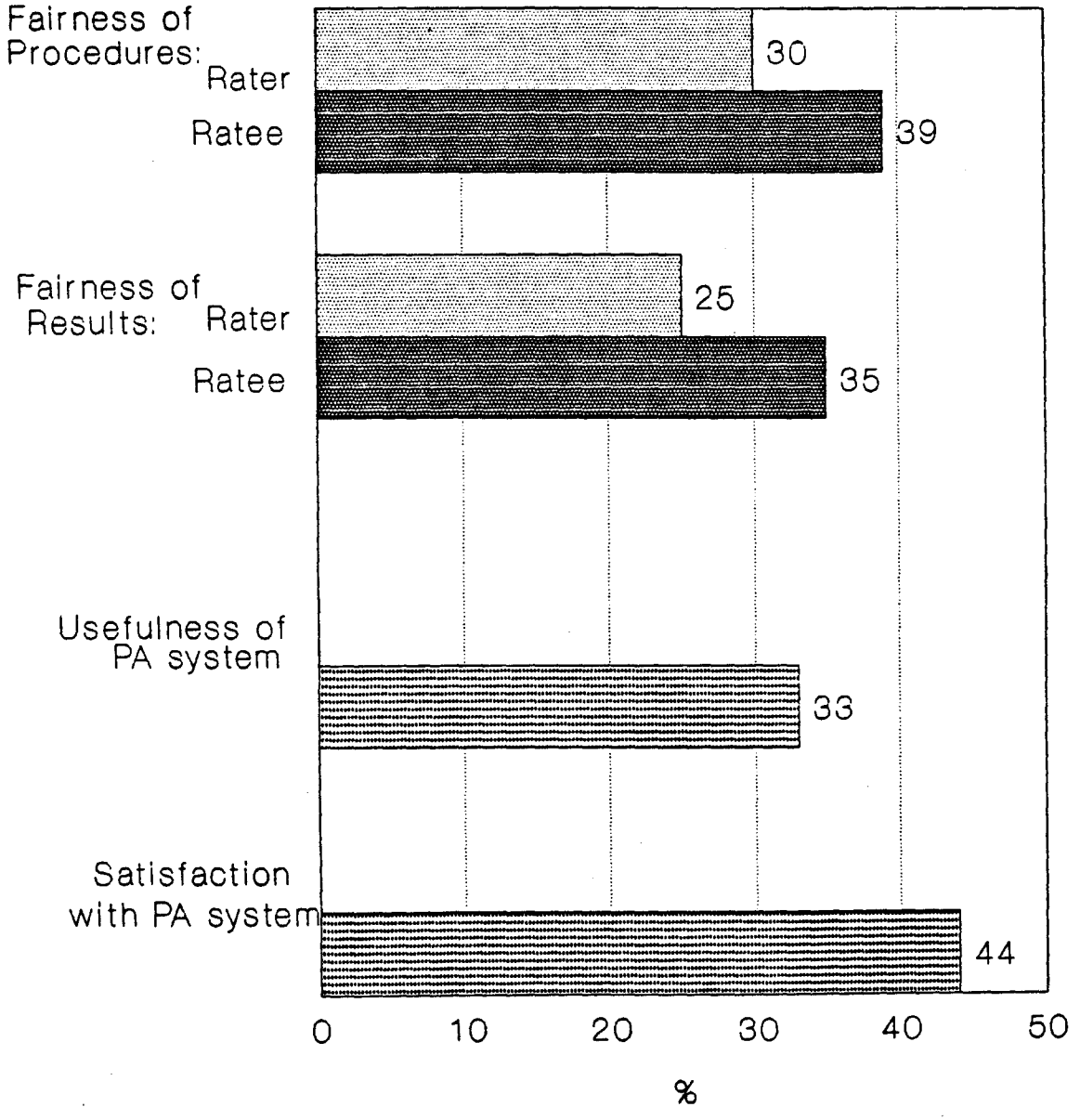
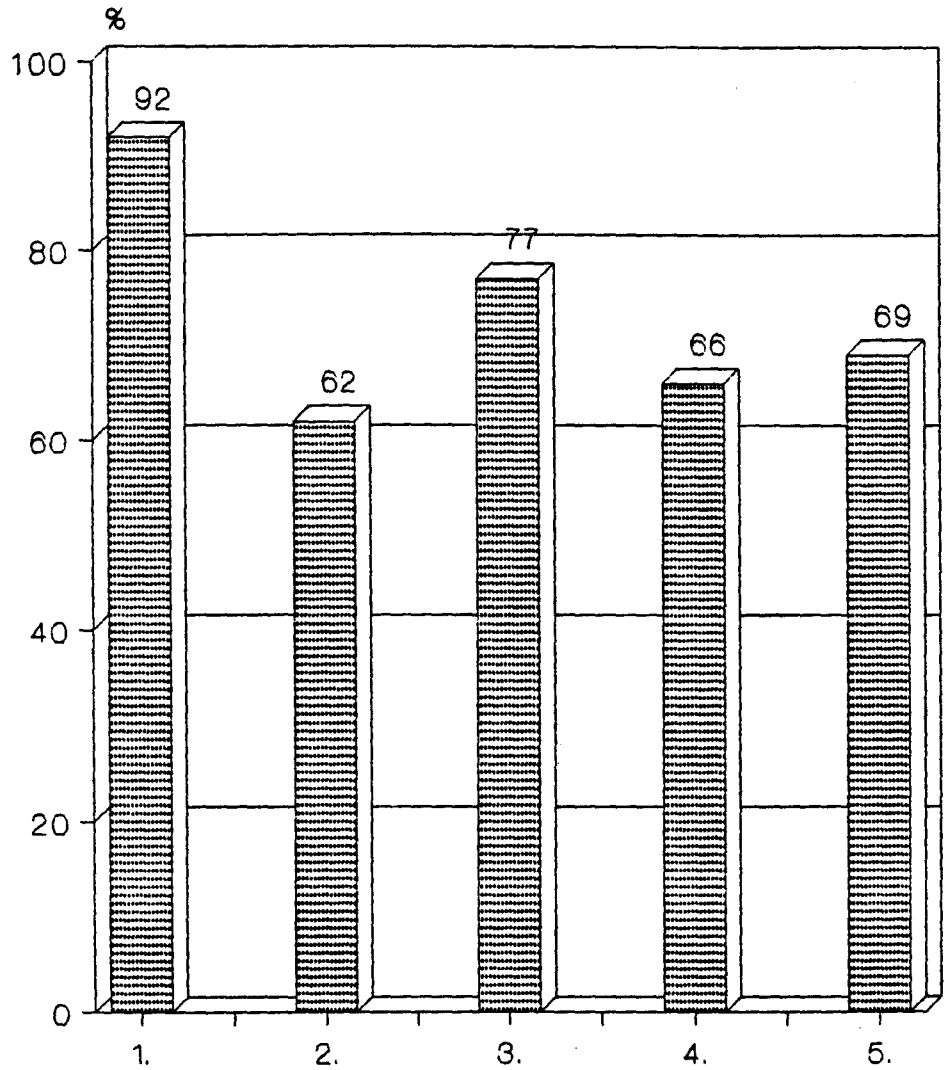


Figure 9.

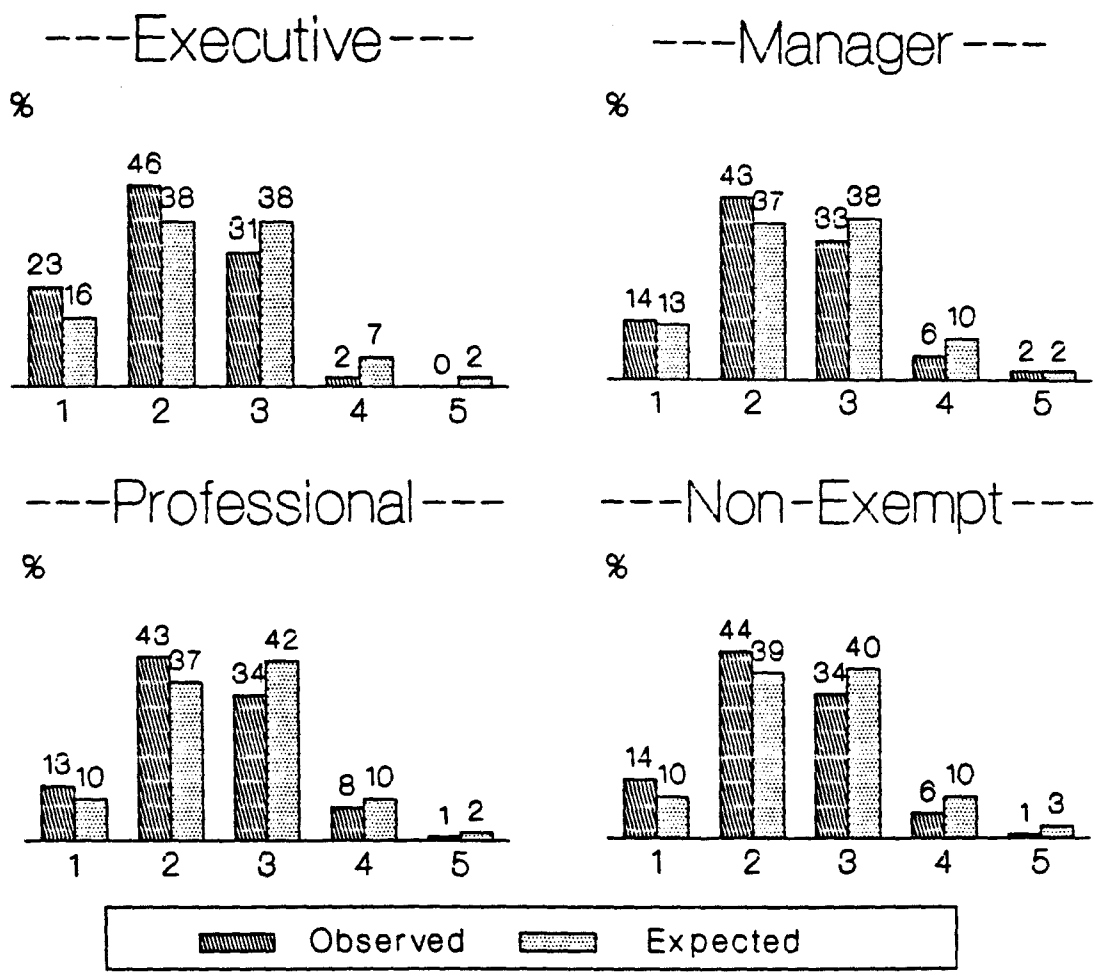
Performance Appraisal and pay delivery.



1. Performance appraisal used to determine merit pay.
2. Merit pay system provides sufficient differentials.
3. Skewed performance distribution is a problem for us.
4. Skew affects pay administration.
5. Skew affects ability to reward best performers.

Figure 10.

Average distribution of observed and expected performance ratings.



1. Far exceeds objectives
2. Exceeds objectives
3. Fully meets objectives
4. Partially meets objectives
5. Unsatisfactory

Figure 11.

Percentage of employees rated in the top two buckets.