

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

**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

This paper has not undergone formal review or approval of the faculty of the ILR School. It is intended to make the results of Center research, conferences and projects available to others interested in human resource management in preliminary form to encourage discussion and suggestions.

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 rater 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 rater 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 rater 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

