

PERFORMANCE
APPRAISAL
RESEARCH AND
PRACTICE

Working Paper 92-15



Center
for
Advanced
Human
Resource
Studies

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

**THE CURRENT STATE OF PERFORMANCE APPRAISAL RESEARCH AND PRACTICE:
CONCERNS, DIRECTIONS, AND IMPLICATIONS**

ROBERT D. BRETZ, JR.

and

GEORGE T. MILKOVICH

Center for Advanced Human Resource Studies
ILR School - Cornell University
393 Ives Hall, Ithaca, NY 14853-3901
(607) 255-5427

WALTER READ

Vice President of Human Resources
Empire Blue Cross and Blue Shield
622 Third Avenue
New York, NY 10017
(212) 476-1979

92 -15

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.

THE CURRENT STATE OF PERFORMANCE APPRAISAL RESEARCH AND PRACTICE: CONCERNS, DIRECTIONS, AND IMPLICATIONS

On the surface, it is not readily apparent how some performance appraisal research issues inform performance appraisal practice. Because performance appraisal is an applied topic, it is useful to periodically consider the current state of performance research and its relation to performance appraisal practice. This review examines the performance appraisal literature published in both academic and practitioner outlets between 1985 and 1990, briefly discusses the current state of performance appraisal practice, highlights the juxtaposition of research and practice, and suggests directions for further research.

The authors gratefully acknowledge support from the Center for Advanced Human Resource Studies, School of Industrial and Labor Relations, Cornell University, and from IBM Corporation. We thank Caroline Weber for her assistance in collecting data about the state of current performance appraisal practice. Finally, we thank the anonymous reviewers for their constructive comments on an earlier draft of this manuscript.

Address all correspondence to Bob Bretz, 393 Ives Hall, Center for Advanced Human Resource Studies, New York State School of Industrial and Labor Relations, Cornell University, Ithaca, New York 14853-3901.

According to some writers, performance appraisal research has done little to improve its usefulness as a managerial decision-making tool (Thorndike, 1949; Banks & Murphy, 1985; Napier & Latham, 1986). Some have suggested that the issues dominating 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 and Murphy (1985) warned that if cognitive process research continued along contemporary lines, the apparent gap between performance appraisal research and practice would increase. Napier and 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 and Villanova (1986) concluded that better understanding of the organizational contexts in which appraisal takes place was necessary in order to improve the degree to which performance appraisal research contributes to performance appraisal practice.

There is a growing concern that much organizational research, while methodologically sophisticated, 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). This statement is representative of some, though certainly not all, of the recent performance appraisal literature. We believe that performance appraisal research can be evaluated both in terms of its theoretical contribution and its ability to inform practice. Since the rating process involves complex cognitive processes, basic research that defines the nature of the phenomenon is clearly needed. However, since performance appraisals occur in applied, social and political contexts, it also is wise to consider the degree to which research is informing practice. Doing so should provide the opportunity for both researchers and managers to critically assess their understanding of the phenomenon. This review is undertaken in that spirit. It (1) summarizes the performance appraisal research published from 1985 through 1990, (2)

provides an overview of the current state of performance appraisal practice in U.S. organizations, (3) compares the trends in performance appraisal research with the issues emerging in practice, and (4) proposes research to address what appear to be the understudied or overlooked issues.

RECENT PERFORMANCE APPRAISAL RESEARCH

Published articles about the performance appraisal process were identified using a computerized literature search (ABI/Inform) augmented by reviewing the tables of contents from several academic and practitioner journals. This literature review is not exhaustive since it does not include technical reports, dissertations, textbooks, or chapters. However, we believe that it serves to indicate, with some precision, the focus of performance appraisal research, and the manner in which appraisal researchers have chosen to allocate their limited resources. Readers interested in reviews covering earlier time periods are referred to Bernardin and Beatty, 1984; Bernardin and Villanova, 1986; DeNisi, Cafferty, and Meglino, 1984; DeNisi and Williams, 1988; Feldman, 1981; Landy and Farr, 1980; and Wexley and Klimoski, 1984. Readers should also note that some of the cognitive processing studies identified below are discussed in some detail in Lord and Maher's (1989) review of the cognitive processing literature.

Overview of Recent Literature

The appendix serves as a guide to performance appraisal research published during this period. The appendix provides an indication of the issues being studied and the methodologies being employed. It should also assist those interested in locating research about particular topics. Recent research has been heavily weighted toward cognitive process issues. The vast majority of these studies were conducted in laboratory settings using student subjects and either paper people or video-tape formats. Rater/ratee characteristics also received considerable attention but the research was not concentrated on any particular characteristic. Study of psychometric issues remained common, with more attention focused

on halo than on other issues. Feedback issues were the one area where field studies were the rule rather than the exception. Research about sources of ratings, rater training, formats, fairness, and appraisal uses and consequences was limited. However, these issues were frequently discussed in the practitioner-oriented outlets, most often in case-study descriptions or in "how to" articles. In the following sections, these major segments of the literature are examined. Although space limitations preclude the discussion of every study, each section attempts to highlight the key issues examined and the methodologies commonly used.

Cognitive Processing of Information

Information processing issues dominated recent performance appraisal research. Laboratory settings and student subjects were very common; only a few studies heeded Banks and Murphy's (1985) call for incorporating non-student subjects and field settings in research designs of cognitive process issues (Hogan, 1987; Huber, Podsakoff & Todor, 1986; Jolly, Reynolds & Slocum, 1988; Mount & Thompson, 1987; Schmitt, Noe, & Gottschalk, 1986). Cognitive processing research concentrated around two issues: (1) how prior expectations or knowledge of prior performance levels affect the way information is processed, and (2) the role of memory in the rating process.

Raters' knowledge of prior performance appears to affect information processing by framing or anchoring current judgements (Huber, Neale, & Northcraft, 1987). Laboratory research indicated that knowledge of prior performance caused contrast effects (i.e., bias away from level of prior performance) rather than assimilation effects (Murphy, Balzer, Lockhart & Eisenman, 1985; Smither, Reilly, & Buda, 1988). Additionally, Steiner and Rain (1989) reported that the order in which good and poor performance was observed affected performance ratings, and that raters biased judgment about inconsistent extreme performance (unusually good or poor) toward the general impression already held.

Raters' expectations also may introduce bias into the rating process. For example, Mount and Thompson (1987) examined the effect of prior expectations on subordinates' ratings of managers whose behaviors were either congruent or incongruent with prior expectations. Results indicated that when behavior was congruent with expectations, appraisal results were more accurate. Similarly, in a longitudinal study of 49 banking supervisor-subordinate diads, Hogan (1987) found supervisors' expectations introduced error into the rating process, and that disconfirmation of prior expectations appeared to lower ratings. However, consistency of ratee performance apparently affects rater ability to form general impressions and categorize information. Padgett and Ilgen (1989) demonstrated that consistent ratee performance led to greater use of categorization while inconsistent performance led to greater retention of behavioral information.

Moreover, it appears that job and ratee knowledge also affect how information is processed. Schmitt, Noe and Gottschalk (1986) studied 153 school administrators to test the degree to which raters used similar methods of combining information, and whether rater agreement was 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 suggested that job and ratee knowledge had significant effects on conceptual similarity and rating covariance, and on halo (Kozlowski, Kirsch & Chao, 1986).

The role of memory has also been important in recent cognitive processing research. Virtually all of this research was conducted in laboratory settings with student subjects. Memory decay introduces bias into the rating process. For example, Kozlowski and Kirsch (1987) suggested that memory decay affected the ability to recall job and ratee information and resulted in halo error and subsequently inaccurate ratings. Under laboratory conditions, ratings recorded one day after performance was observed were already affected by memory

decay (Murphy & Balzer, 1986). When the rater's memory demands are great, bias in favor of general impressions or recent performance may be expected (Murphy, Gannett, Herr & Chen, 1986). Stress has been shown to impact memory by (1) causing less differentiation across dimensions (halo), (2) affecting information retrieval, and (3) possibly affecting categorization as well (Srinivas & Motowidlo, 1987). However, rater characteristics may moderate the degree to which memory decay is problematic. For example, Smither and Reilly (1987) concluded that rater intelligence, not rating delays, affected rating accuracy.

Other cognitive processing studies are very difficult to classify, but illustrate the diversity of this research. For example, information collected from 22 nursing supervisors was used to construct a cognitive map of their appraisal processes (Jolly, Reynolds & Slocum, 1988). Results suggested that values accounted for significant variation in performance ratings. In laboratory settings using student subjects, personality theory (traits) influenced even behaviorally-based ratings (Krzystofiak, Cardy & Newman, 1988), and information acquisition patterns (ranking versus rating) affected how the information was processed (Williams, DeNisi, Megleno & Cafferty, 1986). Sex-role stereotypes did not affect causal attributions of performance and therefore had only small effects on performance ratings (Kinicki & Griffeth, 1985). Williams, DeNisi, Blencoe, and Cafferty (1985) reported that appraisal purpose and outcome had limited effects on how raters utilized information, and Feldman, Camburn, and Gatti (1986) suggested that illusory correlation was not likely to cause bias in ratings. Nathan & Alexander (1985) suggested 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) reported that overemphasis on cognitive determinants of performance ratings has led to neglect of behavioral, social, and affective determinants of bias in the rating process.

Rater/Ratee Personal Characteristics

Research on sex/gender effects has yielded conflicting results. For example, 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), and 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), and raters holding traditional stereotypes of women tended to be less accurate when ratings were made for administrative (versus developmental) purposes. In an experimental setting, Benedict and Levine (1988) demonstrated that females were more lenient with poor performers and delayed performance appraisals and feedback sessions more than males did. However, using both between and within-subject analyses Pulakos, White, Oppler and Borman (1989) concluded that gender and race account for an "extremely small" amount of variance in ratings.

Ratee age received limited research 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 who 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).

A meta analysis of race effects confirmed that rater/ratee similarity results in higher ratings (Kraiger & Ford, 1985). In studies of other personal 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 appeared to influence rating behavior (Tsui & Barry, 1986). However, Cardy & Dobbins (1986) suggested that affect influenced the rating not by increasing leniency but by introducing noise into the process.

Rating Errors and Accuracy

The effect of rating errors on appraisal accuracy continues to attract research resources. Much of this recent research has examined and critiqued competing methods of measuring halo (Pulakos, Schmitt & Ostroff, 1986; Lance & Woehr, 1986; Feldman, 1986). Murphy and Balzer (1986) reported that halo was associated with greater accuracy and speculated that this may be due to categorization schemas that correctly classify the relevant behavioral information and eliminate the noise. Nathan and Tippins (1990) also reported a positive relationship between halo and accuracy, but Fisicaro (1988) concluded that a negative relationship exists. However, Becker and Cardy (1986) argued that the relationship between halo and accuracy was ambiguous, that variance and correlational forms of halo may yield either similar or divergent results. 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), and true halo (true correlations between performance dimensions) was shown to have only a small effect on observed halo (Murphy & Reynolds, 1988).

Based on meta-analytic results, Murphy and Balzer (1989) concluded that the correlation between rating errors and accuracy was very near zero, and therefore, error measures were not good indicators of rating accuracy. Since most performance is multidimensional, some correlation between performance dimensions is expected. Therefore, raters with large observed correlations may, in fact, be accurately rating performance rather than committing halo error. The ambiguity of the relationship may be due to different conceptualizations of accuracy. Lord (1985) proposed that accuracy might

be better understood by distinguishing between classification and behavioral accuracy. Classification accuracy refers to the rater's ability to place ratees into cognitive categories and recall attributes which are most representative of the category prototype. Behavioral accuracy refers to the rater's ability to recall actual ratee behavior and recognize prototypical characteristics that do not necessarily describe the ratee. It seems, therefore, that rating errors such as halo might be positively related to classification accuracy but negatively related to behavioral accuracy.

In other articles on psychometric issues, Smither, Barry, and Reilly (1989) investigated the validity of expert true score estimates, and reported that experts were more accurate than non-experts regardless of the true intercorrelations between performance dimensions. Participation in selection also affected ratings. Specifically, 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 who participated in hiring but viewed the applicant as unfavorable, tended to give more severe ratings. Finally, Sulsky and Balzer (1988) argued that accuracy in performance measurement was lacking due to poor definitions of accuracy, methodological and theoretical limitations of true score development, and the absence of a cohesive theory of performance.

Appraisal Sources

The usefulness of self appraisals may be affected by rating purpose, but conflicting results have been reported. Both laboratory and field studies have concluded that when used for evaluative purposes, self appraisals were susceptible to leniency bias, but leniency decreased when appraisals were expected to be validated (Farh & Werbel, 1986; Farh, Werbel & Bedeian, 1988). Conversely, Fox and Dinur (1988) reported low validity of self ratings regardless of the expectation of validation. Campbell and Lee (1988) suggested that self appraisals were best suited for developmental rather than evaluative purposes, and that self appraisals can improve future performance by creating a self-fulfilling prophecy.

Vance, MacCallum, Coover and Hedge (1988) reported that among a sample of jet engine mechanics, peer, self, and supervisory ratings were equally valid sources but Fox, Ben-Nahum, and Yinon (1989) concluded that rating accuracy was positively related to rater-ratee similarity. Meta-analytic results suggested only moderate relationships exist between self-supervisor and self-peer ratings (Harris & Schaubroeck, 1988).

Appraisal Feedback

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, Stephan & Loveland, 1986). In contrast, Prince and Lawler (1986), reported that salary discussions during the appraisal interview had 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) led to reduced organizational commitment and negative attitudes toward the performance appraisal system.

Using a field study to examine feedback source and message, Earley (1988) reported that self-generated and specific feedback (versus supervisory-generated and general feedback) were positively related to performance. This agrees with Bannister's (1986) experimental results concluding that source credibility and message content influenced recipient response to the feedback. In fact, Becker and Klimoski (1989) reported that feedback from supervisors led to increased performance but feedback from self and peers did not. Ilgen and 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. Message content apparently also affects rater cognition.

Specifically, raters do not like to give negative feedback (Larson, 1989) and are likely to rely on scripts to deliver feedback about poor performance (Dugan, 1989).

The dimensionality of feedback also has been examined. In a longitudinal study of university employees Dorfman et al. (1986) identified three dimensions of performance appraisal feedback (being supportive, emphasizing improvement, and discussing pay and advancement). Furthermore, Russell and Goode (1988) reported that satisfaction with feedback also may be multi-dimensional. Therefore, individuals who are satisfied with the performance appraisal in general, may not be satisfied with the feedback it provides. Rather, satisfaction with feedback may be a function of satisfaction with the supervisor and/or the rating received.

Rater Training

Recent research on rater training has been limited. Hedge and Kavanagh (1988) reported that training focused on minimizing rating errors successfully reduced leniency and halo but also reduced accuracy. They concluded that rater 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 and 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 suggested the training methods best suited for reducing halo, reducing leniency, and improving accuracy (Smith, 1986). Banks and Roberson (1985) argued that rater training programs did not incorporate standard dimensions of test development and the notion of performance appraisals as "tests" violates accepted standards of test construction.

Performance Appraisal Formats

Of the studies examining formats, behaviorally based methods received the most, although limited, attention. Using a sample of mechanics, Hughes and Prien (1986) evaluated alternative scoring methods for mixed standard scales. They found few

differences between the methods and suggested choosing one based on ease of application or explanation. Prien and 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 and Constans concluded that behavioral anchors may lead to biased recall of performance. But two years later, Murphy suggested that the earlier results were not likely to be observed in organizational settings (Murphy & Pardaffy, 1989).

Other Research Issues

Other research-based articles were the sole examples of studies on particular issues. For example, Barrett and Kernan (1987) reviewed court cases since Brito vs. Zia arising from terminations based on performance appraisal, and Miller, Kaspin, and Schuster (1990) discussed performance appraisal practices related to age discrimination cases. Using a managerial sample, Greenberg (1986) reported that perceived fairness of performance evaluations depended 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). Sackett, Zedeck, and Fogli (1988) used a sample of supermarket cashiers to explore 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. Wayne and Ferris (1990) reported that political influence behavior can affect supervisor affectivity and result in more favorable performance ratings than would otherwise be expected, but the type of influence tactic used may affect the performance outcome (Kipnis & Schmidt, 1988). Finally, Napier and Latham (1986) reported that managers (1) perceived no consequences (good or bad) from conducting thorough performance appraisals, and (2) saw little practical value in doing so.

Case Studies and "How-To" Articles

The professional journals are replete with articles discussing the performance appraisal practices in various organizations, and under varying conditions. Case studies and "how to" articles are common. For example, Gellerman and Hodgson (1988) discussed how American Cyanamid Company transformed a ten-level forced distribution performance appraisal system into a three-level system which was deemed to be more consistent with the organization's culture, and Scherkenbach (1985) explained how Ford Motor Company revised their appraisal practice to fit its focus on total quality. Other case studies include reports of the appraisal practices at Xerox (Deets & Tyler, 1986), Control Data (Gomez-Mejia, Page & Tornow, 1985), Merck (Wagel, 1987), many federal and state agencies (e.g. Glen, 1990; Goodell, 1988; Hall, 1987; Laumeier & Beebe, 1988), and several unidentified organizations (Cayer, DiMattia & Wingrove, 1988; Cocheu, 1986; Woods & Dillion, 1985).

The practitioner literature also contains several "recipes" for insuring the effectiveness of various appraisal practices. For example, advice is offered regarding how to construct and implement effective appraisal systems (Levy, 1989; Schneier, Beatty & Baird, 1986a, 1986b), how to make effective use of team appraisals (Edwards & Sproull, 1985; Lanza, 1985), and how to make appraisals more objective (Friedman, 1986; Regel and Hollmann, 1987). One major focus in the practitioner literature addresses transforming performance appraisal from an event to a process. Advice for how to do so is typically discussed under the rubric of "performance management" (Day, 1989; Kirkpatrick, 1986; Levy, 1989; Romanoff, 1989; Schneier, 1989; Schneier, Beatty & Baird, 1986a, 1986b).

Literature Summary

Cognitive processing issues clearly dominated this period 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 field studies but student subjects in laboratory settings 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, assessment of paper people or video scenarios with student subjects in laboratory settings is the norm. The clear exceptions were 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. Some very interesting recent research was found in the "sole example" studies. These include Greenberg's (1986) study of perceived fairness of appraisals as a function of procedural and distributive characteristics, Sackett, Zedeck and Fogli's (1988) use of typical versus maximum job performance and Napier & Latham's (1986) finding that managers perceived 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 is that our knowledge of the rating process has expanded greatly in recent years but remains fragmented. This fragmentation appears to be caused by fundamental differences between the measurement aspects of appraisal research and the organizational purposes of performance appraisal. From a measurement perspective, the necessity to isolate specific effects has resulted in single-issue studies conducted in laboratory settings. Moreover, most of this research addresses the consistency, not necessarily the relevance, of the measurement. The effects of prior expectations, prior knowledge of performance and memory decay have been studied separately from the alternative uses of appraisals (administrative or developmental), the characteristics of raters/ratees or the types of scales and formats employed.

Furthermore, certain appraisal issues have received considerable attention while others have been virtually ignored. The predominance of studies examined information processing, and psychometric issues, yet virtually no systematic research exists on how the organizational context affects the rater/ratee relationship or the cognitive processes of the rater, how raters actually appraise performance, how they use appraisal information, or what issues they believe are important. Moreover, research is only beginning to address how context affects employee perceptions of appraisal, their reactions to appraisal outcomes, and how appraisal purpose (administrative versus developmental) moderates these relationships. Therefore, in order to understand the organizational contexts in which appraisals are made, we turn to a brief examination of the current state of performance appraisal practice.

CURRENT PERFORMANCE APPRAISAL PRACTICES

Organizations vary on a wide array of factors (including for example size, product market, technology, culture, competitive environment, strategy, and union representation) that are likely to affect performance appraisal practices. Therefore, describing the modal setting in which appraisals take place is exceedingly difficult and should be approached cautiously. To reduce reliance on a single survey source, we thought it prudent to integrate the results of three recent surveys of current performance appraisal practices. These included a 1989 survey by the Wyatt Company of the performance appraisal practices in 3052 organizations (Wyatt, 1989), and a 1990 survey by The Conference Board of the performance appraisal practices in 435 of its member organizations (Milkovich & Wigdor, 1991). Additionally, we conducted our own survey to assess the performance appraisal practices of the Fortune Industrial 100. The organizations represented in these three surveys varied in terms of size, location, and industry, and appeared to adequately represent the diversity of U.S. private sector organizations. In all three surveys, the questionnaire was completed by a middle to high-level human resource manager.

For purposes of this review, the survey results are considered along four dimensions; (1) system design and characteristics, (2) system management, (3) important issues and current uses, and (4) performance distributions. Due to space limitations and because the samples differed in size and characteristics, we will present a qualitative integration of their findings. Readers interested in a more detailed presentation of the survey results should see Milkovich and Wigdor (1991) and Wyatt (1989). Details regarding our Fortune 100 survey are available upon request.

System Design and Characteristics

Design. The performance appraisal systems in place in U.S. industry are on average 11 years old. They were designed primarily by personnel specialists with only limited input from the managers who use the system and virtually no input from employees affected by them. In light of the growing interest in employee and customer participation and involvement (Miller & Monge, 1986; Schweiger & Leana, 1986; Wagner & Gooding, 1987), we anticipated that more recently implemented systems might rely on more input from line managers, employees and customers. However, recently implemented systems were no more likely to have them involved in their design than were older systems.

Formats. Management-by-Objectives (MBO) is the preferred format for assessing executives, managers, and professional employees. Other formats, such as trait-based rating scales are far less common among these employee groups. However, "mixed" formats are common, and rating scales or ranking procedures are often used to supplement MBO-based approaches. In contrast, MBO is used to a far lesser extent among non-exempt employees (as defined by the Fair Labor Standards Act). For these employees, trait-based rating scales are the norm. A sizable minority of organizations report using behaviorally-based formats. However, pure behaviorally anchored rating scales (BARS), forced-choice scales, or mixed standard scales are very uncommon. Again, the use of "mixed" formats is common, reflecting the multiple purposes appraisals serve in many organizations. Two employee

groups are far less likely than others to receive formal evaluations. Executives are less likely to be formally evaluated -- perhaps reflecting the difficulty of assessing performance at this level, or the reluctance of executives to submit to the process. Additionally, many hourly employees do not receive formal performance appraisals -- reflecting, in part, organized labor's distrust of the appraisal process and negotiated agreements limiting the use of formal appraisals.

Raters and Sources of Rating Information. The vast majority of performance ratings come directly from the immediate manager. For managerial and professional employees, the second level manager also has significant input. Recent research has demonstrated that non-traditional rating sources, such as self, peers, and subordinates, can provide valid appraisals. Moreover, the popular press has attested to the increasing use of such sources. However, it appears that these types of ratings are still very uncommon, and when they are used the information typically filters through the immediate manager who uses it in making his/her appraisal.

Quantitative indices are used to supply some performance information in most organizations. Profits, sales and costs were frequently cited as important measures for executives and managers. The acquisition and use of job-specific knowledge was considered important for professional positions. Finally, attendance, quality, and quantity of output were important measures for nonexempt and hourly employees.

System Management

Time Spent. It is common to spend about 7 hours per year assessing the performance of each employee at higher organizational levels and about 3 hours per year for each employee at lower levels. However, there is considerable variability on this issue. Many organizations report spending less than 1 hour per appraisal, while a few organizations spend considerably amounts of time. For example, one reported spending between 20 and 40 hours per employee per year on the appraisal process.

Decision Making. Performance appraisal policy decisions (e.g. whether to conduct formal appraisals, whether to link pay to performance, etc.) tend to be made at the corporate level in most organizations, but they are likely to be made at the business unit level in decentralized organizations. Decisions regarding appraisal practices (e.g. type of format to use, rater training issues, etc.), however, are as likely to be made at the business unit level as they are at the corporate level. Very few organizations allow decisions about performance appraisal policies or practice to be made at the facility level.

Training. Most organizations report extensive use of rater training programs. However, training is most likely to occur when new systems are introduced, and few organizations provide rater training on an on-going (yearly) basis. Rater training is most likely to focus on conducting appraisal interviews and providing feedback, proper use of the new forms, setting performance standards, recognizing good performance, and avoiding rating errors. It seems, therefore, that performance appraisal practice has benefited from previous rater training research. However, ratees receive virtually no training in how to best use the process to receive feedback or improve performance. Training remains focused on the rater (manager); preparing employees for their role in the appraisal process simply does not occur.

Rater Accountability. It remains uncommon for managers to be evaluated on how they manage the appraisal process. Basic motivation models suggest 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). In spite of this, only about one-quarter of the organizations surveyed attempted to hold raters accountable for how they managed the appraisal process. Among those that did, the most common method for doing so was to include it as a dimension on the rater's own appraisal form.

Important Performance Appraisal Issues and Uses

Not surprisingly, managers consider fairness and justice issues to be very important. Most organizations report having an informal dispute resolution system (e.g. open door policies) that employees may use to contest the appraisal outcome. About one-quarter report having formalized processes available for this purpose (e.g. binding decisions made by a third party). However, a sizable minority reported that no appeals process was available. While it may be common to have a mechanism for handling appeals, it is far less common to solicit employee opinions about the appraisal process. Most organizations do not systematically collect data to determine either the managers' or the employees' perceptions of fairness of the appraisal process or the results obtained.

Managers identified fairness as the most important performance appraisal issue organizations face. They also tend to be very concerned that the appraisal system be an effective tool to manage future performance, not just one that reflects past performance. Managers indicated that they are most likely to use performance information for improving future performance, making pay distribution decisions, and communicating expectations regarding future performance.

Performance Distributions

Performance appraisal systems typically have five levels to differentiate employee performance. However, even though most organizations report systems with five levels, generally only three levels are used. Both the desired and the actual distributions tend to be top heavy, with the top "buckets" relatively full and the bottom buckets relatively empty. Also, the actual distribution of performance is generally higher than the distribution desired. That is, even though most organizations expect large percentages of employees to receive the top performance ratings and few, if any, to receive the lowest ratings, their expectations tend to be conservative. It is common for 60 to 70 percent of an organization's workforce to be rated in the top two performance levels. This could reflect either actual outstanding

performance on the part of the organization's workforce, or it could be indicative of leniency bias. Since the phenomenon is surprisingly constant across organizations, and it is unlikely that all organizations have predominately outstanding employees, the distributions probably reflect the latter. 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 not only exist, but are common.

THE JUXTAPOSITION OF RESEARCH AND PRACTICE

Cognitive Processing Issues

Research has indicated that expectations of future performance influence the rating process. Since these expectations are formed on the basis of prior knowledge or beliefs, raters will virtually always have some prior performance expectations. It also seems that the type of appraisal process used would affect the degree to which prior expectations became problematic. For example, in MBO-based systems, the mechanism through which a manager and subordinate arrive at mutually agreed upon goals requires that each individual form expectations regarding the level of performance that is achievable. 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 processes used in many organizations. Therefore, performance appraisal practice stands to benefit substantially from this line of research. Additionally, future research should explicitly consider the impact of prior knowledge and expectations under varying conditions of salience created by different appraisal systems.

The conditions under which actual appraisals occur also suggest that continued research on memory characteristics should prove to be valuable. This becomes apparent with the realization that managers report spending only a few hours per year assessing the performance of each employee. This time includes keeping records, completing forms, preparing for the appraisal interview and delivering feedback. 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. Clearly demarcating the limits of memory and recall should lead to processes, such as more frequent appraisals and systematic documentation, that will reduce reliance on memory.

Several other issues emerge from the cognitive processing literature. First, there has been a heavy reliance on student subjects and laboratory settings. While there is some evidence that laboratory settings may provide results that are as valid and generalizable 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, 1986). Laboratory studies are often necessary in order to isolate particular effects. However, sterile environments that dilute the richness and complexity of the environment potentially change the phenomenon of interest. The potential effects of situational and contextual variables must be considered. 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. Therefore, while continued research on isolated cognitive processes is useful, research agendas should be expanded to include attempts to understand how these cognitive processes are affected by the political, social, and affective nature of most rating environments (Dipboye, 1985; Ferris & Judge, 1991).

Rater/Ratee Personal Characteristics

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, race, and ethnicity effects seems particularly timely and important. 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. Finally, the internationalization of the workforce introduces cultural differences regarding job design, performance expectations, and the role of performance feedback. Current research may be culture bound since it assumes a decidedly western approach to these issues. It may be that employees raised in a traditional Japanese environment, for example, may expect a much less directive approach to appraisal and be offended by the confrontational nature of direct feedback. If individual difference characteristics such as these change the way ratings are assigned or interpreted, these trends represent significant issues to be addressed.

A potentially problematic issue that this literature needs to address is the tendency to find 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 experimental settings sufficiently capture the complexity of cross-gender relationships and sex-role stereotypes that exist in work settings.

Psychometric Issues

Researchers and managers appear to have different conceptualizations of accuracy. 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. However, managers tend to 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 true scores.

One potentially fruitful approach to resolving this disparity is Lord's (1985) distinction between classification and behavioral accuracy. Padgett and Ilgen (1989) concluded that behavioral accuracy was both closer to the conceptual criterion and more

descriptive of what accuracy seems to mean in organizational settings. However, since conveying behavioral information is seldom the sole purpose of performance appraisal, Murphy (1991) points out that both types of accuracy have meaning and may be differentially useful depending on the purpose for which the ratings are made (Murphy, Philbin, & Adams, 1989). For example, behavioral accuracy may be more important when providing feedback while classification accuracy may be preferable for salary administration decisions (Murphy, 1991). Therefore, it appears that additional research on the antecedents, definition, and consequences of accuracy are in order.

Appraisal Sources

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? It appears that this will become an even more important issue as the nature of work continues to change. For example, the increasing utilization of self-managed work teams may legitimize peer appraisals and reduce reliance on supervisory ratings. Customer-centered appraisals may also offer potential for understanding performance in specific situations. However, appraisals by customers may be problematic since customers may have limited job knowledge and these ratings are prone to selection bias (since only extreme information is likely to be conveyed).

SUGGESTIONS FOR DIRECTING FUTURE RESEARCH

First, surveys (including our own) that have attempted to delineate the state of practice, appear poorly suited to the task. The problems seem to be caused by respondent characteristics. Most survey methods send a single questionnaire to be completed by a single person within each organization. It is questionable whether a single person knows

enough about the process to adequately convey the nuances of this type of complex human resource system. It is also legitimate to question whether the views expressed by a single organizational respondent are representative of the organizational members for which he/she speaks. Therefore, at the very least, it would seem that multiple perspectives, from multiple organizational levels are necessary to accurately describe the organization's practices.

Furthermore, the respondents to these types of surveys are typically middle or high-level human resource managers that have some kind of policy-making role in the organization. This raises the question of whether their responses are descriptive of the appraisal system as practiced, or as it was intended to be practiced. Given the vested interest the typical respondent has in the appraisal system being described, one might argue that it would be rational to paint as favorable a picture as possible. Therefore, we suggest that future survey research (1) utilize multiple respondents from each participating organization, and (2) clearly distinguish between how the appraisal process was intended to be used and how it is actually used.

The surveys of current practices raise some fundamental research issues. Performance appraisal systems may be considered to be a series of decisions which are affected by environmental, organizational, and dispositional factors. Research is needed to examine the variety of situational variables which affect appraisal design and administrative choices. For example, what are the determinants of managerial choices in performance appraisal design and administration? 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? These types of questions are illustrative of the myriad of decisions that go into implementing any performance appraisal system. Research has tended to focus on the outcomes of these decisions. However, we

seem to know very little about the factors that cause decisions makers to implement certain approaches. Research directed at these types of issues would seem particularly useful for informing future practice.

The surveys revealed some appraisal decisions that varied across employee groups and others that did not. Specifically, while MBO is the most common approach for assessing managers, graphic rating scales are more common among nonexempt employees, and many executives are not subjected to any formal appraisal process. These differences are potentially problematic since the appraisal process appears to become less standardized and systematic at higher organizational levels. It seems that these differences represent potential research opportunities. For example, research could examine whether this practice affects perceptions of procedural and distributive justice both between and within employee groups. We suspect that it does, but research as to the behavioral consequences of these types of practices would prove insightful.

The surveys do not convey a sense of how organizations tie performance appraisal practices to their underlying culture. Differences in this regard may have implications for organizational effectiveness. Are there specific performance appraisal system characteristics that are better suited to particular climates or cultures? An obvious characteristic to consider is the degree to which performance ratings drive pay decisions. Specifically, stronger pay for performance contingencies might be more important under conditions of intense competition than in regulated or cost-plus environments. Also, perhaps administrative uses of performance data should be emphasized in a meritocracy, but developmental uses should be emphasized in public sector/civil servant environments. Finally, different types of organizations certainly reward different types of behavior. Olian and Rynes (1984) speculated, for example, that prospectors and defenders would reward very different types of behavior, but we are not aware of research that has addressed these speculations. Might it also be that particular types of organizations are more willing to set

"stretch goals", forgive performances deficiencies, or establish voice mechanisms? Answers to these types of questions might begin to create a taxonomy of "fit" between performance appraisal system characteristics and organizational environments. Such a taxonomy would certainly allow research to more clearly delineate the most effective appraisal practices for particular organizational settings.

Napier & Latham (1986) found that managers perceived no consequences, positive or negative, of conducting performance appraisals. Conversely, Longenecker, Sims and Gioia (1987) reported that because of actual and perceived negative consequences of accurate appraisal, some managers knowingly make ratings that are inaccurate. However, the rating environment likely affects both rating practices and participant reactions. For example, recent research has shown that contextual variables (such as participation and rating frequency) affect ratee satisfaction (Dobbins, Cardy, & Platz-Vieno, 1990; Giles & Mossholder, 1990). With these exceptions, recent research has not examined the effects of different appraisal system designs or processes on employee attitudes and behaviors.

Some characteristics would appear to influence participant attitudes. For example research also might address whether the use of three, four, or five (or more) rating levels affect employee motivation or work attitudes. Too many levels would supposedly lose the ability to meaningfully differentiate performance. On the other hand, too few levels are likely to create perceptions of inequity. However, what is too many or too few? Consider the proto-typical three-bucket system in which the top and bottom levels are reserved for truly outstanding and truly problematic cases. The middle level is designed to capture 75 to 80 percent of the workforce. If performance is normally distributed, an employee performing one standard deviation below average would likely receive the same rating as an employee performing one standard deviation above average. The better performers would likely consider this to be unfair. What are the perceptions of employees at different positions in the distribution? Are the better performers in the middle bucket more likely to

withdraw than are the poorer ones? Are they more likely to withdraw than are similar performers in say five-level systems in which their performance would be differentially acknowledged?

Deming (1986) argues strongly that performance appraisal has serious negative consequences and urges organizations to cease all individual performance appraisals and to evaluate unit or plant level performance instead (Scholtes, 1987). This approach, reportedly in wide use in high involvement - high commitment facilities, focuses on assisting those whose performance is "out of the system." Deming's notions have received some attention in practitioner performance appraisal literature but no attention from researchers. The possibility of no individual feedback seems difficult to attain. Even without formal individual appraisals, informal appraisals by team leaders and peers seem inevitable and perhaps potentially less systematic and more vulnerable to biases. Accordingly, research might address the effects of informal appraisal processes on employee perceptions, attitudes, and behavior.

The effects of skewed performance distributions on pay allocation and employee attitudes also 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, 1991). Smaller funds may 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. In this regard, tracking changes in appraisal system decisions over time may be useful. For example, some firms report 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.

CONCLUSION

In concluding, we suggest that performance appraisal research and practice seem to converge on many issues and diverge on others. Divergence on some issues is not necessarily a problem since relevancy for decision makers is not the purpose for all research efforts. Yet performance evaluation is an applied subject, and as such research 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 may isolate the effects many researchers wish to investigate, but also limits the generalizability of the results, and removes the issues from the attention and interests of human resource decision makers. If research is to inform practice, interaction between researchers and managers, and application of research results are important.

On the other hand, organizations continue to do things that undermine the effectiveness of the appraisal process. Little time is spent on the appraisal process, raters are not systematically trained and are not held accountable. The employee's role in the performance process is overlooked as are many potentially valuable sources of performance information (self, peers, subordinates). While research has done much to suggest improvements regarding many of the practices noted above, it may ultimately be the changing nature of work that leads managers to implement practices that research has

legitimized. For example, the trend toward self-managed work teams is diminishing the traditional supervisor-subordinate relationship. While on the one hand this may ultimately lead to greater acceptance of peer appraisals, on the other hand it is likely to force research into new directions as well. The cognitive processes involved in peer appraisals are likely to differ from those in supervisory ratings due to the differences in power and social interactions resulting from peer relationships.

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. Cognitive processing research attempts to understand 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, managerial concerns for fairness are being addressed by cognitive processing research. Nevertheless, some very important issues raised by managers are receiving little or no research attention. Most important of these is the need for a more explicit focus on procedural and distributive justice. 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 performance ratings, are judged to be unfair (Greenberg, 1988, 1990).

Therefore, it appears that current performance appraisal research could be expanded to include these concerns. While the current focus is productive and necessary, an expanded research agenda that included the issues discussed above would further allow performance appraisal research to influence human behavior in organizations.

REFERENCES

- Alexander, E. R., Helms, M. M., & Wilkins, R. D. 1989. The relationship between supervisory communication and subordinate performance and satisfaction among professionals. Public Personnel Management, 18: 415-429.
- Angel, N. F. 1989. Evaluating employees by computer. Personnel Administrator. November: 67-72.
- Athey, T. R., & McIntyre, R. M. 1987. Effect of rater training on rater accuracy: Levels-of-processing theory and social facilitation theory perspectives. Journal of Applied Psychology, 72: 567-572.
- Balzer, W. K. 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, C. G., & Murphy, K. R. 1985. Toward narrowing the research-practice gap in performance appraisal. Personnel Psychology, 38: 335-345.
- Banks, C. G., & Roberson, L. 1985. Performance appraisers as test developers. Academy of Management Review, 10: 128-142.
- Bannister, B. D. 1986. Performance outcome feedback and attributional feedback: Interactive effects on recipient responses. Journal of Applied Psychology, 71: 203-210.
- Barrett, G. V., & Kernan, M. C. 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, B. E., & Cardy, R. L. 1986. Influence of halo error on appraisal effectiveness: A conceptual and empirical reconsideration. Journal of Applied Psychology, 71: 662-671.

- Becker, B. E., & Klimoski, R. J. 1989. A field study of the relationship between the organizational feedback environment and performance. Personnel Psychology, 42: 343-358.
- Bedeian, A. G. 1989. Totems and taboos: Undercurrents in the management discipline. The Academy of Management News, 19 (4): 2-6.
- Benedict, M. E., & Levine, E. L. 1988. Delay and distortion: Tacit influences on performance appraisal effectiveness. Journal of Applied Psychology, 73: 507-514.
- Bernardin, H. J. 1986. Subordinate appraisal: A valuable source of information about managers. Human Resource Management, 25: 421-439.
- Bernardin, H. J., & Abbott, J. 1985. Predicting (and preventing) differences between self and supervisory appraisals. Personnel Administrator. June: 151-157.
- Bernardin, H. J., & Beatty, R. W. 1984. Performance appraisal: Assessing human behavior at work. Boston: Kent.
- Bernardin, H. J., & Beatty, R. W. 1987. Can subordinate appraisals enhance managerial productivity? Sloan Management Review, 28: Summer, 63-73.
- Bernardin, H. J., & Villanova, P. 1986. Performance appraisal. In E. A. Locke (Ed.), Generalizing from laboratory to field settings: 43-62. Lexington, MA: Lexington.
- Brannick, M. T., & Brannick, J. P. 1989. Nonlinear and noncompensatory processes in performance evaluation. Organizational Behavior and Human Decision Processes, 44: 97-122.
- Brown, D. G. 1987. Development of performance standards: A practical guide. Public Personnel Management, 16: 93-114.
- Brumback, B. G. 1988. Some ideas, issues and predictions about performance management. Public Personnel Management, 17: 387-402.
- Buford, J. A., Burkhalter, B. B., & Jacobs, G. T. 1988. Link job descriptions to performance appraisals. Personnel Journal. June: 132-140.

- Campbell, D. J., & Lee, C. 1988. Self-appraisal in performance evaluation: Development versus evaluation. Academy of Management Journal, 13: 302-314.
- Cardy, R. L., & Dobbins, G. H. 1986. Affect and appraisal accuracy: Liking as an integral dimension in evaluating performance. Journal of Applied Psychology, 71: 672-678.
- Cayer, M., DiMattia, D. J., & Wingrove, J. 1988. Conquering evaluation fear. Personnel Administrator. June: 97-107.
- Clement, R. W., & Stevens, G. E. 1989. Performance appraisal in higher education: Comparing departments of management with other business units. Public Personnel Management, 18: 263-278.
- Cleveland, J. N., Murphy, K. R., & Williams, R. E. 1989. Multiple uses of performance appraisal: Prevalence and correlate. Journal of Applied Psychology, 74: 130-135.
- Cocheu, T. 1986. Performance appraisal: A case in points. Personnel Journal. September: 48-55.
- Cozzetto, D. 1990. The officer fitness report as a performance appraisal tool. Public Personnel Management, 19: 235-244.
- Czajka, J. M., & DeNisi, A. S. 1988. Effects of emotional disability and clear performance standards on performance ratings. Academy of Management Journal, 31: 394-404.
- Daley, D. M. 1990. The Civil Service Reform Act and performance appraisal: A research note on federal employee perceptions. Public Personnel Management, 19: 245-251.
- Day, D. 1989. Performance management year-round. Personnel. August: 43-45.
- Dearden, J. 1987. Measuring profit center managers. Harvard Business Review. September-October: 84-88.
- Deets, N. R., & Tyler, D. T. 1986. How Xerox improved its performance appraisals. Personnel Journal. April: 50-52.
- DeGregorio, M., & Fisher, C. D. 1988. Providing performance feedback: Reactions to alternate methods. Journal of Management, 14: 605-615.

- Deming, E. 1986. Out of the crisis. Cambirdge, MA: Massachusetts Institute of Technology Center for Advanced Engineering.
- DeNisi, A. S., Cafferty, T. P., & Meglino, B. M. 1984. A cognitive model of the performance appraisal process: A model and research propositions. Organizational Behavior and Human Performance, 33: 360-396.
- DeNisi, A. S., Robbins, T., & Cafferty, T. P. 1989. Organization of information used for performance appraisals: Role of diary-keeping. Journal of Applied Psychology, 74: 124-129.
- DeNisi, A. S., & William, K. J. 1988. Cognitive approaches to performance appraisal. In K. M. Rowland & G. R. Ferris (Eds.), Research in personnel and human resources management, Vol.6: 109-155. Greenwich, CT: JAI Press.
- Derven, M. G. 1990. The paradox of performance appraisals. Personnel Journal. February: 107-111.
- Dickinson, T. L. 1987. Designs for evaluating the validity and accuracy of performance ratings. Organizational Behavior and Human Decision Processes, 40: 1-21.
- Dipboye, R. L. 1985. Some neglected variables in research on discrimination in appraisals. Academy of Management Review, 10: 116-127.
- Dobbins, G. H., Cardy, R. L., & Platz-Vieno, S. J. 1990. A contingency approach to appraisal satisfaction: An initial investigation of the joint effects of organizational variables and appraisal characteristics. Journal of Management, 16: 619-632.
- Dobbins, G. H., Cardy, R. L., & Truxillo, D. M. 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: 551-558.

- Dorfman, P. W., Stephan, W. G., & Loveland, J. 1986. Performance appraisal behaviors: Supervisor perceptions and subordinate reactions. Personnel Psychology, 39: 579-597.
- Drazin, R., & Auster, E. R. 1987. Wage differences between men and women: Performance appraisal ratings vs. salary allocation as the locus of bias. Human Resource Management, 26: 157-168.
- Dugan, B. 1988. Effects of assessor training on information use. Journal of Applied Psychology, 73: 743-748.
- Dugan, K. W. 1989. Ability and effort attributions: Do they affect how managers communicate performance feedback information? Academy of Management Journal, 32: 87-114.
- Earley, P. C. 1988. Computer-generated performance feedback in the magazine-subscription industry. Organizational Behavior and Human Decision Processes, 41: 50-64.
- Eder, R. W., & Fedor, D. B. 1989. Priming performance self-evaluations: Moderating effects of rating purpose and judgment confidence. Organizational Behavior and Human Decision Processes, 44: 474-493.
- Edwards, M. R. 1990a. A joint effort leads to accurate appraisals. Personnel Journal. June: 122-128.
- Edwards, M. R. 1990b. Implementation strategies for multiple rater systems. Personnel Journal. September: 130-138.
- Edwards, M. R., & Sproull, J. R. 1985. Making performance appraisals perform: The use of team evaluation. Personnel. March: 28-32.
- Eyres, P. S. 1989. Legally defensible performance appraisal systems. Personnel Journal. July: 58-62.

- Farh, J. L., & Dobbins, G. H. 1989. Effects of comparative performance information on the accuracy of self-ratings and agreement between self- and supervisor ratings. Journal of Applied Psychology, 74: 606-610.
- Farh, J. L., & Werbel, J. D. 1986. Effects of purpose of the appraisal and expectation of validation on self-appraisal leniency. Journal of Applied Psychology, 71: 527-529.
- Farh, J. L., Werbel, J. D., & Bedeian, A. G. 1988. An empirical investigation of self-appraisal-based performance evaluation. Personnel Psychology, 41: 141-156.
- Fedor, D. B., & Buckley, M. R. 1988. Issues surrounding the need for more frequent monitoring of individual performance in organizations. Public Personnel Management, 17: 435-442.
- Fedor, D. B., Eder, R. W., & Buckley, M. R. 1989. The contributory effects of supervisor intentions on subordinate feedback responses. Organizational Behavior and Human Decision Processes, 44: 396-414.
- Fedor, D. B., & Rowland, K. M. 1989. Investigating supervisor attributions of subordinate performance. Journal of Management, 15: 405-416.
- Feldman, J. M. 1981. Beyond attribution theory: Cognitive processes in performance appraisal. Journal of Applied Psychology, 66: 127-148.
- Feldman, J. M. 1986. A note on the statistical correction of halo error. Journal of Applied Psychology, 71: 173-176.
- Feldman, J. M., Camburn, A., & Gatti, G. M. 1986. Shared distinctiveness as a source of illusory correlation in performance appraisal. Organizational Behavior and Human Decision Processes, 37: 34-59.
- Ferris, G. R., & Judge, T. A. 1991. Personnel/human resources management: A political influence perspective. Journal of Management, 17: 447-488.

- Ferris, G. R., Yates, V. L., Gilmore, D. C., & Rowland, K. M. 1985. The influence of subordinate age on performance ratings and causal attributions. Personnel Psychology, 38: 545-557.
- Fisicaro, S. A. 1988. A reexamination of the relation between halo error and accuracy. Journal of Applied Psychology, 73: 239-244.
- Fox, S., Ben-Nahum, Z., & Yinon, Y. 1989. Perceived similarity and accuracy of peer ratings. Journal of Applied Psychology, 74: 781-786.
- Fox, S., & Dinur, Y. 1988. Validity of self-assessment: A field evaluation. Personnel Psychology, 41: 581-592.
- Friedman, M. G. 1986. 10 steps to objective appraisals. Personnel Journal. June: 66-71.
- Fulk, J., Brief, A. P., & Barr, S. H. 1985. Trust-in-supervisor and perceived fairness and accuracy of performance evaluations. Journal of Business Research, 13: 299-313.
- Galin, A., & Benoliel, B. 1990. Does the way you dress affect your performance rating? Personnel. August: 49-52.
- Gellerman, S. W., & Hodgson, W. G. 1988. Cyanamid's new take on performance appraisal. Harvard Business Review. May-June: 36-41.
- George, D. I., & Smith, M. C. 1990. An empirical comparison of self-assessment and organizational assessment in personnel selection. Public Personnel Management, 19: 175-190.
- Gibb, P. 1985. Appraisal goals and controls. Personnel Journal. August: 89-93.
- Giffin, M. E. 1989. Personnel research on testing, selection, and performance appraisal. Public Personnel Management, 18: 127-137.
- Gilbert, G. R., & Nelson, A. E. 1989. The Pacer Share demonstration project: Implications for organizational management and performance evaluation. Public Personnel Management, 18: 209-225.

- Giles, W. F., & Mossholder, K. W. 1990. Employee reactions to contextual and session components of performance appraisal. Journal of Applied Psychology, 75: 371-377.
- Girard, R. 1988. Is there a need for performance appraisals? Personnel Journal. August: 89-90.
- Glen, R. M. 1990. Performance appraisal: An unnerving yet useful process. Public Personnel Management, 19: 1-10.
- Goddard, R. W. 1989. Is your appraisal system headed for court? Personnel Journal. January: 114-118.
- Gomez-Mejia, L. R., Page, R. C., & Tornow, W. W. 1985. Improving the effectiveness of performance appraisal. Personnel Administrator. January: 74-81.
- Goodell, R. 1988. Room for improvement. Personnel Administrator. June: 132-142.
- 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: Managing impressions of organizational justice. In B. M. Staw & L. L. Cummings (Eds.), Research in organizational behavior, Vol. 12: 111-157. Greenwich, CT: JAI Press.
- Greenhaus, J. H., Parasuraman, S., & Wormley, W. M. 1990. Effects of race on organizational experiences, job performance evaluations, and career outcomes. Academy of Management Journal, 33: 64-86.
- Hall, T. C. 1987. Starting over. Training and Development Journal. December: 60-62.
- Harper, S. C. 1986. Adding purpose to performance reviews. Training and Development Journal. September: 53-55.

- Harris, C. 1988. A comparison of employee attitudes toward two performance appraisal systems. Public Personnel Management, 17: 443-456.
- Harris, M. M., & Schaubroeck, J. 1988. A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. Personnel Psychology, 41: 43-62.
- Hedge, J. W., & Kavanagh, M. J. 1988. Improving the accuracy of performance evaluations: Comparison of three methods of performance appraiser training. Journal of Applied Psychology, 73: 68-73.
- Heilman, M. E., & Stopeck, M. H. 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, R. L. 1986. The relationship between supervisory ratings and results-oriented measures of performance: A meta-analysis. Personnel Psychology, 39: 811-826.
- Heneman, R. L., Greenberger, D. B., & Anonyuo, C. 1989. Attributions and exchanges: The effects of interpersonal factors on the diagnosis of employee performance. Academy of Management Journal, 32: 466-476.
- Heneman, R. L., Wexley, K. N., & Moore, M. L. 1987. Performance-rating accuracy: A critical review. Journal of Business Research, 15: 431-448.
- Herbert, G. R., & Doverspike, D. 1990. Performance appraisal in the training needs analysis process: A review and critique. Public Personnel Management, 19: 253-270.
- Hogan, E. A. 1987. Effects of prior expectations on performance ratings: A longitudinal study. Academy of Management Journal, 30: 354-368.
- Howard, C. G. 1987. Out of sight - not out of mind. Personnel Administrator. June: 82-90.

- Huber, V. L., Neale, M. A., & Northcraft, G. B. 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, V. L., Podsakoff, P. M., & Todor, W. D. 1986. An investigation of biasing factors in the attributions of subordinates and their supervisors. Journal of Business Research, 14: 83-98.
- Hughes, G. L., & Prien, E. P. 1986. An evaluation of alternate scoring methods for the mixed standard scale. Personnel Psychology, 39: 839-847.
- Hyde, A. C. 1988. The new environment for compensation and performance evaluation in the public sector. Public Personnel Management, 17: 351-358.
- Ilggen, D. R., & Moore, C. F. 1987. Types and choices of performance feedback. Journal of Applied Psychology, 72: 401-406.
- Jacobs, R., & Kozlowski, S. W. 1985. A closer look at halo error in performance ratings. Academy of Management Journal, 28: 201-212.
- Jolly, J. P., Reynolds, T. J., & Slocum, J. W. 1988. Application of the means-end theoretic for understanding the cognitive bases of performance appraisal. Organizational Behavior and Human Decision Processes, 41: 153-179.
- Kamouri, A. L., & Balzer, W. K. 1990. The effects of performance sampling methods on frequency estimation, probability estimation, and evaluation of performance information. Organizational Behavior and Human Decision Processes, 45: 285-316.
- Kane, J. S., & Freeman, K. A. 1986. MBO and performance appraisal: A mixture that's not a solution, part 1. Personnel. December: 26-36.
- Kane, J. S., & Freeman, K. A. 1987. MBO and performance appraisal: A mixture that's not a solution, part 2. Personnel. February: 26-32.
- Karl, K. A., & Wexley, K. N. 1989. Patterns of performance and rating frequency: Influence on the assessment of performance. Journal of Management, 15: 5-20.

- Kaufman, R. 1988. Preparing useful performance indicators. Training and Development Journal. September: 80-83.
- Kerr, J. L. 1988. Strategic control through performance appraisal and rewards. Human Resource Planning, 2: 215-223.
- Kingstrom, P. O., & Mainstone, L. E. 1985. An investigation of the rater-ratee acquaintance and rater bias. Academy of Management Journal, 28: 641-653.
- Kinicki, A. J., & Griffeth, R. W. 1985. The impact of sex-role stereotypes on performance ratings and causal attributions of performance. Journal of Vocational Behavior, 27: 155-170.
- Kipnis, D., & Schmidt, S. M. 1988. Upward influence styles: Relationship with performance evaluations, salary, and stress. Administrative Science Quarterly, 33: 528-542.
- Kirkpatrick, D. L. 1986. Performance appraisal: Your questions answered. Training and Development Journal. May: 68-71.
- Klaas, B. S., & DeNisi, A. S. 1989. Managerial reactions to employee dissent: The impact of grievance activity on performance ratings. Academy of Management Journal, 32: 705-717.
- Klimoski, R., & Inks, L. Accountability forces in performance appraisal. Organizational Behavior and Human Decision Processes, 45: 194-208.
- Kozlowski, S. W., & Kirsch, M. P. 1987. The systematic distortion hypothesis, halo, and accuracy: An individual-level analysis. Journal of Applied Psychology, 72: 252-261.
- Kozlowski, S. W., Kirsch, M. P., & Chao, G. T. 1986. Job knowledge, ratee familiarity, conceptual similarity and halo error: An exploration. Journal of Applied Psychology, 71: 45-49.

- Kraiger, K., & Ford, J. K. 1985. A meta-analysis of ratee race effects in performance ratings. Journal of Applied Psychology, 70: 56-65.
- Krein, T. J. 1990. Performance reviews that rate an "A". Personnel. May: 38-40.
- Kroll, M., & Joyce, G. 1989. What is your most important job function? Personnel Administrator. June: 156-160.
- 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: 515-521.
- Lance, C. E., & Woehr, D. J. 1986. Statistical control of halo: Clarification from two cognitive models of the performance appraisal process. Journal of Applied Psychology, 71: 679-685.
- Landy, F. S., & Farr, J. L. 1980. Performance rating. Psychological Bulletin, 87: 72-107.
- Landy, F. S., & Farr, J. L. 1983. The measurement of work performance. Methods, theory, and applications. Orlando, Florida: Academic Press.
- Lanza, P. 1985. Team appraisals. Personnel Journal. March: 47-51.
- Larson, Jr., J. R. 1989. The dynamic interplay between employees' feedback-seeking strategies and supervisors' delivery of performance feedback. Academy of Management Review, 14: 408-422
- Laumeyer, J. & Beebe, T. 1988. Employees and their appraisals. Personnel Administrator. December: 76-80.
- Lawrence, B. S. 1988. New wrinkles in the theory of age: Demography, norms, and performance ratings. Academy of Management Journal, 31: 309-337.
- Lawrie, J. 1990. Prepare for a performance appraisal. Personnel Journal. April: 132-136.
- Lawrie, J. W. 1989. Your performance: Appraise it yourself! Personnel. January: 21-23.
- Lee, C. 1989. Poor performance appraisals do more harm than good. Personnel Journal. September: 91-99.

- Lee, C. 1990. Smoothing out appraisal systems. HR Magazine. March: 72-76.
- Levy, M. 1989. Almost-perfect performance appraisals. Personnel Journal. April: 76-83.
- Locher, A. H., & Teel, K. S. 1988. Appraisal trends. Personnel Journal. September: 139-145.
- Locke, E. A. 1986. Generalizing from laboratory to field settings. Lexington, MA: Lexington Books.
- Longenecker, C. O., & Gioia, D. A. 1988. Neglected at the top - executives talk about executive appraisals. Sloan Management Review, 21: Winter, 41-47.
- Longenecker, C. O., Sims, H. P., & Gioia, D. A. 1987. Behind the mask: The politics of employee appraisal. The Academy of Management Executive, 1: 183-193.
- Lord, R. G. 1985. Accuracy in behavioral measurement: An alternative definition based on raters' cognitive schema and signal detection theory. Journal of Applied Psychology, 70: 66-71.
- Lord, R. G., & Maher, K. J. 1989. Cognitive processes in industrial and organizational psychology. In C. L. Cooper & I. Robertson (Eds.) International review of industrial and organizational psychology: 49-91. New York: John Wiley & Sons, Ltd.
- Lowe, T. R. 1986. Eight ways to ruin a performance review. Personnel Journal. January: 60-62.
- Martin, D. C., & Bartol, K. M. 1986. Training the raters: A key to effective performance appraisal. Public Personnel Management, 15: 101-109.
- McBriarty, M. A. 1988. Performance appraisal: Some unintended consequences. Public Personnel Management, 17: 421-434.
- McEvoy, G. M. 1988. Evaluating the boss. Personnel Administrator. September: 115-120.

- McEvoy, G. M. 1990. Public sector managers' reactions to appraisals by subordinates. Public Personnel Management, 19: 201-212.
- McEvoy, G. M., & Buller, P. F. 1987. User acceptance of peer appraisals in an industrial setting. Personnel Psychology, 40: 785-797.
- McEvoy, G. M., & Cascio, W. F. 1989. Cumulative evidence of the relationship between employee age and job performance. Journal of Applied Psychology, 74: 11-17.
- Metz, E. J. 1988. Designing legally defensible performance appraisal systems. Training and Development Journal. July: 47-51.
- Meyer, J. P., Paunonen, S. V., Gattaly, I. R., Goffin, R. D., & Jackson, D. N. 1989. Organizational commitment and job performance: It's the nature of the commitment that counts. Journal of Applied Psychology, 74: 152-156.
- Milkovich, G. T., & Newman, J. M. 1991. Compensation. Plano, TX: Business Publications, Inc.
- Milkovich, G. T., & Wigdor, A. K. 1991. Pay for performance. Evaluating performance appraisal and merit pay. Washington, DC: National Academy Press.
- Miller, C. S., Kaspin, J. A., & Schuster, M. H. 1990. The impact of performance appraisal methods on age discrimination in employment act cases. Personnel Psychology, 43: 555-578.
- Miller, K. I., & Monge, P. R. 1986. Participation, satisfactions, and productivity: A meta-analytic review. Academy of Management Journal, 29: 727-753.
- Moen, R. D. 1989. The performance appraisal system: Deming's deadly disease. Quality Progress. November: 62-66.
- Moss, S. M. 1989. Appraise your performance appraisal process. Quality Progress. November: 58-60.
- Mount, M. K., & Thompson, D. E. 1987. Cognitive categorization and quality of performance ratings. Journal of Applied Psychology, 72: 240-246.

- Murphy, K. R. 1991. Criterion issues in performance appraisal research: Behavioral accuracy versus classification accuracy. Organizational Behavior and Human Decision Processes, 50: 45-50.
- Murphy, K. R., & Balzer, W. K. 1989. Rater errors and rating accuracy. Journal of Applied Psychology, 74: 619-624.
- Murphy, K. R., & Balzer, W. K. 1986. Systematic distortions in memory-based behavior ratings and performance evaluations: Consequences for rating accuracy. Journal of Applied Psychology, 71: 39-44.
- Murphy, K. R., Balzer, W. K., Lockhart, M. C., & Eisenman, E. J. 1985. Effects of previous performance on evaluations of present performance. Journal of Applied Psychology, 70: 72-84.
- Murphy, K. R., & Constans, J. I. 1987. Behavioral anchors as a source of bias in rating. Journal of Applied Psychology, 72: 573-577.
- Murphy, K. R., Gannett, B. A., Herr, B. M., & Chen, J. A. 1986. Effects of subsequent performance on evaluations of previous performance. Journal of Applied Psychology, 71: 427-431.
- Murphy, K. R., Herr, B. M., Lockhart, M. C., & Maguire, E. 1986. Evaluating the performance of paper people. Journal of Applied Psychology, 71: 654-661.
- Murphy, K. R., & Pardaffy, V. A. 1989. Bias in behaviorally anchored rating scales: Global or scale-specific? Journal of Applied Psychology, 74: 343-346.
- Murphy, K. R., Philbin, T. A., & Adams, S. R. 1989. Effect of purpose of observation on accuracy of immediate and delayed performance ratings. Organizational Behavior and Human Decision Processes, 43: 336-354.
- Murphy, K. R., & Reynolds, D. H. 1988. Does true halo affect observed halo? Journal of Applied Psychology, 73: 235-238.

- Naffziger, D. W. 1985. BARS, RJP's and recruiting. Personnel Administrator. August: 85-96.
- Nanry, C. 1988. Performance linked training. Public Personnel Management, 17: 457-464.
- Napier, N. K., & Latham, G. P. 1986. Outcome expectancies of people who conduct performance appraisals. Personnel Psychology, 39: 827-837.
- Nathan, B. R., & Alexander, R. A. 1985. The role of inferential accuracy in performance rating. Academy of Management Review, 10: 109-115.
- Nathan, B. R., & Tippins, N. 1990. The consequences of halo "error" in performance ratings: A field study of the moderating effect of halo on test validation results. Journal of Applied Psychology, 75: 290-296.
- Northcraft, G. B., & Ashford, S. J. 1990. The preservation of self in everyday life: The effects of performance expectation and feedback context on feedback inquiry. Organizational Behavior and Human Decision Processes, 47: 42-64.
- Olian, J. D., & Rynes, S. L. 1984. Organizational staffing: Integrating practice with strategy. Industrial Relations, 23: 170-183.
- Oliver, J. E. 1985. Performance appraisals that fit. Personnel Journal. June: 66-71.
- Padgett, M. Y., & Ilgen, D. R. 1989. The impact of ratee performance characteristics on rater cognitive processes and alternative measures of rater accuracy. Organizational Behavior and Human Decision Processes, 44: 232-260.
- Pearce, J. L., & Porter, L. W. 1986. Employee responses to formal performance appraisal feedback. Journal of Applied Psychology, 71: 211-218.
- Perry, J., & Petrakis, B. A. 1988. Part I - Developments in organizational performance. Can pay for performance succeed in government? Public Personnel Management, 17: 359-368.

- Podsakoff, P. M. & Farh, J. L. 1989. Effects of feedback sign and credibility on goal setting and task performance. Organizational Behavior and Human Decision Processes, 44: 45-67.
- Pooyan, A., and Eberhardt, B. J. 1989. Correlates of performance appraisal satisfaction among supervisory and nonsupervisory employees. Journal of Business Research, 19: 215-226.
- Prien, E. P., & Hughes, G. L. 1987. The effect of quality control revisions on mixed standard scale rating errors. Personnel Psychology, 40: 815-823.
- Prince, J. B., & Lawler, E. E. 1986. Does salary discussion hurt the developmental performance appraisal? Organizational Behavior and Human Decision Processes, 37: 357-375.
- Pulakos, E. D., 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: 29-32.
- Pulakos, E. D., White, L. A., Oppler, S. H., & Borman, W. C. 1989. Examination of race and sex effects on performance ratings. Journal of Applied Psychology, 74: 770-780.
- Reed, P. R., & Kroll, M. J. 1985. A two-perspective approach to performance appraisal. Personnel. October: 51-57.
- Regel, R. W., & Hollmann, R. W. 1987. Gauging performance objectively. Personnel Administrator. June: 74-78.
- Reinhardt, C. 1985. The state of performance appraisal: A literature review. Human Resource Planning, 8: 105-110.
- Rodgers, T. J. 1990. No excuses management. Harvard Business Review. July-August: 84-98.

- Romanoff, K. E. 1989. The ten commandments of performance management. Personnel. January: 24-28.
- Rothstein, H. R. 1990. Interrater reliability of job performance ratings: Growth to asymptote level with increasing opportunity to observe. Journal of Applied Psychology, 75: 322-327.
- Russell, J. S., & Goode, D. L. 1988. An analysis of managers' reactions to their own performance appraisal feedback. Journal of Applied Psychology, 73: 63-67.
- Sackett, P. R., Zedeck, S., & Fogli, L. 1988. Relations between measures of typical and maximum job performance. Journal of Applied Psychology, 73: 482-486.
- Sahl, R. J. 1990. Design effective performance appraisals. Personnel Journal. October: 53-60.
- Sandler, L. 1990. Two-sided performance reviews. Personnel Journal. January: 75-78.
- Sankowsky, D. 1989. A psychoanalytic attributional model for subordinate poor performance. Human Resource Management, 28: 125-139.
- Scherkenbach, W. W. 1985. Performance appraisal and quality: Ford's new philosophy. Quality Progress. April: 40-46.
- Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. 1986. Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. Journal of Applied Psychology, 71: 432-439.
- Schmitt, N., Noe, R. A., & Gottschalk, R. 1986. Using the lens model to magnify raters' consistency, matching, and shared bias. Academy of Management Journal, 29: 130-139.
- Schneier, C. E. 1989. Capitalizing on performance management, recognition, and reward systems. Compensation and Benefits Review, 21: 20-30.
- Schneier, C. E., Beatty, R. W., & Baird, L. S. 1986a. Creating a performance management system. Training and Development Journal. May: 74-79.

- Schneier, C. E., Beatty, R. W., & Baird, L. S. 1986b. How to construct a successful performance appraisal system. Training and Development Journal. April: 38-42.
- Schneier, C. E., Geis, A., & Wert, J. A. 1987. Performance appraisals: No appointment needed. Personnel Journal. November: 80-87.
- Scholtes, P. R. 1987. An elaboration on Deming's teachings on performance appraisal. Madison, WI: Joiner Associates, Inc.
- Schoorman, F. D. 1988. Escalation bias in performance appraisals: An unintended consequence of supervisor participation in hiring decisions. Journal of Applied Psychology, 73: 58-62.
- Schrader, A. W. 1989. MBO makes dollar sense. Personnel Journal. July: 32-37.
- Schweiger, D. M., & Leana, C. R. 1986. Participation in decision making. In E. A. Locke (Ed.), Generalizing from laboratory to field settings: 147-166. Lexington, MA: Lexington Books.
- Shapiro, G. L., & Dessler, G. 1985. Are self appraisals more realistic among professionals or nonprofessionals in health care? Public Personnel Management, 14: 285-290.
- Shore, L. M., & Thornton, G. C. 1986. Effects of gender on self - and supervisory ratings. Academy of Management Journal, 29: 115-129.
- Sims, R. R., Veres, J. G., & Heninger, S. M. 1987. Training appraisers: An orientation program for improving supervisory performance ratings. Public Personnel Management, 16: 37-46.
- Slattery, P. D. 1985. Performance appraisal without stress. Personnel Journal. February: 49-52.
- Smith, D. E. 1986. Training programs for performance appraisal: A review. Academy of Management Review, 11: 22-40.

- Smither, J. W., Barry, S. R., & Reilly, R. R. 1989. An investigation of the validity of expert true score estimates in appraisal research. Journal of Applied Psychology, 74: 143-151.
- Smither, J. W., Collins, H., & Buda, R. 1989. When ratee satisfaction influences performance evaluations: A case of illusory correlation. Journal of Applied Psychology, 74: 599-605.
- Smither, J. W., & Reilly, R. R. 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: 369-391.
- Smither, J. W., Reilly, R. R., & Buda, R. 1988. Effect of prior performance information on ratings of present performance: Contrast versus assimilation revisited. Journal of Applied Psychology, 73: 487-496.
- Solomon, R. J. 1990. Developing job specific appraisal factors in large organizations. Public Personnel Management, 19: 11-24.
- Srinivas, S., & Motowidlo, S. J. 1987. Effects of raters' stress on the dispersion and favorability of performance ratings. Journal of Applied Psychology, 72: 247-251.
- Steel, B. S. 1985. Participative performance appraisal in Washington: An assessment of post-implementation receptivity. Public Personnel Management, 14: 153-171.
- Steel, R. P., & Mento, A. J. 1986. Impact of situational constraints on subjective and objective criteria of managerial job performance. Organizational Behavior and Human Decision Processes, 37: 254-265.
- Steiner, D. D., & Rain, J. S. 1989. Immediate and delayed primacy and recency effects in performance evaluation. Journal of Applied Psychology, 74: 136-142.
- Sulsky, L. M., & Balzer, W. K. 1988. Meaning and measurement of performance rating accuracy: Some methodological and theoretical concerns. Journal of Applied Psychology, 73: 497-506.

- Tafti, P. M. 1990. Face to face. Training and Development Journal. November: 66-71.
- Tenopyr, M. L. 1988. Artifactual reliability of forced-choice scales. Journal of Applied Psychology, 73: 749-751.
- Thompson, D. E., & Thompson, T. A. 1985. Task-based performance appraisal for blue-collar jobs: Evaluation of race and sex effects. Journal of Applied Psychology, 70: 747-753.
- Thorndike, R. L. (1949). Personnel selection: Test and measurement techniques. New York: Wiley.
- Tsui, A. S., & Barry, B. 1986. Interpersonal affect and rating errors. Academy of Management Journal, 29: 586-599.
- Vance, R. J., MacCallum, R. C., Coovert, M. D., & Hedge, J. W. 1988. Construct validity of multiple job performance measures using confirmatory factor analysis. Journal of Applied Psychology, 73: 74-80.
- Vines, L. S. 1988. How's the boss doing? Human Resource Executive. May: 36-38.
- Vinton, K. L. 1990. Documentation that gets results. Personnel. February: 42-46.
- VonGlinow, M. A. 1985. Reward strategies for attracting, evaluating, and retaining professionals. Human Resource Management, 24: 191-206.
- Vroom, V. H. 1964. Work and motivation. New York: Wiley and Sons.
- Wagel, W. H. 1987. Performance appraisal with a difference. Personnel. February: 4-6.
- Wagner, J. A., & Gooding, R. Z. 1987. Shared influence and organizational behavior: A meta-analysis of situational variables expected to moderate participation-outcome relationships. Academy of Management Journal, 30: 524-541.
- Waldman, D. A., & Kenett, R. S. 1990. Improve performance by appraisal. HR Magazine. July: 66-69.
- Waldman, D. A., Yammarino, F. J., & Avolio, B. J. 1990. A multiple level investigation of personnel ratings. Personnel Psychology, 43: 811-835.

- Warrenfeltz, R. B. 1989. An achievement based approach to evaluating engineering technicians. Public Personnel Management, 18: 243-262.
- Wayne, S. J., & Ferris, G. R. 1990. Influence tactics, affect, and exchange quality in supervisor-subordinate interactions: A laboratory experiment and field study. Journal of Applied Psychology, 75: 487-499.
- Wehrenberg, S. B. 1988. Train supervisors to measure and evaluate performance. Personnel Journal. February: 77-79.
- Wexley, K. N., & Klimoski, R. 1984. Performance appraisal: An update. In K. M. Rowland & G. R. Ferris (Eds.), Research in personnel and human resources management, Vol. 2: 35-79. Greenwich, CT: JAI Press.
- Wexley, K. N., & Snell, S. A. 1987. Managerial power: A neglected aspect of the performance appraisal interview. Journal of Business Research, 15: 45-54.
- Wherry, R. J. 1983. Wherry's theory of rating. In F.L. Landy & J.L. Farr (Eds.), The measurement of work performance. Methods, theory, and applications: 283-303. Orlando, FL: Academic Press.
- Wight, D. T. 1985. The split role in performance appraisal. Personnel Administrator. May: 83-87.
- Williams, K. J., Cafferty, T. P., & DeNisi, A. S. 1990. The effect of performance appraisal salience on recall and ratings. Organizational Behavior and Human Decision Processes, 46: 217-239.
- Williams, K. J., DeNisi, A. S., Blencoe, A. G., & Cafferty, T. P. 1985. The role of appraisal purpose: Effects of purpose on information acquisition and utilization. Organizational Behavior and Human Decision Processes, 35: 314-339.
- Williams, K. J., DeNisi, A. S., Meglino, B. M., & Cafferty, T. P. 1986. Initial decisions and subsequent performance ratings. Journal of Applied Psychology, 71: 189-195.

Woods, J. G., & Dillion, T. 1985. The performance review approach to improving productivity. Personnel. March: 20-27.

Wyatt Company, The. 1989. Results of the 1989 Wyatt survey: Getting your hands around performance management. The Wyatt Communicator. Fourth Quarter: 4-18.
Washington, DC: The Wyatt Company, Research and Information Center.

Zalesny, M. D. 1990. Rater confidence and social influence in performance appraisal. Journal of Applied Psychology, 75: 274-289.

APPENDIX

Summary of Research and Managerial Performance Appraisal Literature

| Primary Issues Explored | Methodology | | | | | | | |
|------------------------------------|---|--|---|------------------------|------------|--|----------|--|
| | Field Study | Lab Study | Lit Review | Survey | Discussion | Case Study | "How To" | |
| Information Processing | Hogan, 1987 Huber et al., 1986 Jolly et al., 1988 Mount & Thompson, 1987 Schmitt et al., 1986 | Balzer, 1986 Brannick & Brannick, 1989 DeNisi et al., 1989 Feldman et al., 1986 Huber et al., 1987 Karl & Wexley, 1989 Kinicki & Griffeth, 1985 Kozlowski & Kirsch, 1987 Kozlowski et al., 1986 Krzystofiak et al., 1988 Murphy & Balzer, 1986 Murphy et al., 1985 Murphy, Gannett et al., 1986 Murphy et al., 1989 Padgett & Ilgen, 1989 Smither & Reilly, 1987 Smither et al., 1988 Srinivas & Motowidlo, 1987 Steiner & Rain, 1989 Williams et al., 1985 Williams et al., 1986 Williams et al., 1990 | Lord & Maher, 1989 | | | Banks & Murphy, 1985 Dipboye, 1985 Murphy, 1991 Nathan & Alexander, 1985 Sankowaky, 1989 Schmidt et al., 1986 | | |
| Rater/Ratee Characteristics | Dobbins et al., 1988 Drazin & Auster, 1987 Fedor & Rowland, 1989 Ferris et al., 1985 Heneman et al., 1989 Kingstrom & Mainstone, 1985 Lawrence, 1988 Pulakos et al., 1989 Shore & Thornton, 1986 Thompson & Thompson, 1985 Tsui & Barry, 1986 | Benedict & Levine, 1988 Cardy & Dobbins, 1986 Czajka & DeNisi, 1988 Dobbins et al., 1988 Galini & Benoliel, 1990 Heitman & Stopeck, 1985 | Kraiger & Ford, 1985 McEvoy & Cascio, 1989 | Greenhaus et al., 1990 | | | | |

| Primary Issues Explored | Methodology | | | | | | "How To" | |
|---------------------------------|--|--|---|--|--|--|--|--|
| | Field Study | Lab Study | Lit Review | Survey | Discussion | Case Study | | |
| Errors and Accuracy | Fulk et al., 1985 Jacobs & Kozlowski, 1985 Kinas & DeNisi, 1989 Nathan & Tippins, 1990 Rothstein, 1990 Schmitt et al, 1986 Schoorman, 1988 | Becker & Cardy, 1986 Kamouri & Balzer, 1990 Karl & Wexley, 1989 Kozlowski & Kirach, 1987 Murphy & Constans, 1987 Murphy & Pardaffy, 1989 Murphy & Reynolds, 1988 Murphy et al., 1989 Padgett & Ilgen, 1989 Smither, Barry & Reilly, 1989 Smither, Collins & Buda, 1989 | Heneman et al., 1987 Murphy & Balzer, 1989 | | | Dickinson, 1987 Feldman, 1986 Fisicaro, 1988 Lance & Woehr, 1986 Lord, 1985 Murphy, 1991 Pulakos et al., 1986 Sankowsky, 1989 | Sulsky & Balzer, 1988 | |
| Feedback | Alexander et al., 1989 Becker & Klimoski, 1989 Dorfman et al., 1986 Earley, 1988 Pearce & Porter, 1986 Prince & Lawler, 1986 Russell & Goode, 1988 Wexley & Snell, 1987 | Bannister, 1986 DeGregorio & Fisher, 1988 Dugan, 1989 Fedor et al., 1989 Ilgen & More, 1987 Klimoski & Inks, 1990 Northcraft & Ashford, 1990 Podsakoff & Farh, 1989 | | | Larson, 1989 | Nanry, 1988 | Harper, 1986 Tafti, 1990 Wight, 1985 | |
| Raters/Appraisal Sources | Eder & Fedor, 1989 Farh et al., 1988 Fox & Dinur, 1988 Fox et al., 1989 George & Smith, 1990 McEvoy & Buller, 1987 Shapiro & Dessler, 1985 Vance et al., 1988 | Farh & Dobbins, 1989 Farh & Werbel, 1986 Zalesny, 1990 | Harris & Schaubroeck, 1988 Heneman, 1986 | McEvoy, 1990 Steel, 1985 | Bernardin, 1986 Bernardin & Abbott, 1985 Campbell & Lee, 1988 McEvoy, 1988 Vince, 1988 | Bernardin & Beatty, 1987 | Edwards, 1990b Edwards & Sproull, 1985 Lanza, 1985 Lawrie, 1989 | |
| Rater Training | Dugan, 1988 Hedge & Kavanagh, 1988 | Athey & McIntyre, 1987 | Smith, 1986 | | Banks & Roberson, 1985 Wehrenberg, 1988 | | Brown, 1987 Kaufman, 1988 Martin & Bartol, 1986 Sims et al., 1987 | |
| Formats | Harris, 1988 Hughes & Prien, 1986 Prien & Hughes, 1987 | Murphy & Constans, 1987 Murphy & Pardaffy, 1989 Tenopyr, 1988 | | Kane & Freeman, 1986 Kane & Freeman, 1987 | Angel, 1989 Naffziger, 1985 | McBriarty, 1988 Schrader, 1989 | Gibb, 1985 Oliver, 1985 | |

| Primary Issues Explored | Methodology | | | | | | |
|-------------------------------------|---|-----------|--|---|---|---|--|
| | Field Study | Lab Study | Lit Review | Survey | Discussion | Case Study | "How To" |
| Reactions To Appraisals | Dobbins et al., 1990 Giles & Mossholder, 1990 Napier & Latham, 1986 Pooyan & Eberhardt, 1989 | | | | | | |
| Legal Ramifications | | | Barrett & Kernan, 1987 Griffin, 1989 Miller et al., 1990 | Daley, 1990 | Goddard, 1989 | | Metz, 1988 Vinton, 1990 |
| Social and Political Aspects | Kipnis & Schmidt, 1988 Longenecker et al., 1987 Wayne & Ferris, 1990 Wexley & Snell, 1987 | | | | | | |
| Performance Management | | | | | Brumback, 1988 Kroll & Joyce, 1989 | Perry & Petrakis, 1988 Schneier, 1989 | Day, 1989 Kirkpatrick, 1986 Levy, 1989 Romanoff, 1989 Schneier et al., 1986a Schneier et al., 1986b Schneier et al., 1987 |
| PA Practices | Clement & Stevens, 1989 | | Reinhardt, 1985 | Cleveland et al., 1989 Locher & Teal, 1988 Warrenfeltz, 1989 Wyatt, 1989 | Derven, 1990 Edwards, 1990a Longenecker & Gioia, 1988 | Cayer et al., 1988 Cocheu, 1986 Cozzetto, 1990 Deets & Tyler, 1986 Gellerman & Hodgson, 1988 Glen, 1990 Goodell, 1988 Gomez-Mejia et al., 1985 Hall, 1987 Laumeyer & Beebe, 1988 Scherkenbach, 1985 Wagel, 1987 Woods & Dillion, 1985 | Buford et al., 1988 Dearden, 1987 Eyres, 1989 Gilbert & Nelson, 1989 Girard, 1988 Howard, 1987 Krein, 1990 Lawrie, 1990 Lee, 1989, 1990 Reed & Kroll, 1985 Regel & Hollmann, 1987 Rodgers, 1990 Sahl, 1990 Sandler, 1990 Solomon, 1990 Waldman & Kenett, 1990 |

| Primary Issues Explored | Field Study | Lab Study | Methodology | | Discussion | Case Study | "How To" |
|-------------------------|---|-----------|---|--------|--|------------|----------------|
| | | | Lit Review | Survey | | | |
| Others | Greenberg, 1986 Kerr, 1988 Meyer, et al., 1989 Sackett et al., 1988 Steel & Mento, 1986 Waldman et al., 1990 | | Herbert & Doverspike, 1990 Murphy, Herr et al., 1986 | | Fedor & Buckley, 1988 Greenberg, 1988 Greenberg, 1990 Hyde, 1988 Lowe, 1986 Moen, 1989 Moss, 1989 Scholtes, 1987 Slattery, 1985 VonGlinow, 1985 | | Friedman, 1986 |