CHAPTER – I

INTRODUCTION
Performance appraisal has long been regarded as one of the most critical yet troubling areas of human resource management.\(^1\) Performance appraisal and feedback is the formal, systematic assessment of how well employees are performing their jobs in relation to the established standards and the communication of that assessment to employees. How appraisal systems are designed, developed and implemented, and how appraisal results are communicated affect morale and organisational climate in significant ways.\(^2\) Performance appraisal and feedback is an emotionally laden process that dramatically affects employees' attitudes towards the organisation and themselves. If used effectively, performance appraisal can improve employee's motivation, performance, and satisfaction. If used inappropriately, the appraisal process can have disastrous effects.

Researchers are fascinated by this subject; yet their inability to resolve definitively the knotty technical problems of performance appraisal has led Heneman (1975) to term it the “Achilles heel” of personnel management.\(^3\) Deming, the guru of the total quality management movement, identified appraisal as one of his infamous


"deadly diseases" of management,¹ and urged that the practice be "purged from the earth."²

Whether or not, one approves of performance appraisal and feedback, there is a sense in which all organisations do practise it, however crudely. In this connection, Schick (1980) maintains that:

"Every company has a performance evaluation monitoring system (PEMS) to measure individual performance against company standards. It may not be recognised, formalised or documented, but the need to evaluate performance is present in all business environments."³

In small concerns with a few employees, the owner or manager knows all his/her people personally, often working closely with them, and thus is well able to assess their performance at first hand. He or she does not need to go through all the motions of filling in formal performance appraisal and feedback documents in order to reach valid conclusions on the current performance, future potential, and salary status of his/her personnel. However, in a larger organisation with several hierarchical levels, first-line and middle-level managers are often quite remote from their top management superiors, thus necessitating the introduction of formal procedures for performance appraisal and feedback to aid the latter make similar decisions.

Organisations are constantly searching for alternative and better ways to appraise performance.¹

**CONCEPTUAL AND THEORETICAL FRAMEWORK**

Systematic employee appraisal techniques came into prominence during and immediately after the World War I. Initially, a "man-to-man" rating system was introduced to evaluate United States military officers. During the 1920's and 1930's, industrial concerns began installing rational wage structures for their 'hourly employees'. It should be noted that the early employee appraisal schemes were called "merit-rating" programmes. In fact, the term merit rating was by far the most widely accepted title until the mid-1950's.²

Commencing in the early 1950’s, great interest developed in the performance appraisal of technical, professional, and managerial personnel. It was recognised that appraisal, on a systematic basis, was an integral part of a well-designed employee development programme. The term 'performance appraisal' is also known by other names such as personnel appraisal, personnel review, progress report, service rating, effectiveness report, and performance evaluation.

Performance appraisal is the systematic evaluation of the individual with regard to his/her performance on the job and his/her potential for development. In this sense, it is the process of determining the relative worth of each and every employee in the


Bernardin and Beatty (1984) defined 'performance' as: The record of outcomes produced on a specified job function or activity during a specified time period. Performance on the job as a whole would be equal to the sum (or average) of performance on the critical or essential job functions. The functions have to do with the work which is performed and not the characteristics of the person performing. Unfortunately, many performance appraisal systems still have confused measures of performance with measures or traits of the person which violates the fundamental objective of performance appraisal: to evaluate performance of the person, not the person in the abstract. Hence, 'performance' refers to the degree of accomplishment of the tasks that make up an individual's job. It reflects how well an individual is fulfilling the requirements of a job.

Effective performance appraisal focuses on the record of outcomes and, in particular, outcomes directly linked to an organisation's mission and objectives. In general, personal traits (e.g., dependability, integrity, perseverance, loyalty) should not be used when evaluating performance because such traits may foster stereotyping and other biases and are difficult to defend should litigation result. Furthermore, people who are evaluated on ambiguous criteria perceive little value in the feedback, and are often less motivated to perform well after the appraisal than before.

---


Few organisations clearly define what it is they are trying to measure in work. Strategic goals are rarely carefully defined or measured. All other things being equal, the greater the specificity in the criterion definition of “performance”, the more effective the appraisal system will be, no matter what purpose it is designed to serve. In addition, the stronger the connection between the strategic goals of a division, strategic business unit (SBU) or functional unit and these criteria, the more effective the appraisal system will ultimately be.

Expectancy Theory and Performance Appraisal

Performance is a vital component of expectancy theory. Specifically, one must be concerned with the linkage between effort and performance, and between performance and rewards. It should be ascertained whether people see effort leading to performance, and performance to the rewards that they value. They have to know clearly what is expected of them. They need to know how their performance will be measured. Furthermore, they must feel confident that if they exert an effort within their capabilities, it will result in a satisfactory performance as defined by the criteria by which they are being measured. Finally, they must feel confident that if they perform as they are being asked, they will achieve the rewards, they value.

In brief, it is obvious that performance appraisal plays a vital role in the expectancy model of motivation. If the objectives that individual employees are seeking are unclear, if the criteria for measuring that objective attainment are vague, and if employees lack confidence that their efforts will lead to a satisfactory appraisal of

---

their performance or feel that there will be an unsatisfactory payoff by the organisation when their performance objectives are achieved, one can expect individuals to work considerably below their potential. If the job of acquiring capable people and developing their basic abilities to do the job has been done, it must also be made sure that they know what behaviours are required of them, understand how they are going to be appraised, believe that the appraisal will be conducted in a fair and equitable manner, and can expect their performance to be recognised by proper rewards.

**Attribution Theory and Performance Appraisal**

In attribution theory, it is suggested that "employee evaluations are directly mediated by managers' attributions as to who is perceived to be in control of the employee's performance - the employee or the manager."\(^1\) Much of attribution theory is consistent with McGregor's Theory X, Theory Y.\(^2\) That is, if a manager has more of a Theory Y orientation and uses a more democratic and participative approach, there is a tendency for that manager to rate the subordinates higher. This higher rating appears to be attributed solely to the manager's perception of worker initiative. Conversely, a more Theory X-oriented manager, one who is more autocratic and makes unilateral decisions, is likely to rate his or her subordinates lower. In this sequence, such managers feel that they controlled the work and accordingly, performance is poorer.

---


One research study mentioned that the two generalisations regarding attribution theory were supported:¹

1. When managers attribute an employee’s poor performance to the employee’s own motivations, the judgements are harsher than when the poor performance is attributed to external factors.

2. When an employee is performing satisfactorily, managers will evaluate the employee favourably if the performance is attributed to the employee’s own efforts than if the performance is attributed to outside forces.

**Performance Management**

Conventional performance evaluation systems are usually built around on an annual event, the formal review, which tended to dwell on the past. This was carried out at the behest of the personnel department, often perfunctorily, and then forgotten. Managers proceeded to manage without any further reference to the outcome of the review and the appraisal form was buried in the personnel record system. This system adopted a reactive, directionless post-mortem treatment of the past performance and in the process, it had lost its utility and credibility. This paved the way for the emergence of a more proactive, purposeful, integrated system of ‘performance management’ with an ideological underpinning and a holistic context.

Performance management is a means of getting better results from the organisation, teams, and individuals by understanding and managing performance within an agreed framework of planned goals, standards and competence requirements. It is a process for establishing shared understanding about what is to be achieved, and an approach to managing and developing people in a way which

increases the probability that it will be achieved in the short and long term. It is owned and driven by line management.

Performance management is concerned with the interrelated processes of work, management, development, and reward. It can become a powerful integrating force, ensuring that these processes are linked together properly as a fundamental part of the human resource management approach which should be practised by every manager in the organisation.

Performance management is a continuous and flexible process which involves managers and those whom they manage acting as partners within a framework which sets out how they can best work together to achieve the required results. It focuses on future performance planning and improvement rather than on retrospective performance appraisal. It provides the basis for regular and frequent dialogues between managers and individuals or teams about performance and development needs.

To a larger extent, performance management is a process for measuring outputs in the shape of delivered performance compared with expectations expressed, as objectives. That is why it focuses on targets, standards and performance measures or indicators. But it is also concerned with inputs and processes – the knowledge, skills, attitudes and behaviour required to produce the required results. It is by defining these input requirements and assessing the extent to which the expected levels of performance have been achieved by using skills and competencies effectively that developmental needs are identified.
The following key activities are involved in a typical performance management process:

(1) The performance agreement or contract – what the individual or team has to achieve in the form of objectives; how performance will be measured and the competencies needed to deliver the required results.

(2) The performance and development plan – sets out an agreement on performance and personal development needs.

(3) Managing performance throughout the year – the continuous process of providing feedback on performance, conducting informal progress reviews, and, where necessary, dealing with performance problems.

(4) Performance review – the formal review of performance over a period covering achievements and problems as the basis for agreeing as necessary revisions to the performance agreement and performance plan.

Performance management aims to enhance what Alan Mumford (1994) calls 'deliberate learning from experience', which means learning from the problems, challenges, and successes inherent in people's day-to-day activities.¹ The premise is that every task individuals undertake presents them with a learning opportunity, as long as they reflect or are helped to reflect on what they have done and how they have done it, and draw conclusions as to their future behaviour if they have to carry out a similar task.

Fletcher and Williams (1992) have aptly concluded that "the real concept of performance management is associated with an

approach to creating a shared vision of the purpose and aims of the organisation, helping each individual employee understand and recognise his/her part in contributing to them, and in so doing manage and enhance the performance of both individuals and the organisation."1 According to them, the following are the four underlying principles of effective performance management:

(1) that it is owned and driven by line management and not by the human resources department;
(2) that there is a emphasis on shared corporate goals and values;
(3) that performance management is not a package solution, it is something that has to be developed specifically and individually for the particular organisation; and
(4) that it should apply to all staff, not just part of the managerial group.

From this discussion, it may be said that 'performance management' is a means whereby employees' work behaviours are aligned with the organisation's goals. There is no one way to manage performance. Whatever system is adopted needs to be congruent with the culture and principles that pervade the organisation.

---

Purposes/Uses of Performance Appraisal

Performance appraisal is used in organisations for a variety of purposes.

According to Beer (1985) the many purposes of performance appraisal can be grouped into two main categories: evaluation goals, and coaching and development goals.¹ Evaluation goals include (1) giving feedback to employees so they know where they stand; (2) developing valid data for pay (salary and bonus pay) and promotion decisions, and providing a means for communicating these decisions; (3) helping the manager to make retention and discharge decisions, and providing a means for warning employees about unsatisfactory performance. On the other hand, coaching and development goals include (1) counselling and coaching employees in order to improve their performance and develop future potential; (2) developing commitment to the organisation through discussion of career opportunities and career planning; (3) motivating employees through recognition and support; (4) strengthening superior-subordinate relations; and (5) diagnosing individual and organisational problems.

When appraisals are used for discipline, pay increases, promotions, discharges, or layoffs, they are likely to be regarded with an apprehension by those employees who tend to underestimate themselves, by the less productive members of the organisation, or by those who feel appraisals will be arbitrary or unjust. Similar feelings result when employees do not know or understand the criteria being used.

used to assess them, or when they see the appraisal as highly critical rather than helpful.¹

According to Glueck (1979) formal performance evaluation "is a system set up by the enterprise to regularly and systematically evaluate employee performance"²; while Kenney, Donnelly, and Reid (1979) hold that "appraisal of an employee's work by his superior once a year, or as appropriate, enables a manager to take stock of his or her staff's strengths and weaknesses."³

McGregor (1960) listed the objectives of performance appraisal as threefold⁴: (1) administrative – in providing an orderly way of determining promotions, transfers, salary increases, etc., (2) informative – in supplying data to management on the performance of subordinates and to the individual on his/her strengths and weaknesses; and (3) motivational – in creating a learning experience for subordinates that motivates them to improve.

Although professional and trade journals in the field of personnel provide extensive anecdotal evidence⁵ concerning the uses

of performance appraisals in organisations, few empirical studies have been devoted to this question.

Levine (1986) found that the most common use for performance appraisal was for determining employee training needs, merit review, and salary administration.1 Rendero (1980) surveyed 24 human resource managers and found that the most frequently mentioned uses included merit review or salary action, employee development, and feedback to employees.2 Campbell, Dunnette, Lawler, and Weick's (1970) survey of 33 organisations suggested that the performance appraisals were used in placement and promotion decisions, training needs assessment, and as a motivational tool.3

It should be noted that performance appraisal systems often involve multiple, conflicting goals.4 For example, an organisation may desire to use performance appraisal as a tool for making promotion related decisions and also as a tool for guiding employee development. Using appraisal to make promotion recommendations, the rater is obliged to focus on comparisons between individuals',

---

whereas 'within-individuals' comparisons are necessary for determining employee training and development needs.¹

Traditional performance appraisals often emphasise performance in terms of individual assessment and rewards or punishments. Besides effectively evaluating the performance of employees, performance appraisal systems also can be used to improve performance, motivate, and develop employees. Under the modern systems-approach, an effective performance appraisal system will not only evaluate employee's abilities, but also opportunities – those elements in the work system that contribute to performance. According to Waldman and Kenett (1990), opportunity factors are more important than individual variables in determining work performance.² Opportunity variables may include the physical environment, such as noise and lighting, available resources, such as human and computer assistance, and social processes, such as leadership. Thus, opportunity variables can influence performance indirectly by affecting individual variables, such as employee motivation.

**Performance Appraisal Process**

Performance appraisal involves two distinct processes: (1) observation, and (2) judgement. Observation processes are more basic, and include the detection, perception, and recall or recognition of specific behavioural events. Judgement processes include the

---


categorisation, integration, and evaluation of information. In practice, observation and judgement represent the last elements of a three-part sequence: job analysis, performance standards, and performance appraisal.

Job analysis identifies and describes work and personal requirements of a particular job. It is management's responsibility to establish performance standards: the levels of performance deemed acceptable or unacceptable for each of the job-relevant, critical areas of performance identified through job analysis.

Performance standards are essential in all types of goods-producing and service organisations, for they help ensure consistency in supervisory judgements across individuals in the same job. Unfortunately, it is often the case that charges of unequal treatment and unfair discrimination arise in jobs where no clear performance standards exist.

Performance appraisal, the last step in the sequence, represents the actual process of gathering information about individuals based on critical job requirements. Gathering job performance information is accomplished by observation. Evaluating the adequacy of individual performance is judgement.

---


Types of Performance Measures

Performance measures may be classified into two general types: objective and subjective. "Objective performance measures" include production data (sales volume and value, units produced, number of errors, amount of scrap) as well as personnel data (accidents, turnover, absences, tardiness). These variables directly define the goals of the organisation, but they often suffer from several glaring weaknesses, the most serious of which are performance unreliability and modification of performance by situational characteristics. The underlying objective in performance appraisal, however, is to judge an individual's performance, not factors beyond his or her control. Moreover, objective measures focus not on behaviour, but rather on the outcomes or results of behaviour. Admittedly, there will be some degree of overlap between behaviour and results, but the two are qualitatively different.1

In short, although objective measures of performance are intuitively attractive, theoretical and practical limitations often make them unsuitable. Although they can be useful when used as supplements to supervisory judgements, correlations between objective and subjective measures are often low2 and when used as basis for personnel decisions, the combination of such measures may be weighed differently for different ethnic groups.3


The disadvantages of objective measures have led researchers and managers to place major emphasis on "subjective measures" of job performance. However, since subjective measures depend upon human judgement, they are prone to certain kinds of errors associated with the rating process. Subjective measures of performance may be *relative* (in which comparisons are made among a group of ratees) or *absolute* (in which a ratee is described without reference to others.) Regardless of their form, subjective performance appraisals frequently suffer from various behavioural barriers that limit their effectiveness.

Behavioural barriers to successful performance appraisal may be political or interpersonal.\(^1\) Political barriers stem from supervisory apprehension concerning the use to which the results of performance appraisal will be put by their peers. Interpersonal barriers arise from the actual face-to-face encounter between subordinate (appraisee) and superior (appraiser).

Political considerations are organisational facts of life. As long as performance appraisal data are neither too good nor too bad, managers have little difficulty competing for resources. Patz (1975) expressed the issue concisely:

> In fact the political barriers form a sort of law of organisation. Its statement is simple: Do not collect accurate information about your subordinates, for it will very likely be held against you.\(^2\)

---


2. Ibid. 77.
In this type of climate, personal values and bias can replace organisational standards. Unfairly low ratings may be given to highly valued subordinates so they will not be promoted out of a rater’s department; likewise, personal bias may lead to favoured treatment for other employees.

Even if political tension can be lessened, significant interpersonal barriers also may hinder the appraisal process. Because of a lack of communication, employees may think they are being judged according to one set of standards, when their superiors actually use different standards. Furthermore, supervisors typically resist making face-to-face appraisals. Such resistance can reduce appreciably the validity of the ratings because rather than confront substandard performers with low ratings, negative feedback, and below average salary increases, supervisors often find it easier to “damn with faint praise” by giving average or above average ratings to inferior performers.

Finally, some managers may complain that formal performance appraisal interviews tend to interfere with the more constructive coaching relationship that should exist between superior and subordinate. They may claim that appraisal interviews emphasise the superior position of the supervisor by placing him or her in the role of judge, thus conflicting with the supervisor’s equally important roles of teacher and coach.

From this it follows that performance appraisal is widely accepted as a potentially useful tool, but behavioural barriers often thwart its successful implementation. Human resource management professionals can contribute by improving the technical aspects of
performance appraisal systems as well as their attitudinal and interpersonal components.

**Categories of Performance Information**

There are three basic categories of performance information. A `trait-based` appraisal system assesses the abilities or other personal characteristics of an employee. A `behaviour-based` system measures the extent to which an employee engages in specific, relatively well-defined behaviours while on the job. Finally, a `results-based` system measures the "bottom line" associated with an employee's work: did the job get done, was a profit made? Each type of appraisal system has some advantages and disadvantages.

**Trait-based appraisals** are used to assess the personality or personal characteristics of employees, such as their ability to make decisions, loyalty to the company, communication skills, or level of initiative. This type of appraisal asks a lot about what a person is, but relatively little about what he or she actually does. While trait-based appraisals are easy to construct, the disadvantages of focusing on traits are significant.

Trait-based approaches have questionable validity. Traits assessed in this type of appraisal often do not relate well to how employees really behave on the job because job behaviour is strongly influenced by situational and environmental factors. An employee who is usually boisterous and aggressive with coworkers may act in a restrained, considerate manner when dealing with clients. Since the link between personal traits and actual job behaviour is weak, trait-

---

based systems are potentially more susceptible to charges of unfairness by minorities, women, and other protected groups. In addition, the interrater reliability of trait-based ratings is often low. The traits are difficult to define accurately, and the different frames of reference used by different raters make agreement among raters unlikely. A final disadvantage is that trait-based appraisals are not helpful for providing feedback to employees. For example, telling an employee that he or she is too shy is not very useful. The information does not indicate how the employee can be less shy.

In behaviour-based appraisals, employees are assessed on what they do on their job. Behaviour measures can be very useful for feedback purposes because they indicate exactly what an employee should do differently. Deficiency may be a problem with some behaviour-based appraisals since they often will not include all behaviours that could lead to job effectiveness.

To avoid the problems inherent in a behaviour-based approach, the result-based appraisals may measure the results of work behaviour. This approach deals with bottom-line issues such as how many cars an employee sold or how much profit the employee brought into the organisation during the month. When it is not important how results are achieved and when there are many different ways to succeed, a ‘results-based’ appraisal would be appropriate.

Despite their intuitive appeal, results-based appraisals pose questions of practicality, contamination, and deficiency. Results-based measures may be difficult to obtain for some jobs. Furthermore, results are not always under the control of an individual
employee. Equipment breakdowns, a change in the economy, bad luck, inadequate budgetary or staff support, or other factors not directly controlled by an employee may greatly affect the job results.\textsuperscript{1} Results-based measures are therefore contaminated by these external factors. Another problem is that results-based appraisals may foster a "results at all cost" mentality among employees.\textsuperscript{2} Teamwork among employees may suffer if individuals are preoccupied with their own personal results and will not take the time to help coworkers. A final disadvantage is that results-based appraisals are less helpful for employee development. Although they may indicate that results are below acceptable levels, they do not always provide clear information on how to improve work performance.

**Sources (Raters) of Performance Appraisal**

Performance appraisals can come from a variety of sources. Although most performance appraisals come from one source, such as the immediate supervisor, this is not necessarily the case. In matrix-type organisations, for example, employees most often have two immediate supervisors and receive ratings from both of them. Further, having more than one rater can increase the reliability of the performance evaluation. The most prominent sources of appraisal found in appraisal literature include: supervisor, co-workers, self, subordinates, computers, customers, and the job itself.

**Supervisor:** The most common evaluator is the employee's immediate supervisor. It has been estimated that over 90 per cent of


all performance appraisals are completed only by the immediate supervisor.¹

Co-workers (Peers): Another possible constituency from which to draw performance information is the employee's peer group or co-workers. Although coworkers may be somewhat uncomfortable and resistant to evaluating their peers, at least one study has shown that peer evaluations are more stable over time and may be the most accurate evaluations of employee performance.² An organisation may choose to encourage peer evaluations, particularly if the contact between supervisor and employee is limited.

Self-appraisal: Although many are reluctant to engage in self-ratings, this information can be extremely valuable to the supervisor. Large discrepancies between the supervisor and employee ratings should be the reason for concern. Discrepancies often occur due to a lack of communication and performance feedback by the supervisor. For example, if employees hear nothing from their supervisor for months, they might erroneously conclude that they are performing well (or poorly). However, when rating comparisons indicate a discrepancy, this information can be used to convince the manager to increase the amount of feedback the employee receives in the future.

Research has found that employees who are given the opportunity to evaluate themselves have a tendency to inflate their


ratings. One study reported that mean self-ratings for a group of workers were at least one standard deviation higher than were the mean supervisory ratings. Further, the range of ratings given by the employees for their performance was much smaller than those provided by their supervisors.

Self-appraisals have been found to be extremely helpful when used for employee development purposes. Self-ratings can encourage discussions about the employee's strengths and weaknesses from both the employee's and supervisor's point of view.

**Subordinates:** Subordinates are valuable sources of information when examining managerial performance. It is important to encourage subordinates to be candid if the information is to be at all useful. Candor, of course, is most likely to occur when subordinates are guaranteed anonymity and have no fear of reprisal. Information from subordinates is not only helpful for determining how well a manager leads, communicates, plans, delegates, and organises, but also for identifying general problem areas within a department.

---


Computers: Computer-aided management involves the use of computers to monitor, supervise, and evaluate employee performance electronically. Computer monitoring must include benefits for both the employees and the employers if it is to be effective. Computer monitoring can work both effectively and efficiently due to the following reasons:

1) **Response-outcome dependency**: Computer monitoring indicates to the worker which responses are correct. In turn, this can lead to an increase in the performance of desired behaviours. It also makes undesired behaviours salient.

2) **Effective feedback vehicle**: Summary reports from the computer can provide immediate feedback to employees so that they can choose to modify their behaviour.

3) **Constructive expectations**: Electronic performance monitoring sets specific standards accompanied by expectations that are incorporated into the daily work routines. Workers have a clear idea of what is expected of them.

4) **Reduced unpredictability**: Computer monitoring allows the employee to track his or her performance throughout the year. Thus, the yearly performance evaluation will be less of a surprise.

5) **Direct accountability**: Performance monitoring by computer is a direct measurement of that work, making it difficult for an employee to cover up errors or to blame others. Thus, those employees who are good performers would be the most likely to accept electronic monitoring.

6) **Better training programmes**: Gathering summary information regarding the types of mistakes most commonly made could lead to training programmes that precisely target the problems.
(7) **Objective documentation:** Electronic monitoring can identify good performers objectively because the computer generates a quantitative appraisal.

(8) **Increased flexibility:** Flexibility can occur because “fast” workers will have more control over scheduling of required work. “Slower” workers can be allowed some additional time to complete the tasks if the computer information shows that the worker output is adequate.

Monitoring employees by computer is open to serious invasions-of-privacy issues. Some lawmakers are already attempting to introduce legislation that would limit the use of computer monitoring in organisations.¹

**Customers:** In jobs that require a high degree of interaction with the public, appraisal sometimes can be done by the “consumers” of the organisation’s services. While the clients served cannot be expected to identify completely with the organisation’s objectives, they can, nevertheless, provide useful information.

Sometimes clients will rate an organisation as a whole. Thus when 1,100 black, Latin, and non-Latin white residents of Miami rated police performance on eleven dimensions (e.g., quick response, making streets safe, recovering stolen property), the police department learnt a great deal about public perceptions of police performance.²


² Jacobs, S. 1976. “Residents Rate Police As Satisfactory”. Miami Herald, November 21: 1D, 12D.
Alternatively, clients might rate particular individuals with whom they interact regularly (e.g., purchasing managers, sales representatives). Such information may affect personnel decisions (promotion, transfer, need for training), but it also can be used in personnel research (e.g., as a criterion in validation studies or in measuring training outcomes on the job), or as a basis for self-development activities.

**The Job itself:** Finally, employees at all levels in the organisation can receive feedback from the job they perform. For example, when a secretary mistakenly presses the wrong button on his/her word processor, a beef is heard. This negative feedback is a reminder that the last behaviour was inappropriate. In order to use this type of performance feedback successfully, organisations must carefully design the jobs their employees perform.

**360-degree Appraisal and Feedback**

360-degree appraisal is also known as multi-rater and multi-source assessment. It is designed to give a complete, multi-dimensional picture of an individual’s performance – from his/her immediate superior, peers, self, subordinates, and customers. Supervisors, peers, subordinates, customers, and employees themselves typically differ in their ability to appraise various dimensions of performance. These raters observe different behaviours and may interpret them with divergent standards. Thus each source has access to unique information about performance.1

---

To assess the degree of inter-rater agreement within rating dimensions (convergent validity), and to assess the ability of raters to make distinctions in performance across dimensions (discriminant validity), a matrix listing dimensions as rows and raters as columns might be prepared.\(^1\) Multiple raters for the same individual may be drawn from different organisational levels, and they probably observe different facets of a ratee's job performance. Hence, across-organisational-level inter-rater agreement for ratings on all performance dimensions is not only an unduly severe expectation; it also may be erroneous.

As a more reasonable alternative, Borman (1974) suggested a hybrid multi-trait-multi-rater analysis, in which raters make evaluations only on those dimensions that they are in good position to rate.\(^2\) In the hybrid analysis, within-level inter-rater agreement is taken as an index of convergent validity. The hybrid matrix provides an improved conceptual fit for analysing performance ratings, and the probability of obtaining convergent and discriminant validity is probably higher for this method than for the multi-trait-multi-rater analysis.

Although the hybrid multi-trait-multi-rater technique provides a method for measuring the convergent and discriminant validity of performance ratings, it does not address directly the extent to which different levels of raters agree with objective indices of performance. It may be that important aspects of job performance are not being captured in subjective ratings.


Zammuto et al., (1982) opine that just as contextual differences may cause raters from different levels to place different emphasis on specific aspects of job performance, organisational and work unit differences (e.g., in climate, policies, tasks, and functions) also affect the nature of appraisal.1

Typically, 360-degree appraisals are used for development and feedback purposes.2 The focus is on evaluation of competencies that are relevant for job performance in useful behavioural terms. Although 360-degree appraisals are being used by many firms, they are most successful in organisations that offer open and participative climates, along with active career development systems.3

Performance Appraisal Methods

Performance appraisal systems can be classified either as relative or absolute. Within this taxonomy, several performance appraisal methods exist. Rank order, paired comparisons, and the forced distribution belong to the relative rating (employee comparisons) systems. The prominent performance appraisal methods in the category of absolute rating systems include: narrative essay, behavioural checklists, forced-choice, critical incidents, graphic rating scales, mixed standard scales, behaviourally anchored rating scales (BARS), behavioural observation scales, management by objectives (MBO), work standards approach, and assessment centers.


Relative Rating Systems (Employee Comparisons)

(a) **Rank order** methods include two variations - simple, and alternation ranking. 'Simple ranking' requires only that a rater order all ratees from highest to lowest, from "best" employee to "worst" employee. On the other hand, 'alternation ranking' requires that the rater initially lists all ratees on a sheet of paper. From this list the rater first chooses the best ratee (#1), then the worst ratee (#n), then the second best (#2), then the second worst (#n-1), and so forth, alternating from the top to the bottom of the list until all ratees have been ranked.

(b) **Paired comparisons:** It should be noted that both simple and alternation ranking implicitly require a rater to compare each ratee with every other ratee, but systematic ratee-to-ratee comparison is not a built-in feature of these methods. This lacuna is overcome by 'paired comparisons'. The number of pairs of ratees to be compared may be calculated from the formula \[ \frac{n(n-1)}{2} \], wherein 'n' is the number of persons to be rated. Hence if 10 individuals were being compared, \( \frac{10(9)}{2} \) then 45 comparisons would be required. The rater's task is simply to choose the better of each pair, and each individual's rank is determined by counting the number of times he or she was rated superior.

Employee comparison methods are easy to explain and helpful in making personnel decisions.¹ Like other systems, however, they suffer from several weaknesses that should be recognised. Employees usually are compared only in terms of a single overall suitability

---

category. The rankings, therefore, lack behavioural specificity, and may be subject to legal challenge. Moreover, it is often impossible to compare rankings across work groups, departments, or locations. Reliability also may suffer, for when asked to re-rank all individuals at a later date, the extreme high or low rankings probably will remain stable, but the rankings in the middle of the scale may shift around considerably.

Absolute Rating Systems

(a) Narrative essay: Absolute rating systems enable a rater to describe a ratee without making direct reference to other ratees. Perhaps the simplest absolute rating system is the `narrative essay', in which the rater is asked to describe, in writing, an individual's strengths, weaknesses, and potential, together with suggestions for improvement. The assumption underlying this approach is that a candid statement from a rater who is knowledgeable of a ratee's performance is just as valid as more formal and more complicated appraisal methods.

The major advantage of narrative essays (when they are done well) is that they can provide detailed feedback to ratees regarding their performance. On the other hand, essays are almost totally unstructured and they vary widely in length and content. Comparisons across individuals, groups, or departments are virtually impossible since different essays touch on different aspects of ratee performance or personal qualifications. Finally, essays provide only qualitative information; yet in order for the appraisals to serve as criteria or to be compared objectively and ranked for the purpose of a personnel decision, some form of rating that can be quantified is essential. Behavioural checklists provide one such scheme.
(b) Behavioural checklists: When using a behavioural checklist, the rater is provided with a series of descriptive statements of job-related behaviour. The rater's task is simply to indicate ("check") statements that describe the ratee in question. In this approach raters are not so much evaluators as they are reporters of job behaviour. Moreover, ratings that are descriptive are likely to be higher in reliability than ratings that are evaluative.1

To be sure, some job behaviours are more desirable than others; checklist items can, therefore, be scaled by using attitude scale construction methods. In one such method, the Likert method of summated ratings, a declarative statement is followed by several response categories, such as "always", "very often," "fairly often", "occasionally," and "never". The rater simply checks the response category he or she feels best describes the ratee. Each response category is weighted – for example, from 5 ("always") to 1 ("never") – if the statement describes desirable behaviour, or vice versa if the statement describes undesirable behaviour. An overall numerical rating for each individual then can be derived by summing the weights of the responses that were checked for each item, and scores for each performance dimension can be obtained by using item analysis procedures.2

(c) Forced-choice systems: A special type of behavioural checklist is known as the 'forced-choice system' – a technique developed specifically to reduce leniency errors and to establish objective

---


standards of comparison between individuals. In order to accomplish this, checklist statements are arranged in groups, from which the rater chooses statements that are most or least descriptive of the ratee. An overall rating (score) for each individual is then derived by applying a special scoring key to the rater’s descriptions.

Forced-choice scales are constructed according to two statistical properties of the checklist items: (1) *discriminability*, a measure of the degree to which an item differentiates effective from ineffective workers, and (2) *preference*, an index of the degree to which the quality expressed in an item is valued (i.e., is socially desirable) by people. The rationale of the forced-choice system requires that items be paired so that they appear equally attractive (socially desirable) to the rater. Theoretically, then, the selection of any single item in a pair should be based solely upon the item’s discriminating power, not its social desirability.

The main advantage claimed for forced-choice scales is that a rater cannot distort a person’s ratings higher or lower than is warranted since he or she has no way of knowing which statements to check in order to do so. Hence, leniency should theoretically be reduced.

Forced-choice scales usually give only a single, global indication of merit, rather than ratings of specific dimensions of performance. Sets of items are usually not developed for specific dimensions because a great deal of time and effort is required to write and scale items and then to develop sets that: (a) are matched

---

on social desirability, and (b) discriminate adequately among individuals.

There are two further problems with forced-choice scales. The most serious is rater resistance. Since control is removed from the rater, he or she cannot be sure just how he/she rated a subordinate. Finally, forced-choice forms are of little use (and may even have a negative effect) in performance appraisal interviews, for the rater is unaware of the scale values of the items he or she chooses. To minimise these weaknesses, it is probably best to limit the use of forced-choice scales to administrative decision making, and not to try to use them as vehicles for personal development.

(d) Critical Incidents: This method of performance appraisal has generated a great deal of interest in recent years, and several variations of the basic idea are currently in use. As described by Flanagan (1954), the critical requirements of a job are those behaviours that make a crucial difference between doing a job effectively and doing it ineffectively. Critical incidents' are simply reports by knowledgeable observers of things employees did that were especially effective or ineffective in accomplishing parts of their jobs. Critical incidents are recorded for each employee by supervisors, as they occur. Thus, they provide a behaviourally based starting point for appraising performance.

The critical incidents method looks like a natural for performance appraisal interviews because supervisors can focus on actual job behaviour rather than on vaguely defined traits.

Performance, not personality, is being judged. Ratees receive meaningful feedback and they can see what changes in their job behaviour will be necessary in order for them to improve. In addition, when a large number of critical incidents are collected, abstracted, and categorised, they can provide a rich storehouse of information about job and organisational problems in general and are particularly well suited for establishing objectives for training programmes.\(^1\)

As with other approaches to performance appraisal, the critical incidents method also has its drawbacks. First of all, it is time-consuming and burdensome for supervisors to record incidents for all of their subordinates on a daily or even weekly basis. Feedback may, therefore, be delayed. Moreover, the rater sets the standards by which subordinates are judged; yet subordinates are likely to be more motivated if they have some say in setting their own behavioural standards.\(^2\) Finally, in their narrative form, incidents do not readily lend themselves to quantification, which, as noted earlier, poses problems in between individual and between-group comparisons as well as in statistical analyses.

(e) Graphic rating scales: Probably the most widely used method of performance appraisal is the ‘graphic rating scale’. In terms of the amount of structure provided, the scales differ in three ways: (1) the degree to which the meaning of the response categories is defined, (2) the degree to which the individual who is interpreting the ratings

---


can tell clearly what response was intended, and (3) the degree to which the performance dimension being rated is defined for the rater.

It should be noted that graphic rating scales may not yield the depth of narrative essays or critical incidents, but they: (a) are less time-consuming to develop and administer, (b) permit quantitative results to be determined, (c) promote consideration of more than one performance dimension, and (d) are standardised and therefore comparable across individuals. On the other hand, graphic rating scales give maximum control to the rater, thereby exercising no control over rating errors (leniency, severity, central tendency, or halo). For this reason, they have been criticised widely. Nevertheless, when simple graphic rating scales have been compared against more sophisticated forced-choice ratings, the graphic scale consistently proved just as reliable and valid\(^1\) and was more acceptable to raters.\(^2\) At this juncture, it is relevant to consider three interesting variations of graphic rating scales – mixed standard scales, behaviourally anchored rating scales, and behavioural observation scales.

(f) Mixed Standard Scales: These are designed specifically to minimise halo and leniency biases and to permit evaluations of the reliability with which each individual is rated, each scale rates, and each rater rates.\(^3\) The procedure is as follows. First, items that discriminate effective from ineffective performance are obtained from

---


knowledgeable persons (usually supervisors). For each performance dimension to be rated, three items are chosen that represent good, average, and poor performance, respectively. These are "standards" or degrees of performance, and the rater must respond to each standard, indicating whether he/she considers the ratee to be better than the description (+), to fit the description (o), or to be worse than the description (-).

The performance dimensions and the standards describing them are then mixed randomly so that no clear order-of-merit descriptions exist for each dimension.

The rationale for mixing standards stems directly from the logic of forced-choice formats – namely, that halo and leniency are likely to be reduced if ratings are not made on a scale with statements arranged in an obvious order-of-merit hierarchy. Empirical findings lend support to this hypothesis.\(^1\) In addition, raters tend to use all seven scale points. A final advantage of mixed standard scales, and one that is not to be taken lightly, is that they are capable of providing indices of the extent of error for different raters and ratees and for different scales.

**(g) Behaviourally anchored rating scales (BARS):** So far, it seems that the performance appraisal systems that are most acceptable to raters (e.g., narrative essays, employee comparisons, graphic scales) are often least acceptable in terms of their psychometric characteristics. Conversely, those that are psychometrically

acceptable (e.g., forced-choice scales) often are resisted by raters. With reference to this resulting dilemma, Smith and Kendall (1963, p. 151) state:

Better ratings can be obtained, in our opinion, not by trying to trick the rater (as in forced-choice scales) but by helping him to rate. We should ask him questions which he can honestly answer about behaviours which he can observe. We should reassure him that his answers will not be misinterpreted, and we should provide a basis by which he and others can check his answers.¹

BARS are rooted in, and referable to, actual job behaviour, not vaguely defined personality traits. A second advantage may be attributed to an aspect of the developmental procedure that is unrelated to the measurement of performance per se – namely, intense rater-ratee involvement throughout all phases of BARS development. High participation serves a dual purpose: (1) it ensures acceptability of the resulting scales and a commitment to make them work, and (2) it guarantees that the scales are relevant to the jobs in question. Performance dimensions and behavioural anchors are defined rigorously and can be distinguished from one another easily and accurately. Moreover, the scales are phrased in the jargon of the users, not the jargon of the psychologist or scale developer. Finally, even though specific behaviours may be observed in different

situations, they are referred to a common set of expectations that serves as a mutual frame of reference.¹

An enormous amount of research on BARS has accumulated over the past decade, and much of it has been summarised by Bernardin and Beatty (1984).² In short, there is little empirical evidence to support the superiority of BARS over other performance appraisal systems. BARS development is a long, painstaking process that may require many individuals. Moreover, separate BARS must be developed for dissimilar jobs. Consequently, this approach may not be practical for many organisations.

(h) Behavioural observation scales (BOS): The basic BARS approach to scale development is shared by BOS.³ Like BARS, BOS consist of clusters of behavioural statements (based on critical incidents) that all describe the same aspect of performance (e.g., "inventory control"). Unlike BARS, however, each statement in a BOS represents a high level of performance. The rater’s task is simply to estimate the frequency with which he or she has observed each behaviour. Psychometrically BOS are similar to BARS in all other respects.⁴

---


At this juncture it is important to stress that rating formats are only one aspect of the larger issue of performance management. As Jacobs et al. (1980, p. 635) pointed out:

What is desperately needed is a more far reaching approach to performance evaluation, one which includes scale construction and utilisation as a starting point, incorporates performances feedback, and culminates in organisational efforts toward training, selection, placement, and promotion. In this light BARS has a future. To continue the current research applications is to uncover nothing new and to do a disservice to the methodology specifically and the area of performance evaluation in general.¹

The reason why formats alone cannot produce valid appraisals is that they cannot specify everything a rater needs to know or consider in order to rate accurately. Unless observation of ratees is extensive and representative, judgements are unlikely to represent a ratee's true performance. Since the rater must make inferences about performance, the appraisal is subject to all the biases that have been linked to rating scales. Raters are free to distort their appraisals to suit their purposes. This can undo all of the painstaking work that went into scale development, and probably explains why no single rating format has been shown to be superior to others.

In this connection, Banks and Roberson (1985) suggest two strategies.² One, build in as much structure as possible in order to


minimise the amount of discretion exercised by a rater. For example, use job analysis to specify what is really relevant to effective job performance, and critical incidents to specify levels of performance effectiveness in terms of actual job behaviour. Two, do not require raters to make judgements that they are not competent to make; do not tax their abilities beyond what they can do accurately. For example, for formats that require judgements of frequency, make sure that raters have had sufficient opportunity to observe ratees so that their judgements are accurate. Above all, recognise that the process of performance appraisal, not the mechanics, determines the overall effectiveness of this essential organisational activity.

(i) Management by Objectives (MBO): The term 'management by objectives' was coined in 1954 by Drucker who suggested that "objectives are needed in every area where performance and results directly and vitally affect the survival and prosperity of the business,"¹ a concept which has received an almost unparalleled level of international acclamation and acceptance in both public and private sector institutions.

At a time when managers, beset by mounting economic, technological, social and political pressures, were desperately seeking better ways of running their enterprises, Drucker’s concept, encapsulating both an attractive managerial philosophy and a viable, commonsense modus operandi of managing, was seized upon by those seeking the holy grail of the ultimate managerial panacea, only to find that it was just another idea, better than most perhaps, but no instant cure-all for managerial problems.

Partly because of such disappointments and partly as a result of close and detailed scrutiny of the reasons for the success or failure of MBO in different circumstances and situations, it has become somewhat fashionable to knock the technique. Many academics and practising managers have jumped on this particular bandwagon, some regrettably coming perilously close to throwing out the baby with the bath water. To obtain a balanced perspective of the subject much depends upon how one regards MBO and what one thinks it really is and can do. As Reif and Bassford (1979) point out, "it is disconcerting to discover how many versions . . . there are of the concept" - a way of setting objectives; a method of performance appraisal and review; a scheme of management development; a technique for reward distribution; or even a form of manpower planning! "MBO" they say, "is none of these, or, more appropriately, is all of these – and more",¹ pointing out that inadequate definition of the concept of MBO and concomitant practice of only one part of it is a common error.

The total concept of MBO comprises both philosophy and methodology. The philosophy, according to Bass and Ryterband (1979) is: "what people want most from work is to succeed at clear and personally valued activities, so concentrate on getting employees to understand the objectives of their particular assignments."² MBO is thus a process founded on a philosophy of participation with personnel heavily involved in – and, through involvement, committed to – the determination of compatible institutional and personal goals.


It is a process, according to Steers and Porter (1979), "by which employees of an organisation, working together, identify common objectives . . . and coordinate their efforts towards goal attainment. The major focus in this "process", they maintain, "is on the future – where the organisation wants to go – not in the past"; objectives being established to facilitate planning and coordination of available resources to maximise the opportunities and minimise the threats in the institution's environment.¹

MBO has several advantages. Observable, results-oriented goals replace the difficult-to-observe, subjective traits that are sometimes used in other measures of performance. MBO can encourage innovation and creativity, since employees are free to determine how they will meet their goals. This approach can also lead to performance improvement, particularly if goals are specific, moderately difficult, and accepted by the employees. A "coordination advantage" can also result: top managers set goals with those immediately below them in the hierarchy, who in turn help set the goals of those below them. This cascading effect may help coordinate the activities of the entire organisation, with all employees engaged in activities related to the organisation's overall objectives and strategy.

Despite the benefits of the MBO approach to performance appraisal and review, even its most committed devotees will concede that it has had its failures. Muezyk (1979), whose penetrating critique gives a detailed and valuable insight into the more important grounds

for disquiet with the technique, holds that “evidence is mounting that MBO has more clear failures than successes,” going so far as to suggest that “we may be approaching an era where more time is spent taking MBO out than putting it in.”¹ Commenting that many organisational adjustments are necessary if a programme of MBO is to be successful, Connor (1980) observes that “there is a tendency for managers to direct their efforts toward meeting only those objectives on which they are measured,” hence other “less quantifiable objectives may be short-changed.”²

He points out that under conditions of rapid change or environmental turbulence it is often quite difficult to set up clearly definable objectives, adding:

There are many difficulties tying performance appraisal into the programme . . . under uncertain and adverse conditions. Executives of companies that have established a management by objective programme declare that the programme worked very well when environmental and competitive forces were favourable. However, during a business downturn, many managers thought they were held strictly accountable for accomplishing objectives over which they had limited control. Their programmes did not adequately respond to new environmental conditions and instead of being a


basis for positive motivation . . . became a source of major conflict.¹

If an MBO system is to be successful, several requirements must be met. First, objectives should be quantifiable and measurable; objectives whose attainment cannot be measured or at least verified should be avoided where possible. Objectives should also be challenging and yet achievable, and they should be expressed in writing and in clear, concise, unambiguous language. MBO also requires that employees participate in the objective-setting process. Active participation by the employee is also essential in developing the action plan. A final requirement for the successful use of MBO is that the objectives and action plan must serve as a basis for regular discussions between the manager and the employee concerning the employee's performance. These regular discussions provide an opportunity for the manager and employee to discuss progress and to modify objectives when necessary.

(jj) The 'work standards approach' to performance appraisal is most frequently used for production employees and is basically a form of goal-setting for these employees. It involves setting a standard or expected level of output and then comparing each employee's performance to the standard. Generally speaking, work standards should reflect the normal output of a normal person. Work standards attempt to define a fair day's output. Several methods can be used for setting work standards, such as average production of work groups, performance of specially selected individuals, time study, work sampling, expert opinion.

¹ Connor, P.E. 1980. op. cit.
An advantage of the work standards approach is that the performance review is based on highly objective factors. Of course, to be effective, the standards must be viewed by the affected employees as being fair. The most serious criticism of work standards is lack of comparability of standards for different job categories.

(k) Potential appraisal through assessment centres: It is now widely accepted that while appraisal of performance against previously agreed goals does produce useful and reasonably accurate assessment of how well, or otherwise, an employee is doing in his/her present job, it has much less relevance when it comes to attempting to predict how he/she will behave in a future promoted post. Stewart and Stewart (1979) argue that “predicting whether someone is going to succeed after a promotion is much more than an academic exercise,” adding:

Saying whether someone will perform well in a new job is a little like examining his tracks after he has been running along a beach. You can look back and infer the speed and direction of running. (Personal history forms and performance appraisal records help you do this). Detailed visual inspection reveals a pair of pink hairy feet in the last set of footprints. (Perhaps your appraisal system singles out ‘present performance’ for comment. Or you can take up references). From this information you can work out how far he will have to run if he keeps up his present speed and direction; so that if tomorrow’s assignment is much the same as
yesterday's, the problem of predicting potential is a simple matter of applied archaeology.

But when you are asking a salesman to become sales manager, or you are sending out a general manager to open up a new operation in Africa, or you are asking a research chemist to become the head of R & D, you are not asking him to continue at the same speed and direction as before; may be you want to take him off the beach altogether and see how he behaves when he is given a pile of rocks to climb up. In these circumstances his behaviour on the beach may be of limited relevance, or even positively misleading.¹

Over the years a number of well-known (and in some instances well-validated) techniques have been devised or developed in continuing attempts to solve the problems of assessing potential. There is a large number of psychological instruments designed to test human attributes such as intelligence, aptitudes, abilities, attitudes, and even personality. Properly used by competent practitioners who are skilled in selecting the appropriate tests for particular purposes, skilled in their administration, and above all, in interpreting the results (usually in conjunction with data from other tests in a battery, personal interviews, and individual histories) well validated psychological tests are useful aids to decision-making on potential.

From an historical standpoint, the first industrial assessment centres, notably those set up by the American Telephone and Telegraph (AT&T) company in the 1950s, were precisely off the job locations in which programmes designed to identify managerial potential through the simultaneous evaluation of several individuals by a number of trained observers were carried out.

Nowadays, however, the term *management assessment programme* is more accurate since the so-called assessment centre is, in fact, a process which uses a variety of evaluative techniques such as simulation exercises, psychological tests, management games, group discussions, individual and group interviews, and feedback sessions aimed at improving the selection and development of managers. As such it may or may not be conducted off the job. It is the procedure and not the location in which it is carried out that is important. While detailed procedures vary widely according to the aims of particular programmes and the needs of the participants and their companies, all are based on the common assumption that the demands likely to be made on individuals by future work situations can be simulated with sufficient accuracy to enable trained observers to judge how well they are likely to be able to cope with them.

In summary, it can be said that there is no one best method of conducting performance appraisal. Before choosing the methods of performance appraisals, human resources professionals should examine two areas with special care: (1) various factors (leadership style, organisational climate, training in appraisal procedures) that can help or hinder the implementation of a particular appraisal programme, and (2) the appropriateness of the appraisal method for the specific jobs to which the appraisal system will apply.
Performance Appraisal Feedback

One of the potentially most important uses of appraisals is providing performance feedback to employees. Most theories of work motivation point out that before employees can improve their performance they must know how well they are currently doing. The performance feedback interview is a discussion between the supervisor and the employee concerning the employee's past performance and how that performance can be improved in the future. Appraisal interviews normally cover the following aspects:

- How effectively the appraisee feels that he/she has performed in his/her key result areas.
- What barriers to more effective performance arose during the period under review and how these were dealt with.
- Areas of work in which he/she feels that he/she has done particularly well or especially badly.
- How he/she feels that the job in general and, in particular, those parts of it which he/she likes and/or dislikes most, and why.
- Any parts of the job about which he/she has feelings of uncertainty or ambivalence.
- What help he/she feels he/she needs from his/her boss and/or the organisation in order to improve his/her performance.
- His/her ambitions for future career development and what steps he/she is taking to achieve these, also what help he/she would like from his/her boss and/or the organisation in this respect.

There are three main approaches to performance feedback discussion, often referred to as (1) tell and sell, (2) tell and listen, and (3) problem-solving.
In a **tell and sell feedback interview**, the supervisor tells the employee how good or bad the employee’s performance has been and then attempts to persuade the employee to accept this judgement. The employee has no input into the evaluation. Because it is very directive and one-sided, the tell and sell interview can lead to defensiveness, resentment, and frustration on the part of the subordinate. For new employees or those who have little desire for participation, the tell-and-sell interview can be effective in providing feedback and improving performance. New employees often feel unqualified to judge their own performance and prefer to be told how they are doing and what is expected of them.¹

In the **tell and listen** approach, the supervisor tells the employee what has been right and wrong with the employee’s past performance but then gives the employee a chance to react. There is evidence that subordinates prefer even very limited participation to none at all.²

In a **problem solving interview** employees evaluate their own performance and set their own goals for future performance. The supervisor is primarily a helper and colleague rather than a judge, and offers observations and advice in a noncritical manner. An active and open dialogue ensues, in which goals for improvement are established mutually. The problem solving interview is more difficult for the supervisor than the other types of interviews, but it is


more likely to result in employee acceptance and commitment to the established goals.

To by-pass some of the difficulties associated with the tell-and-sell and problem-solving interviews, a mixed approach has been recommended. The mixed interview uses the tell-and-sell approach to communicate administrative decisions, whereas the problem-solving approach is reserved for discussing employee development issues and planning for future performance.

Employees who perform poorly use impression management strategies to minimise the amount of negative feedback they receive. Recent research has emphasised the role of impression management in the feedback process.¹ 'Impression management' refers to behaviours by an employee designed to control how he or she appears to the supervisor. For example, impression management tactics include (1) taking sole credit for positive events even when credit should be shared, (2) making the supervisor aware of self-accomplishments, (3) arriving early or staying late to give the impression of being a hard worker, or (4) taking an interest in the supervisor's personal life and perhaps doing personal favours.²

---


Impression management tends to skew the appraisal and feedback interaction.¹

Feedback is most effective in improving performance when it is specific, when it is accepted by the employee, and when it helps define clear goals for future performance.² Research has shown that subordinates prefer specific feedback in the appraisal interview, even if it concerns poor performance, to appraisals that contain only vague performance information.³ For feedback to be accepted by a subordinate, it must come from a credible, trustworthy source. Credibility is enhanced when the evaluator is seen as being knowledgeable about the subordinate’s job, has had adequate opportunity to observe the behaviour of the subordinate, and has clearly taken the time to prepare carefully for the appraisal interview. Research has shown that feedback is more likely to affect subsequent performance when it comes from a believable source.⁴


Rating Errors

Despite the fact that one would like to limit the role of the rater to that of observer-recorder, the fact is that the rater is really an evaluator-judge.¹ Making judgements is simply an inextricable part of managing. The task is complicated, though, when a rater is asked to recall information on several subordinates over the past six or twelve months. Human memory is fallible, and biased ratings can be expected if information was stored in a rater's memory according to irrelevant, oversimplistic, or otherwise faulty categories. Unfortunately, faulty categorisation seems to be the rule rather than the exception.² Research is currently under way in an effort to develop strategies for improving the accuracy of ratings by improving the process of human judgement. Although this research may improve raters' ability to judge accurately, it does not necessarily address the rater's willingness to provide accurate ratings.³ Rating errors occur through both unintentional and intentional processes.

One way to examine why rating errors occur is to use the approach called 'cognitive information processing (CIP)'. Proponents of CIP argue that performance appraisal is best viewed as a complex memory task in which the assessors must (1) acquire performance information about an employee; (2) encode and store that information in their memory; (3) retrieve that information at a later date when asked to assess the employee's performance, and


(4) weigh and combine the information into an overall performance judgement.¹

The amount of performance information available about an employee is enormous. Raters can notice and recall only a limited amount of information on each employee, and research has shown that they use a variety of methods to condense the information they do receive.² This can result in a variety of memory and judgement errors in appraisal. One cognitive short-cut that people use in processing information about others is the schema. "Schemas" are simply mental categories that an individual uses to organise information and classify people.³ For example, a supervisor might use two schemas in classifying employees: the good worker schema and the bad worker schema. Associated with each schema is a set of attributes, called a `prototype’, that represents the essential characteristics needed to be classified into that schema.⁴ Thus, the prototype for the supervisor’s good worker schema might include the following: never absent, writes well, always gets projects done on time, and gets along well with coworkers. A ratee need not exhibit all the prototype characteristics in order to be classified into the schema. For instance, noticing that an employee is never absent from work,


the supervisor might categorise that person as a good worker, even without direct evidence that the employee writes well, hands in projects on time, or gets along well with coworkers.

When asked to assess the performance of an employee, a supervisor searches his/her memory to determine into which schema the employee has been placed. The supervisor recalls information about the employee, including the schema prototype. The supervisor's judgement of the employee is based not only on the precise behaviours observed, but also on the prototype associated with the schema. Some of the commonly found rating errors in performance appraisal literature include: halo, stereotypes, leniency and severity, central tendency, recency effects, and attributions.

**Halo:** Halo bias is perhaps the most pervasive bias in performance appraisal.\(^1\) A rater who is subject to the halo bias assigns ratings on the basis of a global impression of the ratee. An individual is rated either high or low on many factors because the rater knows (or thinks he/she knows) that the individual is high or low on some specific factor. The rater fails to distinguish among levels of performance on different performance dimensions.

Halo bias is doggedly resistant to extinction. In an effort at least to minimise it, numerous rating methods, instrument formats, rater training techniques, and statistical controls have been developed.\(^2\) An assumption that underlies many of these strategies is

---


that increased observation of performance-relevant ratee behaviour will reduce halo bias and thus improve the validity of subsequent ratings. But it is not true. Evidence indicates that when raters have a greater opportunity to observe ratee’s performance and thus to become more familiar with the behaviour to be rated, halo actually increases.\footnote{Jacobs, R., and Kazlowksi, S.W.J. 1985. A Closer Look at Halo Error in Performance Ratings. \textit{Academy of Management Journal}, 28: 201-212.}

**Stereotypes:** Stereotyping occurs when the rater places an employee into a class or category based on one or a few traits or characteristics. For example, an older employee may be stereotyped as being slower, more difficult to train, and unwilling to learn new approaches. Obviously, this perceptual error could negatively affect the overall performance evaluation. Of course, the older employee being evaluated may not fit this stereotype at all and may be quick to pick up new concepts and anxious to participate in new training programmes. Some research has indicated that the composition of the group from which the stereotyped employee comes may influence whether or not the stereotype influences the rating. For example, women received lower ratings when the proportion of women in the group of employees was small; however, the stereotype did not lower ratings for blacks who came from a group that had more whites than blacks. These results suggest that some stereotypes may be stronger than others.\footnote{Sackett, P.R., C.L.Z. DuBois, and A.W. Noe 1991. "Tokenism in Performance Evaluation: The Effects of Work Group Representation on Male-Female and White-Black Differences in Performance Ratings." \textit{Journal of Applied Psychology}, 76: 263-267.}

**Leniency and Severity:** The use of ratings rests on the assumption that the human observer is capable of some degree of precision and
some degree of objectivity.1 “Objectivity” is the major hitch in these assumptions, and it is the one most often violated. Raters subscribe to their own sets of assumptions (which may or may not be valid), and most people have encountered raters who seemed either inordinately easy (lenient) or inordinately difficult (severe). Leniency and severity (strictness) errors occur when the rater tends to use one of the extremes of a rating scale. When leniency errors occur, most employees receive very favourable ratings, even though it is not warranted by their performance. On the other hand, severity errors, which are basically the opposite of leniency errors, occur when the rater erroneously evaluates most employees unfavourably.

Leniency or severity biases can be controlled or eliminated in two ways: (1) by allocating ratings into a forced distribution, in which ratees are apportioned according to an approximately normal distribution; or (2) by reducing ambiguity in the rating scales themselves. This is done by improving the definitions of the dimensions and also by providing anchors (definitions) for the various scale points. The two alternatives differ in their theoretical underpinnings. Forced distribution removes a substantial amount of control from the rater by telling him or her, for example, that “10 per cent of your ratees must be rated poor (or outstanding)”. On the other hand, improvement in scale definitions and anchors is based on the notion that in order to minimise leniency or severity, the rater should be helped, as much as possible, by specifying clearly what is to be rated and what the various scale points mean. In this way, he or she will be better able to distinguish the various levels of

performance. No less important, however, are rater cooperation and interest in avoiding leniency, severity, and other biases.

**Central Tendency:** When political considerations predominate, raters may assign all their subordinates ratings that are neither too good nor too bad. They avoid using the high and low extremes of rating scales and tend to cluster all ratings about the centre of all scales. "Everybody is average" is one way of expressing the central tendency bias. The unfortunate consequence, as with leniency or severity biases, is that most of the value of systematic performance appraisal is lost. The ratings fail to discriminate either within or between people, and the ratings become virtually useless as managerial decision-making aids, as predictors, as criteria, or for giving feedback. Central tendency biases can be minimised by specifying clearly what the various anchors mean.

**Recency Effects:** Recency errors occur when performance is appraised based on performance information that occurred most recently. Essentially, supervisors rate the employee’s most recent behaviour. Recency errors are most likely to occur when there is a long period of time between performance and appraisals (such as a year).

**Attributions:** Another perceptual error that can affect the validity of the performance appraisal involves the attributions the rater makes about employee behaviour. Making an attribution means to assign causation for another’s behaviour.¹ For example, if a supervisor

attributes an employee's good performance to external causes, such as luck, holding an easy job, or receiving help from coworkers, then the performance appraisal will not be as positive as if the supervisor had attributed good performance to internal causes, such as effort or ability. Similarly, if the supervisor attributes poor performance to external causes rather than internal causes, the performance appraisal will not be as negative.

**Rater Training**

In the context of rater training there are three broad objectives: (1) to reduce or eliminate judgemental biases; (2) to improve the observational skills of raters by teaching them what to attend to; and (3) to improve the ability of raters to communicate appraisal information in an objective, constructive manner with ratees.

Traditionally, rater training has focused on teaching raters to eliminate judgemental biases such as leniency, central tendency, halo, and contrast effects. With respect to rater accuracy or validity, inaccurate performance appraisals may result from the way in which performance information is selected, attended to, and organised by the rater. Emphasis should be placed on teaching raters what to observe, rather than showing them "how to" or "how not to rate". In addition, however attention should be paid to matching examples of effective or ineffective performance to the exact job(s) for which appraisals will be made, not to a "generic" job. Doing so will enable

---


raters to store job-related information in mental categories that match actual job requirements.¹

Developing consensus among raters regarding performance standards and relative levels of effective or ineffective behaviour, encouraging raters to record specific examples of behaviour, providing an opportunity for raters to practice providing feedback and reinforcement where appropriate, and offering periodic refresher training to maintain desired rater behaviours may well be the keys to improving the appraisal process.

Several rater training programmes have been developed that aim to help evaluators produce reliable and accurate performance ratings. These programmes can generally be classified into three types: rater error training, frame-of-reference training, and information processing training.

People can be taught how to reduce rating errors, such as leniency, severity, central tendency, and halo errors.² In rater error training, evaluators view examples of the common errors and receive suggestions on how to avoid them. Several critics have argued that error reduction often comes at the expense of rating accuracy.³


Attempts have been made to reduce errors by developing a common frame of reference among raters for evaluating performance. Examples of actual employee behaviour in the organisation are used to develop norms for good and poor performance. Raters are then trained to observe these behaviours and use them as the standard against which to judge the performance of employees. In a review of the effectiveness of rater training programmes, four of seven studies using frame-of-reference training were reported to have reduced leniency error. Only three of ten studies using frame-of-reference training to reduce halo error reported positive results.

Some training efforts focus on how performance raters observe, store, recall, and use information. In a 1980 study, raters were trained to avoid eight different "errors of observation". Their accuracy in observing and remembering behaviour improved, but no measures of rating error or rating accuracy were taken.

Since rater error training and frame-of-reference training have had limited success, information processing approaches may be the most promising methods for improving rating accuracy. However, regardless of which training method is used, raters who are actively involved in their training programmes are more likely to learn to

---


reduce rating errors. Neither discussion nor lecture without practice and feedback has been shown to reduce rating errors.¹

**REVIEW OF LITERATURE**

Performance appraisal is an applied topic, and it is useful to periodically consider the current state of performance appraisal research and its relation to performance appraisal practice. An extensive review of appraisal research presented below 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. Major relevant earlier studies on various aspects of performance appraisal are briefly discussed in the following nine sections: (1) cognitive processing of information, (2) rater/ratee personal characteristics, (3) rating errors and accuracy, (4) appraisal sources, (5) appraisal feedback, (6) rater training, (7) performance appraisal formats, (8) legal ramifications, and (9) case studies and "how-to" articles.

**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)² call for incorporating non-student subjects and field settings in

---


research designs of cognitive process issues.\textsuperscript{1} Cognitive processing research concentrated around two issues: (a) how prior expectations or knowledge of prior performance levels affect the way information is processed, and (b) the role of memory in the rating process.

Rater’s knowledge of prior performance appears to affect information processing by framing or anchoring current judgements.\textsuperscript{2} Laboratory research indicated that knowledge of prior performance caused contrast effects (i.e., bias away from level of prior performance) rather than assimilation effects.\textsuperscript{3}

Additionally, Steiner and Rain (1989) reported that the order in which good and poor performance was observed affected performance ratings and that raters biased their judgement about inconsistent extreme performance (unusually good or poor) towards the general impression already held.\textsuperscript{4}

Rater’s 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 behaviours were either congruent or incongruent with prior expectations. Results indicated that when behaviour was congruent with expectations, appraisal results were more accurate.¹

Similarly, in a longitudinal study of 49 banking supervisor-subordinate dyads, Hogan (1987) found supervisors' expectations introduced error into the rating process, and that disconfirmation of prior expectations appeared to lower ratings.² However, consistency of ratee performance apparently affects rater ability to form general impressions and categorise information. Padgett and Ilgen (1989) demonstrated that consistent ratee performance led to greater use of categorisation but inconsistent performance led to greater retention of behavioural information.³

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.\(^1\) Laboratory research using student subjects and paper-people also suggested that job and ratee knowledge had significant effects on conceptual similarity and rating covariance as well as on halo.\(^2\)

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.\(^3\) Dipboye (1985) reported that overemphasis on cognitive determinants of performance ratings has led to neglect of behavioural, social, and affective determinants of bias in the rating process.\(^4\)

**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, and no gender differences were reported when rating familiar tasks in work situations where feedback was available. Conversely, students tended to give women professors higher ratings, 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. However, using both between and within-subject analysis 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


supervisor's causal attributions appeared to be related to subordinate's age.¹

A meta analysis of race effects confirmed that rater/ratee similarity results in higher ratings.² In studies of other personal characteristics, (a) emotional disability was shown to inflate ratings when clear professional standards were not present³; (b) attractiveness inflated ratings of non-managerial women, deflated ratings of managerial women, and had no effect on men⁴; and (c) the effects of rater-ratee acquaintance may depend on rating format.⁵ Rater affect also appeared to influence rating behaviour.⁶ However, Cardy and Dobbins (1986) suggested that affect influenced the rating not by increasing leniency but by introducing noise into the process.⁷

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.\(^1\) Nathan and Tippins (1990) reported a positive relationship between halo and accuracy,\(^2\) but Fisicaro (1988) concluded that a negative relationship exists.\(^3\) 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.\(^4\)

Schoorman (1988) reported that supervisors who had a say in the hiring decision and who viewed the applicant as favourable, subsequently tended to give lenient performance ratings, whereas those who participated in hiring but viewed the applicant as unfavourable, tended to give more severe ratings.\(^5\) 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.\textsuperscript{1}

\textbf{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.\textsuperscript{2} Conversely, Fox and Dinur (1988) reported low validity of self-ratings regardless of the expectation of validation.\textsuperscript{3} 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.\textsuperscript{4}

Vance, MacCallum, Coover, and Hedge (1988) reported that among a sample of jet engine mechanics, peer, self, and supervisory


ratings were equally valid sources\(^1\), but Fox, Ben-Nahum, and Yinon (1989) concluded that rating accuracy was positively related to rater-ratee similarity.\(^2\) Meta-analytic results suggested that only moderate relationships exist between self-supervisor and self-peer ratings.\(^3\)

**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.\(^4\) In contrast, Prince and Lawler (1986), reported that salary discussions during the appraisal interview had either no relationship or a positive relationship with future behaviour.\(^5\) However, Pearce and Porter (1986), reported that feedback describing an employee as "satisfactory" (as compared to above average or outstanding) led to reduced organisational


commitment and negative attitudes towards the performance appraisal system.¹

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.² This agrees with Bannister's (1986) experimental results concluding that source credibility and message content influenced recipient response to the feedback.³ In fact, Becker and Klimoski (1989) reported that feedback from supervisors led to increased performance but feedback from self and peers did not.⁴

Ilgen and Moore (1987) explored feedback content in a laboratory setting and found that feedback about quantity led to higher quantity, feedback about quality led to higher quality and feedback about both led to both.⁵ Message content apparently also affects rater cognition. Specifically, raters do not like to give negative


feedback\(^1\) and are likely to rely on scripts to deliver feedback about poor performance.\(^2\)

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, emphasising improvement, and discussing pay and advancement).\(^3\) Furthermore, Russell and Goode (1988) reported that satisfaction with feedback also may be multi-dimensional.\(^4\) 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**

Hedge and Kavanagh (1988) reported that training focused on minimising rating errors successfully reduced leniency and halo but also reduced accuracy. They concluded that the rater training should emphasise observation and decision-making processes rather than

---


simply error reduction.¹ 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.² A review of 24 rater training studies suggested the training methods best suited for reducing halo, reducing leniency, and improving accuracy.³

Banks and Roberson (1985) argued that rater training programmes did not incorporate standard dimensions of test development and the notion of performance appraisals as "tests" violates accepted standards of test construction.⁴

Performance Appraisal Formats

Of the studies examining formats, behaviourally based methods received the most, but 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.⁵ Prien and Hughes (1987), using a State government sample, showed that mixed standard scales can be used to identify and minimise, individual rater error and system-wide


problems.¹ In 1987, Murphy and Constans concluded that behavioural anchors may lead to biased recall of performance.² But two years later, Murphy suggested that the earlier results were not likely to be observed in organisational settings.³

Legal Ramifications

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,⁴ and Miller, Kaspin, and Schuster (1990) discussed performance appraisal practices related to age discrimination cases.⁵ Using a managerial sample, Greenbery (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).⁶


Case Studies and "How-To" Articles

The professional journals are replete with articles discussing the performance appraisal practices in various organisations 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 10-level forced distribution performance appraisal system into a 3-level system that was deemed to be more consistent with the organisation'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, Control Data, Merck, many Federal and State agencies, and several unidentified organisations.


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, how to make effective use of team appraisals, and how to make appraisals more objective. 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".

**Literature Summary**

Cognitive processing issues clearly dominated performance appraisal research over the years. 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. 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.

From this review of appraisal research it follows that the knowledge of 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 organisational 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.

These studies point to important issues that have barely been addressed or under-researched such as: what determines the design, development and implementation of a particular performance appraisal system in an organisation; how the raters and ratees feel and react to the existing performance appraisal system and processes; and in what way the organisational characteristics affect performance appraisal practice in different organisations.
NEED FOR THE PRESENT STUDY

Performance appraisal research can be evaluated both in terms of its theoretical contribution and its ability to inform performance appraisal practice. Because the rating process involves complex cognitive processes, basic research that defines the nature of the phenomenon is clearly needed. However, because performance appraisals occur in applied social and political contexts, it is also wise to consider the degree to which research is informing practice.

Performance appraisal systems encompass measurement issues as well as attitudinal and behavioural issues. Traditionally, the researchers tended to focus their research efforts on measurement issues per se; yet any measurement instrument or rating format probably has only a limited impact on performance appraisal scores. Broader issues in performance appraisal must be addressed since appraisal outcomes are likely to represent an interaction among organisational contextual variables, rating formats, and rater motivation.

According to some writers, performance appraisal research has done little to improve its usefulness as a managerial decision-making tool.1 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 organisational 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.¹

Napier and Latham (1986) suggested that progress on performance appraisal practice has lagged because the research that might inform practice has ignored Thorndike's (1949) call for practicality in its quest for measurement elegance.² Bernardin and Villanova (1986) concluded that better understanding of the organisational contexts in which appraisal takes place was necessary in order to improve the degree to which performance appraisal research contributes to performance appraisal practice.³

Bedeian (1989) expresses a growing concern that much organisational research, though methodologically sophisticated, lacks substantive application and is directed towards increasingly selective audiences of researchers, to the neglect of other audiences such as policy makers and managers.⁴ This statement is representative of some, though certainly not all, of the recent performance appraisal literature.

There has been a heavy reliance on student subjects and laboratory settings. Although there is some evidence that laboratory settings may provide results that are as valid and generalisable as

¹ Banks, C.G., and Murphy, K.R. 1985, op. cit.
those obtained in field settings, 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 behaviour. 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 ongoing relationship exists in both conceptually and operationally different from the rating task presented in laboratory settings. Therefore, though 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 effective nature of most rating environments.

Many researchers would suggest that accurate appraisals 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 organisational effectiveness within the context of the organisation and the constraints imposed by the regulatory environment in which it operates.


Organisations vary on a wide array of situational or contextual factors (including age of the establishment; ownership pattern – private, public, MNC; nature – manufacturing or service; number of employees; technology; culture; competitive environment; strategy; and union representation) that are likely to affect performance appraisal practices.

Furthermore, certain appraisal issues have received considerable attention but others have been virtually ignored. The predominance of studies examined information processing and psychometric issues, yet virtually no systematic research exists on how the organisational context affects the rater or ratee relationship or the cognitive processes of rater, how raters actually appraise performance, how they use appraisal information, or what issues they believe are important. Moreover, research is needed to address how organisational context affects employee perceptions of appraisal, their reactions to appraisal outcomes, and how appraisal purpose (administrative versus developmental) moderates these relationships. It is to be further noted that Indian research on performance appraisal is relatively sparse as compared to the European or Western studies. No authoritative and comprehensive study has so far been conducted in India to evaluate the performance appraisal system adopted by the Indian private and public sector organisations. Hence, there is an imperative need to address major performance appraisal issues related to the design, development, and implementation of performance appraisal system in private and public sector organisations; views, perceptions, and reactions of the rater and ratees in respect of the existing performance appraisal system; how the appraisal process was intended to be used and how it is actually used in these organisations; and the moderating effect of
organisational characteristics (age of the organisation, ownership pattern, nature – manufacturing vs. service, ISO status, and number of employees) on performance appraisal practices.

**HYPOTHESIS**

It is increasingly acknowledged in the available performance appraisal literature that an organisation cannot simply borrow and implement a new appraisal system from another organisation. This is precisely because the process of designing, developing, and implementing an appraisal system is organisation-specific. Several organisational and situational characteristics are believed to exert their influence on the success or failure of a given performance appraisal system. Accordingly, the researcher formulated the following hypothesis:

- \( H_1 : \) Basic organisational characteristics tend to determine the key dimensions of appraisal practices.
- \( H_0 : \) Basic organisational characteristics do not determine the key dimensions of appraisal practices.

Basic organisational characteristics essentially include: ownership pattern, age of the organisation, organisational output – manufacturing or service, ISO status, and the number of employees. On the other hand, the key dimensions of appraisal practices consist of uses of performance appraisal, type of appraisal – formal or informal, frequency of formal appraisal, categories of performance information collected, sources of appraisal, rating instruments used, type of rating medium, existence of super-review and approval, practice of performance feedback and counselling, system of potential appraisal, and the instances of rater and ratee training.
OBJECTIVES

The chief objective of this study is to describe the current appraisal system and practices prevailing in private and public sector enterprises of Karnataka State, and to highlight the user (rater and ratee) reactions to the existing appraisal system, processes and practices in these organisations. The specific objectives of the study are:

(1) to describe and analyse the key dimensions of current appraisal system and practices in private and public sector units;
(2) to identify and discuss the major determinants of appraisal practices;
(3) to project the views and perceptions of the HR professionals mainly responsible for designing, developing, and implementing appraisal system;
(4) to elicit and report the user (rater and ratee) reactions in respect of the existing appraisal system and practices; and
(5) to raise and present managerial and organisational implications based on the findings of this study to enhance the effectiveness of performance appraisal system, processes, and practices.

RESEARCH METHODOLOGY

The nature of this study is partly explanatory and partly exploratory. This study is coordinated mainly by using primary data and information. However, certain books and periodicals were consulted to understand the relevant conceptual and theoretical aspects relating to the field of study, and also for the purpose of reviewing earlier studies on performance appraisal.

The primary data and information were collected by using two structured, comprehensive, and pre-tested questionnaires — one for
the human resource (HR) professionals, and the other one for the appraisers/raters and appraisees/ratees in the chosen respondent private and public sector units in Karnataka State.

The first questionnaire consisted of relevant questions intending to elicit the responses from HR professionals in respect of the key dimensions of designing, developing, and implementing appraisal system in their respective organisations, and the questionnaire also included certain questions seeking the views and perceptions of these HR professionals about the current appraisal system and practices (Appendix-1). The researcher personally administered this questionnaire to one each HR professional mainly responsible for designing, developing, and implementing performance appraisal system in 52 private sector and 31 public sector enterprises of Karnataka State. This part of the sample size, therefore, consisted of 83 HR professionals representing 83 respondent units. The researcher had very carefully chosen the respondent units fairly and adequately representing private and public sector units in Karnataka State duly reflecting varied organisational characteristics such as ownership pattern, age of the organisation, manufacturing and service sector, ISO certified and non-ISO companies, and the number of employees.

The second questionnaire included two parts. The first part contained questions intending to obtain the views, perceptions and reactions of the raters (appraisers) in respect of using their current appraisal system and procedures for rating their subordinates. The second part of this questionnaire dealt with certain questions seeking the views, perceptions and reactions of the ratees (appraisees) about
their experience with the implementation of the current appraisal system in their respective organisations (Appendix – II).

The researcher personally administered this two-part questionnaire to 921 respondents in 83 sample respondent private and public sector units. This sample size covered 567 employees (399 managers and 168 workers) in 52 private sector units, and 354 employees (243 managers and 111 workers) in the remaining 31 public sector units in Karnataka State. It should be noted that all the managerial personnel acted as both `raters` to their subordinates and `ratees` to their immediate supervisors. However, workers essentially performed the role of only `ratees` in these respondent units. Hence, for the purpose of coordinating data and information pertaining to the ratees’ reactions, the relevant sample size was all the 921 (567 in the private sector and 354 in the public sector) respondents. But the relevant sample size for tabulating and discussing the rater’s reactions was only 642 (399 managers in the private sector and 243 managers in the public sector).

The sample survey of HR professionals, and the raters and ratees was conducted during June-December, 2000. Appropriate and relevant statistical tools and techniques were used such as percentages, mean, standard deviation, and chi-square test for the purpose of analysing and interpreting survey data. Interesting and critical revelations were projected through charts and diagrams.
SCOPE AND LIMITATIONS

This study has been conducted in respect of 83 private and public sector units in Karnataka State. It is widely acknowledged in the performance appraisal literature that the multinational corporations have enough resources and willingness to design, develop, and implement a highly advanced and innovative appraisal system. Of course, it may be argued that other organisations can study and benchmark appraisal system and practices. At the same time, even though nothing prevents from benchmarking “best practices” of appraisal in MNCs, such practices cannot be simply adopted by other organisations due to the organisation-specific and situation-specific attributes of an appraisal system. Therefore, owing to these aforementioned reasons, the MNCs were kept outside the purview of this study.

The researcher had to make several personal visits to the respondent units to collect the relevant information from the HR professionals. More often than not, HR professionals were quite reluctant to freely share their views about their current appraisal system and practices due to the “confidential” nature of the appraisal formats and rating instruments. The researcher was also intending to do the content analysis of the appraisal formats. But the idea was dropped due to the refusal of many HR professionals to part with their appraisal documents and instruments.

In balance, even though the results and discussions of this study cannot be generalised to speak about the appraisal system and practices in other States of India or elsewhere, a clear idea as to the determinants of appraisal practices and also the user reactions to the
current appraisal system and practices in private and public sector organisations can be obtained from this study.

CHAPTER CLASSIFICATION

This study entitled “AN EVALUATION OF PERFORMANCE APPRAISAL SYSTEMS IN PRIVATE AND PUBLIC ENTERPRISES – A Study With Reference to Selected Units in Karnataka State” is coordinated in five chapters as shown below:

CHAPTER - I : “INTRODUCTION” presents a conceptual and theoretical framework, review of literature, need for the present study, hypothesis, objectives, research methodology, and scope and limitations. It also provides the chapter classification at the end.

CHAPTER - II : “PROFILE OF THE SAMPLE” gives the organisational characteristics of the respondent units, a brief profile of the human resource professionals in-charge of designing, developing, and implementing performance appraisal systems in their respective organisations, and the socio-economic-demographic characteristics of the raters and ratees considered for the sample survey.

CHAPTER - III : “PERFORMANCE APPRAISAL SYSTEM – Design, Development and Implementation” describes and analyses the organisational policies and practices pertaining to the design, development, and implementation of appraisal systems in the chosen 52 private and 31 public sector units in
Karnataka State. It also projects the views and perceptions of HR professionals in respect of the key dimensions of current appraisal system. Finally, the chapter concludes by identifying and discussing the major determinants of appraisal practices in these respondent units.

CHAPTER - IV : “CURRENT APPRAISAL SYSTEM AND PROCESSES, AND OUTCOMES – User Reactions” intends to provide a detailed description and analysis of the users’ (raters and ratees) reactions about the current appraisal system, processes, and outcomes. The users’ reactions are presented in two parts – raters’ reactions, and ratees’ reactions. The raters’ reactions (N = 642) focus on their perceptions in respect of key dimensions of current appraisal system, subordinate appraisal and its utility, type of approaches and behavioural patterns used in the last feedback session, and the perceived uses of feedback interview. The reactions of ratees (N = 921) related to their views and perceptions about the major aspects of current appraisal system, degree of awareness about the appraisal issues, nature and perceived utility of the last feedback session, situational constraints impeding performance and its impact on the fairness of appraisal results, and the overall degree of satisfaction in respect of the major components of current appraisal system.
CHAPTER - V : "SUMMARY AND CONCLUSION" provides a capsule summary of the major results and discussions presented in the earlier chapters and raises some important managerial and organisational implications arising therefrom. Finally, the chapter concludes by suggesting some directions for further research in this field.