GENERAL DISTRIBUTION

OCDE/GD(93)177

PUBLIC MANAGEMENT

OCCASIONAL PAPERS

PERFORMANCE APPRAISAL: PRACTICE, PROBLEMS AND ISSUES

by

Robert Wood
and
Verena Marshall

University of Western Australia

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Paris 1993

COMPLETE DOCUMENT AVAILABLE ON OLIS IN ITS ORIGINAL FORMAT


**Occasional Papers on Public Management** (free of charge)

Aspects of Managing the Centre of Government (1990)
Financing Public Expenditures through User Charges (1990)
Serving the Economy Better (1991)

**Market-type Mechanisms Series:**

No.1. Issues and Strategy: Managerial Aspects (1992)
No.2. The United States Experience with Contracting Out: User Charges for Prescription Drugs in Italy; Market-type Mechanisms and Health Services in the UK (1992)
No.3. Property Rights Modifications in Fisheries (1992)
No.4. Vouchers: Social Housing and Nursing Homes (1992)
No.5 Complex Contracting Out Information Technology Services (1992)
No.6 Internal Markets (1993)
No.7 Market Emulations in Hospitals (1993)

**Regulatory Management and Reform Series**

No.1 Current Concerns in OECD Countries (1992);
No.2. Controlling Regulatory Costs: The Use of Regulatory Budgeting (1992)
No.3. Improving Regulatory Compliance: Strategies and Practical Applications in OECD Countries (1993)

**Value and Vision. Management Development in a Climate of Civil Service Change (1993)**

**Internal Management Consultancy in Government (1993)**

**Information Technology in Government: Management Challenges (1993)**

In order to obtain the above papers, please contact:
OECD Public Management Service,
2 rue André-Pascal,
75775 Paris Cedex 16
France
Tel: 45 24 90 52 or 45 24 90 73
Fax: 45 24 87 96 or 45 24 17 06

**Publications** (for sale)

Administration as Service: The Public as Client (1987)
The Control and Management of Government Expenditure (1987)
Flexible Personnel Management in the Public Service (1990)
Public Management Developments: Survey 1990 (1990)
Private Pay for Public Work: Performance Related Pay Schemes for Public Sector Managers (1993)
Public Management: OECD Country Profiles (1993)
Pay Flexibility in the Public Sector (1993)

For these publications, please contact:
OECD Publications Service
2 rue André-Pascal
75775 Paris Cedex 16, France
FOREWORD

An increasing number of OECD countries are developing and using performance appraisal systems in public sector organisations. At the same time, in both the public and the private sectors, there is evidence of widespread dissatisfaction with appraisal, much of it stemming from ineffectiveness of appraisal systems in achieving the purposes for which they were established. Problems with appraisal are attributable, at least in part, to flaws in the design, implementation or operation of schemes. However, the persistence and prevalence of certain problems points to a failure of research on performance appraisal to deal adequately with the difficulties confronted by practitioners who must design and operate appraisal schemes.

This paper reviews the research on performance appraisal, deals with a number of critical issues that need to be addressed in order to improve the practice of appraisal, and suggests ways of dealing with these issues.

The paper has been prepared at the request of the Public Management Service, in the context of work by the Public Management Committee on improvement of human resource management in the public sector. It is hoped that the paper will be of value to practitioners and researchers seeking to develop more effective approaches to performance appraisal.

The views expressed in this paper are those of the authors and do not necessarily represent those of the OECD. The report is published on the responsibility of the Secretary General.

___________________

Copyright OECD, 1993
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUMMARY</td>
<td>5</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>7</td>
</tr>
<tr>
<td>RESEARCH OVERVIEW</td>
<td>8</td>
</tr>
<tr>
<td>The Measurement School</td>
<td>9</td>
</tr>
<tr>
<td>The Appraisal Interview School</td>
<td>12</td>
</tr>
<tr>
<td>Social Cognitive Process School</td>
<td>15</td>
</tr>
<tr>
<td>Performance Management School</td>
<td>17</td>
</tr>
<tr>
<td>IDENTIFYING THE PROBLEMS</td>
<td>19</td>
</tr>
<tr>
<td>Technical Design Problems</td>
<td>19</td>
</tr>
<tr>
<td>Organisational Change and Development Problems</td>
<td>20</td>
</tr>
<tr>
<td>Decision-Support Problems</td>
<td>20</td>
</tr>
<tr>
<td>Strategic Integration Problems</td>
<td>21</td>
</tr>
<tr>
<td>THE ISSUES</td>
<td>23</td>
</tr>
<tr>
<td>The Purpose of Appraisal</td>
<td>23</td>
</tr>
<tr>
<td>Judgements versus Ratings</td>
<td>25</td>
</tr>
<tr>
<td>Accuracy versus Effectiveness</td>
<td>27</td>
</tr>
<tr>
<td>The Organisational Context</td>
<td>29</td>
</tr>
<tr>
<td>CONCLUSION</td>
<td>31</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>32</td>
</tr>
<tr>
<td>FIGURES</td>
<td></td>
</tr>
<tr>
<td>Figure 1. Framework for Analysis of Performance Appraisal</td>
<td>8</td>
</tr>
<tr>
<td>Figure 2. Dual Roles in Performance Evaluation</td>
<td>13</td>
</tr>
<tr>
<td>Figure 3. Characteristics of Feedback for Effective and Ineffective Interpersonal Communication in Human Resources Management</td>
<td>14</td>
</tr>
<tr>
<td>Figure 4. Framework for the Integration of Organisation Strategies, Employee Role Behaviours and Performance Appraisal Systems</td>
<td>22</td>
</tr>
<tr>
<td>Figure 5. Purposes of Performance Appraisal Viewed from the Organisational and Managers’ Perspectives</td>
<td>24</td>
</tr>
</tbody>
</table>
SUMMARY

Within private and public sector organisations in OECD countries, there is widespread dissatisfaction of managers with the processes of performance appraisal. This dissatisfaction often arises from a failure of systems to achieve the outcomes for which they were established. Research on performance appraisal has failed to deal adequately with the practical problems that lead to the shortcomings and the ineffectiveness of performance appraisal systems. Specific criticisms have included an over-emphasis on the design of rating instruments, a failure to consider the organisational settings in which performance appraisals are conducted, and a failure to consider adequately the motivations, capacities and relationships of managers and the employees whose performance they must rate.

This paper provides an overview of the research on performance appraisal from the standpoint of four distinct schools: Measurement, Appraisal Interview, Social Cognitive Processes and Performance Management. The major problems confronted by practitioners who must design, implement and operate performance appraisal systems are considered, and the critical issues which underline the continuing gap between research and practice in performance appraisal are analysed. Recommendations are proposed to resolve the issues confronted in practice.

In the 1950s, performance appraisal ratings were used almost exclusively for administrative decisions such as promotions and salary increases and this was reflected in the concerns for valid rating data of the Measurement School. During the 1960s and 1970s, performance appraisal interviews began to be used for employee counselling and feedback, and research expanded to include the body of studies labelled as the Appraisal Interview School. During the 1970s and 1980s, performance appraisal processes became incorporated into Management by Objectives (MBO) systems that were an extension of corporate planning and budgeting processes, leading to the Performance Management School.

With the emergence of performance management as an influential part of human resource management, performance appraisal has come to be viewed as an integral part of managers’ day-to-day activities. Within the context of the performance management approach, appraisal is a key element in an interrelated set of processes aimed at improving the performance of work units and organisations. The focus is on using appraisal to raise performance levels and to develop employees to their full capacity, in order to achieve organisational goals.

Elements of the performance approach have been adopted by many organisations in both the public and private sectors across the OECD. A substantial part of this report is devoted, therefore, to an analysis of the problems that practitioners confront in applying performance appraisal in the context of performance management systems. There are four broad categories of problems: technical design of appraisal systems; organisational change and development requirements for the effective introduction of performance appraisal systems; the use of performance appraisal data as inputs to administrative decisions and ensuring the suitability of appraisal schemes to function as decision support systems; and strategic integration of appraisal with other human resource management systems, with organisational planning, review and accountability systems and with the overall strategy of the organisation.

While all four areas are important, this report suggests that the two areas that require most attention are decision support systems and strategic integration. These two problems raise several critical issues for the practice of performance appraisal. Issues that relate to the decision support functions of appraisal are the purpose of appraisal, judgements of performance versus performance ratings, and accuracy of appraisals versus effectiveness of schemes. Strategic integration of performance appraisal raises the issue of the organisational context within which appraisal schemes are applied.
There is a growing body of evidence which shows that the purposes of appraisal, that is the decisions or other ends for which appraisal information is used, are the key determinants of the effectiveness of appraisal systems within different contexts. In exploring the different purposes for which appraisal can be used, the report identifies several potential areas of conflict that may arise in organisations and suggests ways of resolving these.

The report goes on to suggest that understanding the relationship between performance judgements and performance ratings is important for the decision-support functions of appraisal. The rating of an employee presented by a manager often does not in practice reflect accurately the manager’s judgement of that employee’s performance; inflation of performance rating is a widespread phenomenon which undermines the usefulness of ratings for decisions on pay or other matters. The report explores the factors that influence ratings, suggests areas where more practical research is needed and points to possible methods of conducting appraisals in order to ensure that ratings provide a more accurate reflection of judgements about performance.

If performance management systems are to have lasting effects, they must be consistent with the way the organisation is managed, with other structural features of the organisation and, in particular, with the strategies of the organisation. Very little field research has been done to establish the effectiveness of different systems across different organisational contexts. This is an area of considerable interest to those who must ask the question: "what is the correct performance appraisal system for my organisation?" The answer to this question will often come from the answer given to questions regarding the strategy and the culture that management wishes to implement within an organisation and the nature of the tasks performed by those to be appraised. The report suggests several aspects of the relationship between organisational context and effective performance appraisal that need to be researched in order to provide guidance to practitioners on the design and operation of appraisal schemes. It also indicates changes that may be required in organisational structures and management systems in order to improve performance appraisal.

The suggestions contained in this report are not exhaustive. However, they are based on an analysis of the issues as they relate to the practice of appraisal and not on theoretical models. Until recently, recommendations for improving performance appraisal have focused on rater training, instrument design, and policy recommendations. It is now generally accepted that many recommendations in these areas, particularly those based on research in the Measurement School, have reached the limits of their usefulness and that a broader range of issues needs to be addressed if research is to provide improvements in the practice of performance appraisal. At the same time, performance appraisal has been incorporated into performance management systems and needs to be studied in that context to be properly understood.
In spite of the widespread and increasing adoption of performance appraisal systems in both public and private sector organisations across OECD countries, there is widespread dissatisfaction with appraisal. The question is often asked, "is performance appraisal worth the effort?" Many managers don’t think so. Napier and Latham (1986) found in a survey of US managers that most believed there was little practical value in doing performance appraisals thoroughly and consistently. In short, many felt that the effort involved in completing performance appraisals was simply not justified. Proponents of the increasingly popular Total Quality Management approaches have suggested that individual performance appraisals should be totally replaced by objective measures of performance that focus on the plant or the work group level (e.g. Deming, 1986). Supporters of performance appraisals argue that it is the lack of objective measures of performance for many jobs that makes performance appraisals necessary.

Much of the dissatisfaction with performance appraisal in practice comes from ineffectiveness in achieving some of the purposes for which schemes are established. For example, it is a common finding that performance appraisal ratings suffer from extreme leniency effects, resulting in 70-80 per cent of those assessed being rated in the top two categories of rating scales (Landy and Farr, 1980; Longenecker, Sims and Gioia, 1987; OECD, 1993). The resulting lack of discrimination makes the data from performance ratings ineffective for making the necessary discriminations in decisions such as salary, promotion or training. However, as we will argue later, while leniency may limit the usefulness of appraisal data for administrative decisions it may serve other purposes. For example, a manager may use lenient ratings to obtain the support and commitment of employees.

Some of the problems with performance appraisal are attributable to flaws in the design, implementation or operation of performance appraisal schemes in organisations. Evidence from surveys consistently shows that managers spend little time on the appraisal process, that managers are not systematically trained in how to use rating formats or how to conduct appraisal interviews, and that managers are not held accountable for either the manner in which they conduct appraisals or the outcomes (Bretz, Milkovich and Read, 1992; Milkovich and Wigdor, 1991; Wyatt, 1989). From these findings it could be argued that problems with performance appraisal, such as leniency in ratings, are due to design faults and the lack of proper implementation of the process. According to this line of argument, solutions to problems with performance appraisal should be found in improvements to the rating instruments, training programmes, policies and procedures of specific schemes.

However, the documentation of problems such as leniency effects is both widespread and persistent, dating back to the 1950s (Bernardin, Orban and Carlisle, 1981; Heron, 1956; Sharon and Bartlett, 1969; Taylor and Wherry, 1951). The persistence of leniency effects, despite all the improvements in training programmes, policies and procedures that have occurred over the last 40 years, suggests that there may be some fundamental problems with the practice of performance appraisal in organisations. Alternatively, it could be argued that the lack of progress may be due to performance appraisal researchers studying the wrong problems.

There is a body of literature that attributes the continuing dissatisfaction and problems in performance appraisal practice to shortcomings in the research on performance appraisals (e.g. Banks and Murphy, 1985; Ilgen, Barnes-Farrell and McKellin, 1993; Landy and Farr, 1980; Murphy and Cleveland, 1991). According to this line of argument, performance appraisal research has failed to contribute significantly to developments in the practice of performance appraisal because it has failed to deal adequately with the practical problems that managers confront in the conduct of appraisals. Specific criticisms have included an over-emphasis on the design of rating instruments, a failure to consider the organisational settings in which performance appraisals are conducted and a failure to consider adequately
the motivations, capacities and relationships of managers and the employees whose performance they must rate (Bernardin et al., 1982; Murphy and Cleveland, 1991). These criticisms are not new, having been clearly articulated by Landy and Farr (1980) more than a decade ago. However, research in the ensuing period has been criticised for displaying many of the same shortcomings as those identified by Landy and Farr in 1980 (Murphy and Cleveland, 1991).

The noted lack of a practical orientation in much performance appraisal research may be a product of the manner in which research problems are formulated. Banks and Murphy (1985) have argued that the questions asked in performance appraisal research have been motivated more by research models and the researchers’ desire to improve their knowledge of basic judgement processes and less by practical problems of managers who must conduct appraisals. This basic research orientation has contributed to knowledge about specific judgement effects that have been identified and validated in laboratory settings (Ilgen et al., 1993). For example, research based on social-cognitive theories, which have provided the guiding framework for much performance appraisal research in the last 10 years, has established that race and gender of the individual being rated have very small impacts on performance ratings compared to actual performance (Pulakos, White, Oppler and Borman, 1989; Schmitt and Lappin, 1980). This finding is of interest to researchers who study the impact of stereotypes on judgement but adds little of value to the practice of performance appraisal. The social cognitive school of performance appraisal research has made several similar contributions to the understanding of social cognition and the way it functions in judgement processes (Ilgen et al., 1993). However, some argue that such discoveries have little practical relevance for the design or conduct of performance appraisals and have not led to significant improvements in practice (e.g. Banks and Murphy, 1985).

In the subsequent sections of this paper we will provide an overview of four identifiable schools of research in performance appraisal, followed by a discussion of the problems that organisations confront in the design, implementation and operation of performance appraisal systems. In the final section we deal with several issues that we believe should be the focus of future research if the practice of performance appraisal is to be improved. A framework for the analysis of the schools of research, problems and issues associated with performance appraisal is presented in Figure 1.

![Figure 1. Framework for analysis of performance appraisal](image)

<table>
<thead>
<tr>
<th>SCHOOLS</th>
<th>PROBLEMS</th>
<th>ISSUES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Development</td>
<td></td>
</tr>
</tbody>
</table>

**RESEARCH OVERVIEW**

Performance appraisal research has a relatively long history among studies of human resource management systems and practices. The formal study of performance appraisal has been traced to work on selection by industrial psychologists in the United States during the period leading up to the Second...
World War (e.g. Scott, Clothier and Spriegel, 1941, cited in Murphy and Cleveland, 1991, p. 3) and to the seminal statement on selection by Thorndike (1949). Since the early 1950s, there has been a highly focused and active body of researchers studying the measurement instruments, cognitive and behavioural processes and, to a lesser degree, broader contextual issues of performance appraisal. In providing a brief overview of this research, we have found it convenient to separate studies into four broad schools or perspectives on performance appraisal. The schools are based on the questions studied and the assumptions made about the practice of performance appraisal. We have labelled the four bodies of research as the Measurement, Appraisal Interview, Social Cognitive Processes and Performance Management schools to reflect the major focus of each school. The bodies of research in each of these schools are more notable by their beginnings than their endings. For example, the Measurement School, which has focused on the design of appraisal instruments since at least the seminal statement by Thorndike in 1949, still continues today.

The development of the different schools does demonstrate one important link between research and performance appraisal practices in that changes in the problems studied by the different schools reflect changes in the uses of performance appraisal data over the last 40 years (DeVries et al., 1986). As the uses of performance appraisal have increased in organisations, the issues studied by researchers have expanded to include topics that relate more directly to the new aims. In the 1950s, performance appraisal ratings were used almost exclusively for administrative decisions such as promotions and salary increases and this was reflected in the concerns for valid rating data of the Measurement School. During the 1960s and 1970s, performance appraisal interviews began to be used for employee counselling and feedback, and research expanded to include the body of studies labelled as the Appraisal Interview School. During the 1970s and 1980s, performance appraisal processes became incorporated into Management by Objectives (MBO) systems that were an extension of corporate planning and budgeting processes, leading to the Performance Management School.

The Measurement School

In most performance appraisal systems an important end product is a rating of one individual by another. Ideally, assessments of performance would be based on objective measures of outputs and outcomes. However, the reality in organisations is such that most jobs have many important outcomes that cannot be assessed through reliable and valid quantitative indicators. Many aspects of jobs have either delayed outcomes or are inputs to other jobs. Under these circumstances, subjective assessments of work become a necessity if decisions are to be based on the performance of individuals. The design of performance appraisal instruments was an attempt to accommodate this reality and to ensure that the subjective assessments of one person’s work by another person were the most reliable and most valid that could be made.

The measurement perspective has been traced to a statement on the criterion model of measurement by Thorndike (1949), although studies of rating forms based on trait psychology had been conducted even earlier (e.g. Scott et al., 1941). In the early research on the topic, performance appraisal was treated as a measurement problem which could be solved through the proper design of instruments or rating scales used to record performance judgements. Other factors which might have been expected to influence a manager’s rating of an employee, such as the manager’s skills and motivation, characteristics of the employee and the organisational context in which appraisals were conducted, were considered sources of error that could be minimised through the correct design of rating instruments and the training of raters (Murphy and Cleveland, 1991).

From the measurement perspective, performance appraisal formats are chosen for the accuracy of the ratings. It is assumed that a manager’s major concern is to provide ratings which accurately represent the performance of the individuals being appraised. This assumption has been challenged by those who
argue that a concern for accuracy is only one of many motives that may influence the performance rating a manager gives an employee. Longenecker and colleagues (1987), for example, established that people seek to protect their self-interest when giving performance appraisal ratings. The protection of one’s own position has been shown to be a major determinant of the flow of information and the transmission of judgements such as performance appraisal ratings in organisations (Lawler and Rhode, 1976). Managers will also often give ratings that increase the probability of an employee receiving a merit award in competition with other employees. In doing so, managers are using ratings as a means of rewarding their staff and may also increase their own power in this process. The performance rating of a staff member will reflect on the manager’s own performance; if the subordinate is seen to be a good performer, the manager may be considered more effective by his or her colleagues.

The overwhelming majority of the research on rating scales has been concerned with the development of criterion-referenced formats which require the rating of an individual’s performance against some established standards, usually represented as points on a rating scale (Landy and Farr, 1980). Norm-referenced criteria which require comparison between individuals, such as rankings and paired comparisons, are infrequently used in practice and have received relatively less research attention, particularly in recent times. Reviews of the research on norm-referenced criteria formats are available in Carroll and Schneier (1982, pp. 124-129) and Landy and Farr (1983). It is unfortunate that there is not more widespread use of norm-referenced criteria formats such as rankings, which force a distribution on the performance judgements about different employees. The distributions that are produced by rankings could make performance appraisal data more useful in administrative decisions that require multiple levels of discrimination between staff (e.g., allocation of bonuses when there are a limited number of awards available). Ranking distributions also help eliminate the clustering of employees at the top of the distribution (rater leniency), at the bottom of the distribution (rater severity), or in the middle (central tendency). Ranking distributions are most useful when a large number of employees must be appraised and there is more than one appraiser (Cascio, 1992). As performance pay allocations remain the most common use of performance appraisal data in several OECD countries, further research on alternative ranking formats could have practical value. Attempts to obtain the needed levels of discrimination for performance pay allocations by forcing distributions on to data from criterion-referenced rating scale formats have proven most unpopular (e.g. OECD, 1993).

The criterion-referenced rating scale formats that have been studied most frequently are graphic rating scales, behaviourally anchored rating scales (BARS), mixed standards scales (MSS) and behavioural observation scales (BOS).

**Graphic Rating Scales** are the simplest formats used to record judgements about aspects of an individual’s performance. Ratings are recorded along either an unbroken continuum or in discrete categories within a continuum. The end points and categories within a scale may be labelled with either simple numerical anchors (i.e. 1, 2, 3, 4, 5, etc.) or adjectival anchors (e.g. very poor, average, outstanding). The scales are simple to develop and require little training to use. The criteria rated are usually either important dimensions of job performance (e.g. quality, quantity, timeliness) or personal characteristics (e.g. co-operation, oral skills, honesty). Critics of graphic scales point to the lack of clarity in the criteria used and the inadequate definitions of standards. The frequent use of traits as criteria was another source of criticism of graphic rating scales. However, traits have proven to be valid criteria when given the proper behavioural definitions (Bernardin and Beatty, 1984).

**Behaviourally Anchored Rating Scales (BARS)** were a major focus of research on performance appraisal in the 1970s. BARS represented an attempt to enhance the quality of graphic rating scales by providing clear behavioural definitions of criteria and by replacing numerical and adjectival anchors with behavioural examples for each of the performance level categories on the
The criteria in BARS’s are job dimensions which cover the different categories of duties, responsibilities and activities for a job. BARS are developed by a process which includes systematic analyses of jobs and extensive consultation with users. In the United States, the interest in BARS was strongly influenced by the fact that the BARS methodology for developing job-related criteria complied with the requirements of equal employment opportunity (EEO) legislation in that country (Carroll and Schneier, 1982; Smith and Kendall, 1963). The assumption that defining performance in behavioural terms would produce more accurate ratings has not been supported in subsequent research (Murphy and Cleveland, 1991). The involvement of users in the development of BARS has made them more costly to develop than alternative formats but has also increased their acceptance amongst users.

Mixed Standards Scales (MSS) also use behavioural examples to define the categories of performance levels but they have a different organisation of items and a different response format from BARS. The behavioural examples which define poor, average and good performance in a MSS are developed through a consultative analysis of jobs which is the same as the equivalent steps in the BARS methodology. In a MSS, however, the behaviour examples for the different levels of performance for all criteria are presented on the appraisal forms in a randomly mixed order. Managers and staff are not told which items measure which job dimensions or how their ratings of items are converted into a performance score. This lack of transparency is intended to minimise the biases created when a person is rated against known categories or job dimensions. MSS’s simplify the rating task and are easy to use but their lack of transparency about the derivation of performance scores makes it difficult to give feedback based on the ratings.

Behaviour Observation Scales (BOS) are another variation of the BARS approach. BOS’s are similar in format to MSS’s but require a different type of judgement by the rater. Instead of an evaluative judgement, BOS’s ask for ratings of the frequency with which an employee engages in individual behaviours (e.g. very often, often, occasionally, seldom, almost never).

Further behaviour-oriented appraisal methods include the narrative essay, in which the rater describes in writing an employee’s strengths, weaknesses and potential, together with suggestions for improvement. The behavioural checklist asks the rater to ‘check’ from a list of statements describing job-related behaviour the statements most closely resembling the staff member’s behaviour. Another methodology is critical incidents which consist of brief written reports by raters of employee behaviours and accomplishments that were particularly effective or ineffective in relation to parts of their job.

Extensive research comparing the different formats has led to the conclusion that no single approach is consistently superior to any other in accuracy of the ratings given (Bernardin and Beatty, 1984; Carroll and Schneier, 1982; Landy and Farr, 1980; Murphy and Cleveland, 1991). No rating format has been fully successful as a basis for merit pay, bonuses or promotional decisions. All suffer from the previously mentioned leniency effects. The rating scores obtained from different formats tend to be highly correlated, suggesting that all formats measure similar things (Bernardin and Beatty, 1984). The correct choice of format is contingent upon the purposes for which performance ratings are to be done. For example, behavioural approaches produce ratings that are most amenable to interpretation and therefore the most useful for feedback and counselling purposes. BARS are very good for these purposes but are the most difficult to develop, maintain and use.

Many specific recommendations for the development and design of rating instrument formats have emerged from the measurement body of research. Research on the specific components of performance appraisal formats, including the number of response categories, the types of anchors used on rating scales, the source of ratings (supervisor, peer or self) and the processes for development of rating scales have been extensively reviewed elsewhere (e.g. Bernardin and Beatty, 1984; Carroll and Schneier, 1982; Landy and
Farr, 1980). Readers are referred to those sources, particularly Bernardin and Beatty (1984), for answers to specific questions such as "how many points should there be on a rating scale?" Practitioners are warned that the choice of specific design options, like the choice of an overall format, frequently depends upon what the appraisal instruments are to be used for.

In recent times, research in the Measurement School has been criticised for lack of attention to practical criteria, such as the ease of use of scales, and for studying the accuracy of ratings in contrived settings (Latham, 1986; Murphy and Balzer, 1986). Much of the research into the accuracy of ratings has been conducted in laboratory settings with students as subjects. Field studies of performance appraisal rating formats in operating organisations are far less frequent (Murphy and Cleveland, 1991). As a result of these types of criticisms, the research attention given to rating formats has diminished to almost nil since Landy and Farr (1980) first pointed out that no one rating scale format was superior to the others and went on to suggest that further development of rating scale formats would not improve the quality of performance ratings.

The Appraisal Interview School

During the 1950s a body of work was started which focused on the communication processes in appraisal interviews and other managerial dynamics of the appraisal process. An important early contributor to this school was Maier (1958a), who identified three different communication styles in his studies of appraisal interviews: tell and listen, tell and sell, and problem-solving. The problem-solving approach, in which the manager and employee share information and work constructively towards solving problems and improving performance, has been shown to produce more positive employee reactions and better work-related outcomes than the other two approaches (Burke, Weitzell and Weir, 1978; Maier, 1958b).

Since the early work of Maier and others (e.g. Meyer, Kay and French, 1965) studies of appraisal interviews and performance reviews have come to focus on the structure of the interview process and the content of the communications. The latter has dealt extensively with the nature of feedback and the role of goal-setting in the performance appraisal process (e.g. Latham and Wexley, 1981). The facilitative effects of goal-setting and feedback in appraisal processes have been well documented (Ilgen, Fisher and Taylor, 1979; Latham and Wexley, 1981).

The role of the manager in the appraisal interview process is seen as that of a coach or counsellor who must support and develop employees, as distinct from the role of a judge whose primary function is evaluation of employees (McGregor, 1957). The latter, of course, is the role assigned to the managers in the Measurement School. Studies of appraisal interviews highlighted the differing roles that managers must play in the performance appraisal process and the conflicts inherent in these roles. The conflict inherent in the manager’s roles of judge and counsellor in pursuit of the different purposes of appraisals is outlined in Figure 2.

Appraisal interview researchers view personal development as one of the central purposes of performance appraisal. To serve this purpose, appraisal interviews require open communication and trust between managers and their staff. Studies of the relationships between managers and their staff have also shown that the quality of those relationships can affect the accuracy of performance ratings. For example, the greater the openness and trust between a manager and his or her staff, the less inflated the performance ratings given by the manager (Padgett, 1988, cited in Ilgen et al., 1993). This is consistent with the evidence that supportive management styles lead to the acceptance of higher performance standards by employees (Carroll and Tosi, 1970; Likert, 1967). If the managerial process is supportive and there is trust and open communications between managers and employees, the standards set in appraisals will be higher and the performance ratings are more likely to be accurate.
Figure 2. **Dual roles in performance evaluation**

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Counselling</th>
<th>Judging</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance feedback</td>
<td>Reassurance, involvement, recognition</td>
<td>Making employees aware of impact measures, performance adequacy</td>
</tr>
<tr>
<td>Changing performance plans</td>
<td>Mutual agreement on expectations</td>
<td>Introducing new (higher) standards</td>
</tr>
<tr>
<td>Developing capabilities</td>
<td>Motivating, focusing development for future improvement</td>
<td>Identifying gaps and development needs</td>
</tr>
<tr>
<td>Staffing and compensation actions</td>
<td>Build understanding of link of performance to actions</td>
<td>Provide sound defensible basis for actions</td>
</tr>
</tbody>
</table>


Studies of the dynamics of appraisal processes have also shed some light on the motives of both managers and the staff they must evaluate. Interestingly, performance appraisal interviews are often more threatening for the manager who must conduct the appraisal than for the employee who is being evaluated. Longenecker and Gioia (1988) have identified several sources of appraisal apprehension for managers. These include the level of trust between the manager and the employee; the extent to which the employee’s job responsibilities, authority and goals are defined; whether or not there has been regular and specific feedback throughout the appraisal period; the political nature of appraisals; and the manager’s perception of a hidden agenda. The manager’s confidence or self-efficacy regarding his or her ability to make accurate judgements and communicate them to the employee, and the degree to which the manager feels that he or she will be held accountable for the accuracy of the ratings, influence both the accuracy of ratings and the perceived quality of the overall appraisal process (Bernardin and Cardy, 1982; Marshall, 1993; Murphy and Cleveland, 1991).

A considerable body of practical recommendations regarding the structuring of appraisal interviews, the methods for presenting feedback to staff and communication strategies has also emerged from the studies of appraisal interviews. Figure 3 presents a summary of the recommendations regarding positive and negative feedback that have been developed from this research. The general thrust of these recommendations is to encourage managers to give frequent and specific feedback to employees and to avoid negative emotional comments.

In discussions of performance, the quality of communication strategies and the level of trust that underpins exchanges between a manager and an employee are critical to the success of performance reviews and appraisal interviews. Communication strategies which protect the self-esteem of both parties have generally proven to be most productive (Carroll and Schneier, 1982). Recommendations such as those outlined in Figure 3 are presented to managers as a strategy for giving negative feedback in a manner that is constructive and which presents minimal threats to recipients. Defensive reactions to negative or even somewhat-less-than positive feedback can create problems in the relationship between a manager and his or her staff. There is evidence that labelling employees as satisfactory, as distinct from outstanding or above average, leads to reduced commitment and negative attitudes in subsequent dealings between the manager and employee (Pierce and Porter, 1986). In general, it has been found that employees are more
satisfied and more accepting of performance appraisal outcomes the greater their inputs into the process of appraisal and into the design of instruments used to appraise them (Dobbins et al., 1988; Silverman and Wexley, 1984). Managers also do not like to give negative feedback and those who are effective at doing so tend to practise or rehearse strategies for what they want to say (Larson, 1989).

Figure 3. Characteristics of feedback for effective and ineffective interpersonal communication in human resources management

<table>
<thead>
<tr>
<th>Effective Feedback</th>
<th>Ineffective Feedback</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Intended to help the employee</td>
<td>1. Intended to belittle the employee</td>
</tr>
<tr>
<td>2. Specific</td>
<td>2. General</td>
</tr>
<tr>
<td>3. Descriptive</td>
<td>3. Evaluative</td>
</tr>
<tr>
<td>4. Useful</td>
<td>4. Inappropriate</td>
</tr>
<tr>
<td>5. Timely</td>
<td>5. Untimely</td>
</tr>
<tr>
<td>6. Employee readiness for feedback</td>
<td>6. Makes the employee defensive</td>
</tr>
<tr>
<td>7. Clear</td>
<td>7. Not understandable</td>
</tr>
<tr>
<td>8. Valid</td>
<td>8. Inaccurate</td>
</tr>
</tbody>
</table>


The consistent message that emerges from research in the Appraisal Interview School is that good management practices lead to more accurate and more acceptable appraisals. Good management practices include open communication, clear expectations and a supportive supervisory style. Trust is a key intervening factor between a manager’s actions and the effectiveness of a performance appraisal. Some have argued, however, that the emphasis on participative work cultures is more rhetoric than reality and that performance appraisal systems have stressed individual accountability and managerial control to the exclusion of participation and self-management (Guest, 1990; OECD, 1993).

A limitation of research into the conduct of appraisal interviews, goal-setting and feedback practices is that it assumes an apolitical model of management. However, the processes of performance appraisal often involve highly political manoeuvres by individuals. Managers often rate employees leniently to ensure they receive maximum merit pay increases and the employees, in turn, repay the managers with their support and acceptance (Longenecker et al., 1987).

Performance appraisal interviews can also lead to significant conflicts over values, information and resources. Recommendations regarding the structuring of feedback assume that clarity and specificity are sufficient to gain staff acceptance of ratings and other performance-related information. However, the practical realities of many organisations are such that conflict resolution will involve all manner of trade-offs and negotiation to establish an agreed performance rating. Managers have ongoing exchange relationships with employees and are dependent upon them for commitment and support. Maintaining the support of employees will often take precedence over openness of communications and the accuracy of ratings during an appraisal interview.
Social Cognitive Process School

In the early 1980s there was a major shift in the focus of performance appraisal research away from instrument design and rater training to understanding the rater as a motivated decision maker who processes a range of social cues in making judgements about other people (Feldman, 1981; Landy and Farr, 1980). This shift in focus was concurrent with an upsurge in the interest in social cognitive theories at the time and frameworks adopted from social cognitive theories dominated performance appraisal research in the 1980s. The social-cognitive process school of performance appraisal research has focused on different stages of the performance appraisal judgement process, including the acquisition of information, the organisation and storage of information, and the retrieval and integration of information into a rating (Ilgen et al., 1993).

In the Social Cognitive Process School, rating accuracy was retained as a key criterion, but the determinants of accuracy studied go beyond the rating formats of the Measurement School to include the rater’s ability and his or her motivation to rate accurately. There is considerable evidence from social cognitive bodies of literature that the information processing and recall capabilities of individuals are limited relative to the demands of the typical performance appraisal task in most organisational settings (Landy and Farr, 1980; Murphy and Balzer, 1986). Another important result from the research has been the recognition that factors which affect the motivation of the appraiser, such as the purpose of the rating and accountability for the outcomes, along with the rewards and punishments associated with appraisals, will influence the effort put into the appraisal process and the quality of the outcomes (Marshall, 1993). The accuracy of ratings will depend more upon the motivation, or lack thereof, of managers to rate accurately than upon the specific format.

Managers vary in their motivation to make accurate appraisal judgements for different purposes, and they require different information in reaching those judgements. People’s observations, and therefore their appraisal judgements, vary depending upon the purpose to which the performance appraisals are put (Williams et al., 1985). For example, salary decisions lead to greater requests for information during the appraisal process. It appears that the purpose of appraisal may shape a schema which then influences the amount and nature of what is observed by managers in their assessments of employees, including the level of detail of behaviours observed (DeNisi et al., 1984; Foti and Lord, 1987). Much of the variance in ratings discussed as rater error may in fact be rater reaction to multiple, competing and conflicting demands or purposes (Murphy and Cleveland, 1991). Simple-to-use ratings formats that are linked to a clear purpose and with clear accountability will produce more accurate rating than those that lack these motivational characteristics, irrespective of the other technical properties of the formats.

The motivation of a manager to rate accurately is also related to his or her knowledge about the job of the person being appraised and the rating scale to be used (Bernardin, Cardy and Abbott, 1982). If categories used in the evaluation of employees are matched to the performance dimensions of the job then raters will be more accurate (Ilgen et al., 1993). Consistent with this basic finding is the recommendation that scales be based on job analysis and therefore focused on the dimensions of the task.

In most appraisal systems, ratings are conducted annually. Delays between the observation of data and the recording of ratings has been shown to affect the recall of behaviours but, in some cases, to not seriously affect the accuracy of the overall performance ratings or rankings of employees (Barnes-Farrell and Coutoure, 1984; Murphy and Balzer, 1986). The research evidence suggests that managers form general impressions of staff and these categorical judgements (e.g., good, average, or poor performer) may be quite accurate as a basis for an annual performance rating. At the same time, the recall of specific information and detail on which the judgements were based may be lost to the manager. The loss of specific detail on behaviours and performances may make it difficult for the manager to give feedback on why an employee received a particular rating even if that rating is accurate. Other studies have established
that memory decay for specific information happens quickly and can lead to bias in ratings if the original impressions are inaccurate (Kozlowski and Kirsch, 1987; Murphy and Balzer, 1986).

Stress also affects memory and the recall of information (Srinivas and Motowidlo, 1987). Performance appraisal interviews are often stressful for the manager and therefore he or she may be placed in a position of trying to justify a rating but not able to recall the information that relates to why he or she made the judgement. Research on memory clearly demonstrates the importance of continuous reviews and record keeping if performance of several employees is to be observed, recalled and evaluated and reported in the form of ratings. The information processing demands of such tasks simply exceed the memory and recall capacity of most individuals.

Following a review of the social cognitive process research conducted throughout the 1980s, Ilgen and his colleagues (1993) have identified four major contributions. First, it has raised awareness of the importance of observation in performance appraisal. It has now become accepted that what managers attend to is influenced by a complex range of factors, including prior knowledge, attitudes towards an employee, the job and the purposes for doing the rating. More work is needed on identifying techniques for improving the observation of performance data. Observation techniques need to be practical and cost-efficient for the managers to use in the organisational context. Diaries and other data-recording techniques may be needed for managers to have valid data available when making annual appraisal judgements. Although little research has been done in this area, diary-keeping skills and techniques offer promise as a method for improving the effectiveness of performance appraisals (Bernardin and Cardy, 1982).

Second, social cognitive process research has corrected the belief that rating biases are, necessarily, evidence for inaccuracy in ratings. There has been a shift in training away from rater error training towards training managers how to actually rate tasks and to agree on the norms and standards used within an organisation. The new training approach is referred to as frame-of-reference training and seeks to provide all managers with a common set of standards for the behaviours that define effective performance for each of the criteria on the rating scale (e.g. Borman, 1979). This shift in emphasis is of considerable importance when we consider that most ratings data are used for purposes of comparison between different staff. Therefore, individual biases may be less relevant in determining the overall validity of rating data than a common agreement on the standards to be used.

The aim in frame-of-reference training is to provide managers with greater knowledge of the specific standards which are attached to performance criteria to be used in appraisals. These training programmes reflect the basic principles of the social learning model (Bandura, 1986). The technique is quite straightforward. Managers are brought together, preferably in the time period immediately preceding the annual appraisals, and are given examples of good and poor performance for the rating scales to be used. Managers then practice giving ratings on a set of case examples and discuss standards used to rate the different cases presented. The managers are then given feedback on the accuracy of their judgements and the relevance of the standards used. Through this process, managers who participate in the training programme develop a common set of norms as to what constitutes different levels of performance for the criteria used in appraisals and a common schema for assessing performance.

Third, a manager’s expectations about an employee and the purposes for which the ratings are obtained will strongly influence the ratings themselves. The predictive judgements of a manager are not independent from the evaluative judgements to be made, when rating an employee’s performance. Unfortunately, research has little to say beyond this more general statement about the specific purposes such as salary increment, development and feedback and how they affect ratings in different contexts. This is an area where further research is clearly needed.
Finally, according to Ilgen and colleagues (1993), the Social Cognitive Process School builds on the earlier Measurement School to demonstrate that no matter what scale is used, raters do not simply report what they observe. People actively evaluate information and are influenced by their own motivations and goals. Having established this general principle, more research is needed on the different purposes that individuals and organisations may have for appraisals and how different motives and goals influence the judgements of managers. Perhaps more importantly, work needs to be done on how to design performance appraisal systems which properly align organisational purposes and the goals of individual managers in performance appraisals.

Social cognitive research has helped to clarify and refocus research on performance appraisal. However, the models adopted have made few contributions to the practice of performance appraisals in organisations because, according to Murphy and Cleveland (1991) and others:

1. They have paid insufficient attention to the context in which appraisals occur.
2. They have failed to identify issues of concern to practitioners as well as researchers.
3. They have failed to demonstrate links between the research and the practice of performance appraisals.
4. They have had an overly narrow focus on the information-processing and judgement aspects of appraisal.
5. They have focused rating accuracy as the major effectiveness criterion and have ignored other more relevant criteria.

**Performance Management School**

More recently, it has been recognised that performance appraisal is a tool for managing future performance and not just a means of scoring past performance (Schneier, Beatty and Baird, 1986). According to the Performance Management School approach, performance appraisal is an integral part of a manager’s day-to-day activities and a means of shaping expectations, encouraging and motivating staff, identifying and solving problems and allocating resources. This more broadly framed and integrative view of performance appraisal treats the appraisal interview and the rating of performance as components in the performance management process. Performance ratings have been de-emphasised, relative to the planning, reviewing and problem solving aspects of that process. Performance management is an attempt to transform performance appraisal from an annual or six monthly event to a continual process that is integrated with other management systems and the every day work of the manager.

Several factors have contributed to the development of appraisal as a sub-component within a broader performance management system. First has been the widespread adoption of MBO-based performance appraisal systems in which the annual appraisal rating is part of a cycle that includes goal-setting, action plans, regular performance reviews and continuous informal feedback (Levinson, 1976). Second is the fact that many of the administrative decisions that have become linked to performance appraisal ratings are in fact critical to the development and performance of staff and, therefore, are critical activities in the manager’s role. Finally, there is the recognition that in most organisations performance appraisal is one of several planning and accountability systems that direct managerial work and these need to be linked to one another.
Critical to the successful implementation of performance management systems is recognition of the need to integrate the performance appraisal process with the day-to-day job of the manager (Schneier, 1989; Schneier, Beatty and Baird, 1986). Incorporating the observation, recording and feedback aspects of performance appraisal into the day-to-day activities of the manager’s role is often difficult due to the dynamic and fragmented nature of that role. Performance appraisal systems often assume a sequential attention to tasks, while the research evidence demonstrates that managers often work on tasks in episodic, disconnected ways (Kotter, 1982; Mintzberg, 1973). Regular performance review discussions in which goals are adapted to current demands and resources can help overcome the potential mismatch between the formal appraisal process and realities of a manager’s work. Focusing on outputs and outcomes and delegating the responsibility for the means of achieving them to the individual manager also shifts the focus of the appraisal away from the less observable flow of managerial work activity.

Integrating performance appraisal with other planning and accountability systems will mean that appraisal standards for managers will often be taken from budgets and other plans. Performance management systems may provide strategic integration through the use of common performance indicators in corporate, departmental and individual managers’ plans and performance agreements. Customer service, profitability and quality, for example, may be incorporated in goals at all levels as a means of ensuring that corporate goals in these areas are pursued throughout the organisation. The performance appraisal process then becomes one of several tools for implementing and achieving departmental or organisational goals and must, therefore, be co-ordinated with the organisational planning and budgeting processes. Coordination of different planning processes through the individual, operational and strategic levels will include issues of timing and the common usage of jargon, as well as the hierarchical linking of goals, resources and accountability processes. Goals and resources need to cascade down from corporate and departmental plans to individual performance agreements. Accountability then flows back up the hierarchy as review processes aggregate results at successive levels.

The orientation of the Performance Management School is toward organisational level performance. Systems will be assessed against perceived contributions to the achievement of work-unit and organisational goals. The general criterion is often referred to as "value-added". Accuracy of ratings is not considered a critical criterion unless related to the achievement of the system goals. For example, it may be considered perfectly reasonable for a manager to rate an employee more leniently if, as a result, the employee becomes more committed to a work team and puts in higher levels of effort than he or she might otherwise have done. From the manager’s perspective, performance ratings may be judged in terms of their more immediate contribution to the performance of staff and the work unit and not, necessarily, for their validity as inputs to decisions about promotions or performance pay (Landy and Farr, 1980; Murphy and Cleveland, 1991).

The MBO-based systems of appraisal that are commonly used in performance management systems have been criticised for their lack of comparability in the ratings of employees because of differences in the difficulty of goals set and other factors which can determine levels of performance (e.g. Bernardin and Beatty, 1984). Such differences can limit the usefulness of the rating data for administrative decisions that require valid comparisons to ensure equitable and merit-based differences in the outcomes for different employees. As a result, the performance appraisal component of a performance management system may not provide adequate data for decisions such as promotions or salary awards (Bernardin and Beatty, 1984; Murphy and Cleveland, 1991). This conclusion, which is often stated as though common wisdom in the performance appraisal literature, needs further research.

In most MBO-based systems of appraisal, judgements of performance against agreed goals are summarised in a global five-point rating scale (e.g., OECD, 1993). The research questions that need to be addressed to establish the validity of these ratings for administrative decisions include "How do assessments of performance against agreed goals compare in accuracy with observations of a manager’s behaviour?" and
"How does the use of a single global scale compare in accuracy with other rating formats?" Given the popularity of the MBO-based format for performance appraisal, research directed specifically at ways of improving the validity of ratings based on goal achievements would be of potentially great practical value.

There has been little direct research on the effectiveness of performance appraisal formats within performance management systems. However, there is a body of research on MBO systems which is relevant, even though it does not deal directly with the appraisal components of MBO systems (cf. Barton, 1981; McConkie, 1979; Rodgers and Hunter, 1989). Goal-setting and feedback are the most validated of all management practices and have consistently been shown to lead to significant gains in individual productivity (Locke and Latham, 1990). Goal-setting and work planning systems that incorporate many of the features of performance management systems have also been shown to provide significant improvements in organisational productivity, particularly when they are used and supported by senior managers (Latham and Wexley, 1981; Rodgers and Hunter, 1989). Therefore, although there is little research that deals directly with the effectiveness of performance management systems, the general approach and various components of the system have been well validated as means for improving performance. However, as mentioned previously, more research is needed to establish whether the performance appraisal component of performance management systems have the necessary validity for administrative decisions that require comparisons between employees.

From the performance management perspective many of the factors that were either ignored or treated as sources of error in earlier research on performance appraisal become quite central determinants of the effectiveness of the appraisal process within the broader performance management process. In a later section we will discuss several of these issues and outline their relevance for performance appraisal when conducted in the context of a performance management process. In discussing these issues, particular emphasis will be given to appraisal within performance management systems.

IDENTIFYING THE PROBLEMS

Clearly, for such a widely practised managerial activity, performance appraisal suffers from many shortcomings. But what are the problems with performance appraisals in performance management systems? We would like to suggest that practitioners confront four broad categories of problems in the design, implementation and operation of performance appraisal systems and that some of these problems are not adequately dealt with in the extant research. These four categories of problems are: technical design; organisation change and development; decision support systems; and strategic integration. Managers must seek solutions to all four sets of problems and must recognise that the solutions to one set of problems will constrain their options in solving other problems.

Technical design problems

The first set of problems are those that deal with the technical design of the appraisal system. Technical problems will arise in the design of the rating instruments and the performance appraisal process. Questions to be addressed in the design of the appraisal instrument will include the number of levels in the rating scale, the specific criteria used to rate performance, the identification of standards and a wide range of other issues. Decisions about the process to be followed will need to consider issues such as the timing and nature of the appraisees’ involvement. Solutions to these problems are important because they influence both the quality of the judgements made and the acceptability of the system to users. If the rating instrument and procedures are seen as valid, then the performance appraisal system is more likely to be accepted by the managers and the employees. Findings from the Measurement School of research have provided answers to most of the technical design problems, particularly those that arise in the design of
performance appraisal forms. The Appraisal Interview School research findings offer advice on the
structure of the appraisal process and the procedures to be followed, particularly in the conduct of
performance reviews and appraisal interviews.

With access to either the literature or expert advice, practitioners should have little difficulty in
finding answers to most of the technical questions raised in the design of rating instruments and appraisal
procedures. However, problems may arise in the specific choices made because of the over-emphasis on
rating accuracy as the criterion for determining the best options in research on rating formats. Also, much
of the research has failed to consider the difference between judgements and ratings, an issue which is
discussed in more length later in this paper. Finally, many of the technical design questions for appraisal
formats cannot be properly considered until the purpose of the appraisal system has been clearly defined
(Landy and Farr, 1983; Marshall, 1993).

Organisational change and development problems

A second set of problems that arise in the implementation and operation of performance appraisal
systems relate to the organisational change and development requirements of performance appraisal
systems. The effective introduction of a performance appraisal system will require multiple skills from the
managers who must complete the rating forms and conduct appraisals interviews. The effectiveness of the
appraisal system will also be dependent upon the skills of staff who are being appraised. The skills
required of managers may include communication skills, observation skills, negotiation skills, coaching
skills and others. The specific skills required will depend upon the design and operation of an appraisal
scheme and the purposes for which it is used. Developing and maintaining the skills required for the
effective operation of performance appraisal schemes may involve a major process of organisational change
and development. In some organisations, performance appraisal schemes will represent a shift in the
organisational culture and require changes in the basic beliefs and values of staff, in addition to the
development of the skills needed for effective appraisals.

The design of appraisal instruments and the uses to which they are put do not of themselves
provide managers with the necessary skills to conduct the processes that surround and support the
performance appraisal rating. These skill requirements and the associated beliefs and values may go well
beyond the requirements of the performance appraisal process, especially when performance appraisals are
introduced as part of a broader strategy of cultural change aimed at enhancing the performance culture of
an organisation (Beer et al., 1985). As such, appraisals can become a carrier of the beliefs, values and
skills of the organisation, providing cues and reinforcements to the appropriate behaviours. Performance
appraisal systems that lack credibility or degenerate into bureaucratic form-filling may undermine the
broader cultural orientation toward performance that is intended. Fitting the performance appraisal system
to the broader organisational context is an issue that requires greater attention and will be discussed in a
later section of this paper.

Decision-support problems

A third set of problems relates to the use of performance appraisal data as inputs to administrative
decisions, such as promotions, pay allocations and job assignments. Performance appraisal schemes are
often introduced as decision-support systems to assist managers to make the necessary discriminations and
other judgements required in administrative decisions. Ensuring the suitability of performance appraisal
data for the decisions that have to be made is often a problem in organisations. As we have said already,
data from performance ratings have consistently been found to lack the necessary discrimination for
decisions such as performance pay allocations and promotions because of leniency effects in ratings.
Linking performance appraisal to administrative decision processes creates an interdependency between the two systems that goes beyond the use of performance ratings as a data input for administrative decisions. The perceived fairness and validity of the performance appraisal system will determine the equivalent perceptions and acceptance of any decisions which are based on the performance appraisal data. Problems can also arise because of the feedback effect that can occur when important decisions are attached to performance appraisal. Judgements required in performance appraisal come to be driven by the prospective decision outcomes and not the other way around as intended. Solutions to these problems may not be in the design of the appraisal formats but, instead, in the decision processes by which appraisal judgements and ratings are converted to decision outcomes such as promotions or salary increases. For example, research might focus on the impacts of committee versus individual decision processes on the quality and acceptance of administrative decisions.

**Strategic integration problems**

When identifying solutions to the technical, organisational change and development and decision-support problems of performance appraisal, it is important to keep in mind that performance appraisal schemes represent a third- or fourth-order set of decisions in the design and management of organisations. Before deciding on options for a performance appraisal scheme decisions must be made in relation to the strategic orientation of the organisation, the specific business and functional strategies that are to be adopted, and the organisational culture needed to support the strategies. Although there are numerous ways by which companies can gain a competitive advantage, one way often overlooked is through human resource management practices (Schuler and MacMillan, 1984). However, for human resource practices to contribute to an organisation’s competitive advantage, systems such as performance appraisal must be derived from or otherwise linked to the strategies of a company.

The key human resource management (HRM) policy goals in writings of the 1980s were those of high commitment, high quality, flexibility and *strategic integration* (Guest, 1990). According to the HRM line of argument, a coherent HRM strategy, based on a performance management system and directed towards the HRM policy goals must be fully integrated into the business strategy of the organisation and fully supported by line management at all levels if high productivity and related outcomes sought by industry are to be achieved (e.g., Miles and Snow, 1984; Porter, 1985; Schuler, 1989; Walton, 1985).

The full strategic integration of performance appraisal systems requires integration at three levels. First, they must be integrated with other HRM systems. The traditional personnel activities of planning, staffing, appraising, compensation, training, labour-management relations and performance appraisal are all integral components of the HRM strategy and must be linked to one another in ways that send consistent signals and reinforcements to staff (Schuler, 1989). Secondly, as previously mentioned, the appraisal process provides the framework for many activities and decisions in the performance management system and, therefore, it should also be integrated with other planning, review and accountability systems, such as business plans and budgets.

Thirdly, the performance appraisal system must be consistent with the business strategies of the organisation. It must identify, develop and reinforce the role behaviours needed of employees for the organisation to achieve its strategic goals. The performance appraisal system can have powerful formative effects on the power structure, workforce skill profiles and corporate culture of an organisation. Appraisals and performance reviews can reinforce hierarchical control and the directiveness of supervision by instituting regular top-down assessments of employees’ behaviours and performances. Alternatively, highly participative, non-standardised systems may be used to decentralise decision-making, increase employee involvement and encourage initiatives in different parts of the organisation. It is important to decide
whether a proposed performance appraisal system should be designed to reinforce current power structures, and their inherent role behaviours, or be used to shift power to those who must hold power in three, five or ten years time (Dunphy and Hackman, 1988).

**Figure 4. Framework for the integration of organisation strategies, employee role behaviours and performance appraisal systems**

<table>
<thead>
<tr>
<th>STRATEGY</th>
<th>EMPLOYEES’ ROLE BEHAVIOURS/ COMPETENCIES</th>
<th>PERFORMANCE APPRAISAL CHARACTERISTICS</th>
</tr>
</thead>
<tbody>
<tr>
<td>COST REDUCTION</td>
<td>Efficiency orientation</td>
<td>Top-down approach</td>
</tr>
<tr>
<td></td>
<td>Repetitiveness</td>
<td>Short-term focus</td>
</tr>
<tr>
<td></td>
<td>High predictability</td>
<td>Individual accountability</td>
</tr>
<tr>
<td></td>
<td>Concern for quantity</td>
<td>Quantitative</td>
</tr>
<tr>
<td></td>
<td>Low error rates</td>
<td>Fixed job duties</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Low employee participation</td>
</tr>
<tr>
<td>QUALITY ENHANCEMENT</td>
<td>Concern for quality</td>
<td>Considerable participation by employee</td>
</tr>
<tr>
<td></td>
<td>Concern for quantity</td>
<td>Individual or team-based approaches</td>
</tr>
<tr>
<td></td>
<td>Customer orientation</td>
<td>Behavioural &amp; process criteria</td>
</tr>
<tr>
<td></td>
<td>Goal commitment</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Problem-solving</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Constant improvement</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Personal responsibility</td>
<td></td>
</tr>
<tr>
<td>INNOVATION</td>
<td>Creativity</td>
<td>Longer-term focus</td>
</tr>
<tr>
<td></td>
<td>Risk taking</td>
<td>Participative approach</td>
</tr>
<tr>
<td></td>
<td>Entrepreneurship</td>
<td>Developmental approach</td>
</tr>
<tr>
<td></td>
<td>Low predictability</td>
<td>Supportive approach</td>
</tr>
<tr>
<td></td>
<td>Flexibility</td>
<td>Qualitative and results criteria</td>
</tr>
<tr>
<td></td>
<td>Highly independent</td>
<td></td>
</tr>
</tbody>
</table>


Properly designed HRM practices such as performance appraisal can provide an effective means for linking business strategies to needed employee role behaviours. For example, if a firm in a service industry is committed to a quality strategy and decides that a team approach by staff will be the most effective way of implementing the strategy, then the performance appraisal system should be designed and implemented to support the cohesiveness of work teams. A system which stressed the accountability of individuals, as most performance appraisal systems do, would be inconsistent with the desired role behaviours of employees and may undermine the quality response strategy. The widespread mimicry in the design of performance appraisal systems in both public sector (e.g., OECD, 1993) and private sector
organisations (Milkovich and Wigdor, 1991; Wyatt, 1989) suggests, however, that organisations are not basing the design of their performance appraisal systems on considerations of strategy or desired role behaviours. HRM scholars are beginning to suggest frameworks for the linking of HRM strategies, HRM practices and desired role behaviours but, at this point, there is little empirical research regarding the validity of these models. An example of such a model for performance appraisal systems, adapted from Schuler (1989), is provided in Figure 4.

The failure to consider decisions about performance appraisal systems in the proper context of the solutions to the higher order problems of strategies and culture will contribute to the dissatisfaction of managers and staff. It is not uncommon in organisations for managers to receive signals from one set of plans that tells them teamwork, co-operativeness and high responsiveness are expected while, at the same time, the performance appraisal system is sending a contradictory set of signals that stress individual accountability, efficiency and other criteria. Proper strategic integration of performance appraisal systems can minimise the problems that arise due to the conflicting and competing demands on a manager from different planning and accountability systems.

While all four problem areas discussed above are important, we suggest that the two areas that require the most attention are decision support systems and strategic integration. Issues that relate to decision support are purpose, judgements versus ratings, and accuracy versus effectiveness; the issue concerning strategic integration is organisational context. A discussion of these issues follows.

THE ISSUES

The purpose of appraisal

Performance appraisals, whether done in isolation or as a component in a performance management system, are done for a purpose. Earlier, we have argued that it is widely believed among managers that performance appraisals are not worth the effort. From this it could be inferred that performance appraisal systems are failing in the purposes set for them. The problem, however, may be due to either poor or conflicting definitions of the purpose of appraisal. There is a growing body of evidence which shows that the purposes of appraisal, that is the decisions or other aims to which the information is put, are the key determinants of the effectiveness of appraisal systems within different contexts (Marshall, 1993; Meyer, Kay and French, 1965; Sharon and Bartlett, 1969; Williams, et al., 1985). Organisations may expend too little effort in defining the specific purposes for appraisal systems and in resolving the conflicts arising between different purposes.

The purposes attached to performance appraisals within performance management systems have several influences. First, they determine the criteria that are used to assess the effectiveness of the performance appraisal system. As we have already pointed out in our previous discussion, the types of criteria used by different research schools for performance appraisal have been quite different. Secondly, the purposes of appraisal reflect the interest of different parties such as personnel departments, individual managers and ratees and, therefore, are central to consideration of politics and conflicts that may arise around the performance management process. The designers of performance appraisal formats often base their choices on what they believe will give them the most accurate and valid rating. Managers who conduct appraisals want a system that is easy to use and enables them to motivate their staff.

Clearly defining the purposes of an appraisal system and designing the system to meet those purposes can obviously enhance the system’s effectiveness if the actors involved understand and accept the stated organisational purposes. However, the purposes of performance appraisal can be viewed from several different, potentially conflicting perspectives. Figure 5, from Murphy and Cleveland (1991), shows lists
of possible purposes from an organisational perspective and from the individual manager’s perspective. The potential for conflict between these two sets is apparent, without the addition of the goals of the employee who is to be rated.

Managers may also pursue goals such as increasing their power with subordinates, enhancing their own position with superiors or maximising their own feelings of competence. This is not to imply that all participants in performance appraisal systems are totally self-interested or cynical about the organisational objectives. However, the reality of political relationships between individuals in organisations cannot be ignored (Longenecker, Sims and Gioia, 1987). The achievement of political goals, through performance appraisal or other means, may increase the influence of the manager and, thereby, make him or her more effective.

Figure 5. Purposes of performance appraisal viewed from the organisational and managers’ perspectives

**Organisational Goals**
- Distinguish good from poor performers
- Provide developmental feedback
- Document and justify decisions
- Provide input for other personnel decisions

**Managers’ Goals**
- Avoid interpersonal tension in workgroup
- Motivate poor performer
- Maintain current level of motivation in good performer
- Achieve or restore equity in workgroup
- Use ratings to convey your effectiveness as a supervisor


It is also possible for managers to be motivated to achieve the organisation’s purposes such as improved performance, but to use the performance appraisal system in unintended ways to achieve these aims. For example, an inflated rating which is recognized as such by both the employee and the manager may be given in order to motivate the employee to a higher work standard. The implicit signal the manager sends and which may be made explicit in direct conversations with the employee is that, "I will not punish you but in return I would expect higher levels of work output". Again, such interpretations become obvious when performance appraisal is viewed as a communication process within the context of a broader performance management system and not just as the integration of information into a rating judgement.

Murphy and Cleveland (1991) have highlighted the differences between those purposes of appraisal which involve comparisons between persons, such as pay awards and promotions, and those which focus on a single person, such as motivation, feedback and personal development, and do not require any comparisons between people. From this distinction flows other issues such as the publicity of the rating and the requirements of a common yardstick for comparative purposes, which we will be taking up in later discussion. The use of appraisal ratings for comparative purposes requires a clarity of standards
and justification for ratings which can withstand a challenging public scrutiny. The widespread incidence of leniency effects shows how few systems adequately meet these requirements.

The purposes of appraisal are also related to the contexts in which the performance management system operates. This will include the behavioural styles of the different participants, the organisation’s structure, strategy and culture, and may also extend to the external legal environment. For example, in the United States there is a strong legal environment surrounding the design and operation of performance appraisal systems and documentation for legal purposes is a major purpose of appraisal systems (Cleveland, Murphy and Williams, 1989). It is also the case that, in the United States, legal requirements have dictated the design of appraisal systems and the emphasis on the accuracy of ratings by the various schools of research.

The justification required as a result of the legal imperatives relating to performance appraisals in the United States may actually limit the suitability of performance appraisals as decision-support systems. Precedents from case law have resulted in the increasing use of performance appraisal information to justify personnel decisions that require comparisons between people and discriminations in the allocations of opportunities and rewards (Cleveland et al., 1989). At risk in these circumstances is that the justification determines the actual decision, and not the other way around. Appraisal ratings may be based on the quality of justifications that they provide as inputs into personnel decisions. The vast amount of evidence regarding leniency effects in performance appraisal systems that are used for purposes of pay and promotion decisions is evidence of how the pressure of public justification can undermine the quality of appraisal data provided for the personnel decisions (cf. Bernardin and Buckley, 1981; Landy and Farr, 1980, 1983; Latham, 1986; Murphy et al., 1984).

The nature of conflicts that may arise as a result of the differing purposes of performance appraisal and techniques for their resolution have not been well researched. Several levels of conflict may arise in performance appraisals and different techniques may be needed for the resolution of each type of conflict. For example, conflicts may arise between the different organisational purposes attached to a performance appraisal system. The existence of these intra-role conflicts for managers (e.g., Figure 2) has led to suggestions that appraisal activity for purposes such as counselling be separated from performance ratings which are used for decisions such as promotions and pay (DeVries, 1983; Meyer, Kay and French, 1965). Separation of activities can diminish this role conflict. Conflicts may also arise between the organisational purposes and managers’ purposes (Figure 5), and between the demands of different planning and accountability systems. The latter conflict is one of the issues to be resolved in solving the decision-support and strategic integration problems of performance appraisal systems referred to earlier. Another type of conflict is that between individual participants in appraisals. For example, research has shown that staff want pay discussed during appraisal interviews while managers often do not (Mohrman, 1986), and that managers tend to select one purpose over others in their conduct of the performance appraisals (Bjerke et al., 1987).

**Judgements versus ratings**

Although it is well understood in practice, it is often not acknowledged in performance appraisal research that performance judgements and performance ratings are not the same. Performance ratings are a public record of one individual’s judgements of another’s performance and, as such, they represent a form of performance feedback. Judgements are private records about another person’s performance, and therefore are more likely to be a direct consequence of the observation and integration of information regarding that person’s performance. Ratings introduce a range of motivational factors that influence public presentations and which go beyond the demands of a privately recorded judgement. Therefore, the rating of an employee presented by a manager may also reflect his or her confidence as a manager in areas such as public
speaking or conflict management. Managers are expected to justify performance ratings. Private judgements because they may go unnoticed by others do not demand the same level of justification (Mohrman and Lawler, 1983). An accurate private appraisal of an individual’s performance by a manager by no means guarantees that an accurate rating will enter the formal performance appraisal system.

Understanding the relationship between judgements and ratings is important for decision-support purposes. Ratings can be viewed as the output of goal-directed behaviour by the manager and will therefore reflect the purposes that the manager hopes to achieve by this behaviour. As we have already argued, the goals of a manager in providing a rating may not be consistent with the organisational goals for the performance appraisal rating (refer to Figure 3).

It is also the case that ratings of individuals may reflect a wide range of judgements and not just the evaluation that they are intended to convey. Ratings, as a record of evaluation judgements, are intended to describe the performance levels of an individual. However, in the cognitive and motivational milieu surrounding the reporting of ratings, other factors come into play when a manager is making these judgements. These include diagnostic and predictive judgements. Diagnostic judgements are those in which the manager makes attributions about the causes of an employee’s performance level. The diagnosed causes of performance can have a range of differing implications for the actions required and for the ratings of performance. For example, if a manager attributes poor performance to factors that are beyond the control of the individual employee he or she is more likely to feel that an accurate rating of the performance level is unfair and to inflate the performance ratings of that person. If this logic is correct then a general explanation for the consistent leniency in ratings may be that managers lack confidence in the diagnostic judgements regarding the true causes of performance and their ability to attribute those causes to factors that are controlled by individual employees. Diagnostic judgements may influence evaluations in other ways. When the cause of a performance outcome is seen as temporary or aberrant in nature and not directly related to the individual, evaluation judgements may be adjusted to reflect the expected steady state.

Predictive judgements about a person’s future performance can also influence evaluations of that person’s past performance. When prior experience with an employee leads to expectations that they will perform better or worse in the future, these expectations influence a manager’s categorisation and storage of information about the employee (Foti and Lord, 1987). When rating the employee, a manager’s judgements will be biased up or down by the expectations they hold for that person’s future performance. Predictive judgements will also enter into evaluative judgements when a manager is seeking to use the rating to signal to the employee to remain committed and motivated on the job. Predictive judgements are also influenced by diagnostic judgements; for example, if a particular employee performs badly in a given work period, and this is attributed to temporary factors which are now no longer seen to be operating, then the manager may expect improvements in future performances by the employee. These two judgements may then influence the rating given to the employee.

The impact of these different judgements on performance ratings may diminish the validity of ratings as an input to administrative decisions and create uncertainty about the true criteria being used when rewards, promotions and the like are handed out. More research is needed on the relationships between different types of judgements that managers make when rating an employee’s performance and how these judgements impact on the rating. Decision processes that allow for the introduction of predictive and diagnostic judgements without invalidating the recorded evaluative judgements also need to be developed and studied. Further research is also needed to examine those specific factors that lead to disconnections between the private evaluative judgements and recorded performance ratings by managers. An important question for study is "how can the appraisal rating and administrative decision processes be structured and linked together to ensure managers are confident enough to give accurate reports of their private judgements?"
A problem-focus approach related to this research could consider, for example, how the recording of ratings for, say, performance pay decisions by a committee compare to those prepared by individual managers who must confront an employee. Judgements recorded by a committee lead to a diffusion of the responsibility for the particular rating and therefore reduce the pressure on the individual manager to inflate the rating in order to maintain the support and commitment of the individual employee.

**Accuracy versus effectiveness**

Related to the purposes of appraisal are the criteria used to determine the effectiveness of a performance appraisal system. As we have already discussed, different research schools have used different criteria in determining the effectiveness of performance appraisal. These have included criteria such as the validity or accuracy and reliability of appraisals, plus the practicality, use and acceptability of appraisal forms to staff who must appraise others and those who are appraised. The criterion which has dominated the design of performance appraisal systems, particularly in the Measurement and Social Cognitive Process schools, has been accuracy of ratings. In these two schools of research, accuracy is a narrowly defined effectiveness criterion which focuses on the degree to which performance ratings are free from rater errors such as halo effects or leniency. In the Performance Management School effectiveness is a more encompassing criterion which includes acceptance of the system and contributions to the performance of the individual, group and organisation.

Performance appraisal systems chosen for their accuracy may not necessarily contribute to the effectiveness of performance management systems and, in some cases, may even detract from effectiveness. The argument that accuracy is the expected prerequisite for a performance appraisal system to contribute to the performance of organisations is based on several assumptions: that the primary use of performance ratings is for administrative decisions; that administrative decisions depend on performance ratings and not other sources of data; and that greater accuracy will usually lead to more effective administrative decisions. As we have already argued, however, performance ratings serve a range of purposes in performance management systems. For many of these purposes, such as motivating staff, our concern with accuracy may run counter to the aims being sought by the manager.

Reliability, validity and accuracy are relevant criteria when performance appraisal is used for decision support purposes (Sulsky and Balzer, 1988). However, it is also the case that the level of accuracy required for many administrative decisions is far less than that which is sought in performance ratings. Accuracy in performance rating is often only important up to the point that it leads to and supports the correct decision. Broad discriminations may be enough for correct decisions. Promotion decisions for example, only require a level of discrimination that divides employees into promotable and non-promotable groups. A less than perfect appraisal system may generate this dichotomy with the same level of accuracy as a scheme which can validly and reliably differentiate between multiple levels of performance (Feldman, 1986). The pursuit of accuracy may produce unnecessary over-discrimination between individual staff members.

Lord (1985) has identified two different types of accuracy in performance appraisals. The first is behavioural accuracy which refers to a manager’s ability to accurately recall performance against specific criteria in behavioural terms. The second type is classification accuracy, which refers to a manager’s ability to classify employees into the relevant categories such as promotable, or not promotable. The specific types of accuracy required will depend upon the purposes to which the appraisal information is to be put.

For managers, the communication of highly accurate ratings may be detrimental to the other purposes of performance appraisal such as improving the performance of a work unit. The reactions of
raters, the perceived fairness of rating and discouragement of employees are all factors that a manager must take into account when deciding on the rating to be given and communicated to a staff member. Rating errors, therefore, may not reflect an inability to discriminate but may indicate that a manager was unwilling to discriminate amongst staff. Failure to discriminate in the ratings communicated to staff may be highly adaptive behaviour from a manager’s point of view when he or she must maintain the commitment and support of staff. The relative emphasis to be given to accuracy versus acceptability of ratings will depend, among other things, upon the nature of the task and the levels of interdependence between the manager and the staff member. Where effective performance of a task is highly dependent upon the ongoing commitment and co-operation of staff members, acceptability will be a relatively more important determinant of the effectiveness of performance appraisals.

Accuracy will also depend upon the level of detail required in appraisal judgements. In teamwork situations individuals may be less concerned about the accuracy of ratings for individuals and more concerned about the overall performance of the team. Aggregate measures of performance, such as team-based measures, are generally more quantitative and more reliable than individually focused appraisals. The more interdependent the work performed within a team, the more difficult it becomes to isolate and appraise the contributions of individuals. Therefore, individual ratings may be used for their encouragement and feedback effects and less for their accuracy in discriminating between the relative performance level of individuals.

The relative emphasis on effectiveness versus accuracy will also influence the training programmes used to support the introduction of performance appraisal systems. Performance appraisal training frequently includes segments which are intended to make raters more aware of rating biases, such as halo, leniency, similar-to-me and contrast errors, which are believed to lead to inaccurate ratings (Wexley and Latham, 1981). There is evidence that rater-error training does reduce the incidence of particular errors but the question is now being asked whether such training is really beneficial in terms of the effectiveness, or even the accuracy, of performance appraisals. This questioning is a result of the consistent finding of a low positive correlation between halo, the most frequently researched source of rating error, and rating accuracy (Borman, 1979; Borman and Kerney, 1976; Hedge and Kavanagh, 1988; Murphy and Balzar, 1981). If an error such as halo is positively related to accuracy in ratings, training efforts which attempt to improve accuracy by removing errors are of questionable value. Training which develops the skills required to conduct appraisal-related activities as part of the day-to-day role of a manager, although more costly to conduct, is now generally considered more effective than rater error training. Frame-of-reference training, mentioned earlier, is also proving a more effective means of developing common rating standards among managers.

Any training aimed at improving the capabilities of the manager may be an ineffective means of improving the accuracy and effectiveness of performance appraisal without changes in the incentives for managers to conduct effective appraisals. Deficiencies in ratings, such as leniency, are often a product of the manager’s (lack of) motivation to provide accurate ratings and not just a function of his or her capacity or the design of the rating instrument (Banks and Murphy, 1985). Managers often have very few good reasons to give accurate ratings. In this context, leniency in ratings is not an error but a functional behaviour by the manager that enables him or her to do their job more effectively. Among the incentives that influence a manager’s tendency to provide lenient ratings in the management of performance are negative reactions by subordinates to low ratings, such as a lack of co-operation, and a lack of significant rewards or other positive outcomes for the manager. (DeCotiis and Petit, 1978; Mohrman and Lawler, 1983). Techniques for minimising these effects might include the use of committees which diffuse responsibility for specific appraisals when they are connected to decisions that may be rejected by the employees.
Unless there are clear accountability mechanisms and rewards for the effective conduct of performance appraisals, managers will not be motivated to achieve the desired organisational outcomes. Specifically, there is evidence that increasing the accountability of managers for their performance ratings will enhance the accuracy of those ratings (Klimoski and Inks, 1990; Marshall, 1993). Accountability mechanisms which require justification of ratings to senior management and provide feedback on the accuracy of ratings may also help to increase the dispersion of ratings (Ford and Weldon, 1981; Rozelle and Baxter, 1981). Another technique for increasing a manager’s accountability is to include the effective conduct of performance appraisals as a goal in his or her own performance appraisal. Evidence on the effects of these and other techniques would be a useful contribution to performance appraisal practices.

Other important effectiveness criteria to consider when designing and implementing a performance appraisal system are the acceptability and the practicality of the system. Research on acceptability has focused on factors that affect the perceptions of fairness and accuracy of appraisal systems (Bernardin and Beatty, 1984; Marshall, 1993). In general, this research indicates that performance appraisals are most likely to seen as accurate and fair, and therefore acceptable, when they are conducted frequently by informed managers against relevant criteria and allow the employee some input into the process. Related to the acceptability and effectiveness of performance appraisal systems is their practicality.

Performance appraisals are often considered impractical and, therefore, a waste of time by managers because of their failure to capture adequately the key contributions of employees or because they do not fit with other demands on a manager’s time. The perceived practicality of a system will be related to the degree to which performance appraisal is integrated with other planning, review and accountability systems, such as corporate and work unit plans and budgets. Minimising the level of duplication of work required for different systems will change the practical value of performance appraisal.

The organisational context

If performance management systems are to have lasting effects, they must be consistent with the way the organisation is managed, with other structural features of the organisation and, in particular, with the strategies of the organisation. Organisations vary on a complex range of factors including size, technology, degree of bureaucratisation, nature of staff, strategies pursued and work cultures. Many of these factors could be expected to impact on the observability of the work, and the agreed standards and judgements of work within organisations (Lee, 1985). Additionally, they affect the managerial systems and styles adopted by organisations. Despite the array of factors which can lead to differences between organisations, there is a marked similarity in the types of appraisal systems being adopted in public and private sector organisations across a range of OECD countries. In particular, at the managerial level, MBO-based appraisal systems with five-level rating scales have become the standard (OECD, 1993). This, in itself, is not necessarily a bad thing. There is widespread evidence for the effectiveness of goal setting and feedback, and that evidence points to the universalistic benefits of systems that promote higher levels of goal setting and feedback within organisations (e.g., Locke and Latham, 1990). However, the specific performance management systems adopted within organisations will need to vary if they are to reflect the differences in managerial styles and other factors related to the context of different organisations.

Very little field research has been done to establish the effectiveness of different systems across different organisational contexts. This is an area of considerable interest to those who must ask the question: "what is the correct performance appraisal system for my organisation?" As was previously argued, the answer to this question often will come from the answer given to higher order questions regarding the strategy and the culture that management wishes to implement within an organisation.
A key consideration in the adoption of a performance appraisal system is the nature of tasks performed by those to be appraised. As the complexity and interdependence of an organisation increases, the identification of individual work contributions becomes more difficult. Interdependence may be a product of the flow of work between staff or due to deliberate efforts at co-operation between staff in order to do a job or set of jobs more effectively. Under the former, isolating individual efforts for appropriate purposes may require the redesign of tasks so that individuals complete whole tasks with identifiable outcomes. Whether such work redesign is worthwhile will depend upon the technology and related costs of the work done. When interdependence is a result of co-operation to get the job done more effectively, individually focused appraisals undermine the effectiveness of the work group. This assumes, of course, that co-operation does actually contribute to performance and is not simply a social activity that drains resources away from productive effort.

When co-operation contributes to performance, performance appraisal decisions may be more appropriately used for motivational purposes that focus on the individual, such as providing feedback and identifying training needs and solving problems. Uses of the data for comparative purposes, as required in decisions such as promotion or salaries, are more difficult to sustain and justify because of the difficulty in identifying the contributions of individuals. The levels of interdependence between staff have been found to affect ratings (Liden and Mitchell, 1983). Much of the managerial rhetoric of the 1980s emphasised participative work cultures, teams and empowerment of individuals through delegation. Great stress was placed on self-development and self-management. At the same time, however, there was a marked growth in the use of performance appraisal systems which stressed individual accountability and which were used for discriminatory decisions such as those required in performance-based pay (Guest, 1990). Further research is needed on the impacts of different types of appraisal systems on levels of co-operation and other aspects of work cultures.

Research is also needed on the relationships between performance appraisals and other human resource management systems such as the recruitment and selection systems. Proper integration of human resource management systems will ensure that workers receive consistent goals from the different systems. It is often the case that the norms of practice developed within different systems are in conflict with one another. For example, selection standards built up through time may affect the distributions of performance ratings (Murphy and Cleveland, 1991). In particular, high selection standards can lead to a general reluctance to rate people as average. That is, when an organisation has high selection standards and reinforces the elitism or high quality of its staff as part of its induction process, it then becomes very difficult for the appraisal system to impose a distribution of performance ratings on those staff which includes average or below average ratings.

Yet another aspect of organisational context requiring study is the relationship between performance appraisals and the dominant management styles of an organisation. Clearly there is a need to ensure that performance appraisal processes are consistent with the way an organisation is managed. The choice of who should conduct appraisals, for example, needs to be related to the culture of the organisation. Organisations that emphasise high involvement and high commitment of employees may wish to use self-appraisals to emphasise the role of the individual and to de-emphasise the role of the hierarchical structure. Organisations with self-managed work teams may rely more heavily on peer appraisal to reinforce the desired culture. Peer appraisals may also be more appropriate for tasks which involve high levels of interdependency. As previously mentioned, appraisals that focus on individual accountability may be a poor fit with both the desired organisational culture and the task.

In MBO-based performance management systems, the risks associated with goal-setting may be best matched to a highly supportive management style which is not commonly associated with a stress on individual accountability. Pushing individuals to set challenging goals increases the likelihood of their failing to achieve them. Therefore, goal-setting systems need to manage the discouragement that may
follow when individuals fail. A system which combines encouragement to set challenging, even difficult goals, needs to be balanced by a supervisor who plays the role of resource provider and coach.

The context of an organisation may also be used as a means of improving performance appraisal. That is, changes to the context rather than the appraisal instrument or individual training may be used to make it easier for managers within the organisation to conduct performance appraisals. Our earlier comments on the incentives for managers to conduct effective appraisals are one example of adapting the context to the needs of performance appraisal. Other examples could include the redesigning of jobs and the development of measurement systems, such as customer satisfaction surveys, which allow for easier assessment of levels of performance.

Further work is needed to understand how performance appraisals can influence organisational contexts in unplanned ways. Over time the performance appraisal process can become an important part of the culture within an organisation. Communication rules may develop around the regular performance reviews and annual performance appraisal within a performance management system (Schall, 1983). For example, it may become a communication rule that, in the interest of harmony, a supervisor does not directly criticise or challenge a worker’s claims about his or her own performance. Another aspect of culture that can become embodied in the performance appraisal process is the ritual associated with process. The annual cycle of performance appraisal punctuates different periods of the year with steps in the flow of activities and paper work associated with the process. Therefore, for example, the annual performance appraisal becomes a time for clearing away last year’s tasks and signalling an end to those activities and moving on to a new set of objectives and tasks (Schein, 1985). Understanding the relationships between the organisational context and the effectiveness of performance appraisals is an important step in achieving strategic integration of human resource management systems.

CONCLUSION

In the previous section we have suggested several areas for future research in performance appraisal which will deal more directly with the practical problems that arise in the use of performance appraisals within organisations. These suggestions are not exhaustive but have been based on our perceptions of the issues as they relate to the practice of appraisal and not on theoretical models that might be used to guide the research. Theoretical models are an important source of insights and explanations about performance appraisal activities. However, the starting point for the research needs to be the problems that require solution in the practice of appraisal.

Until recently, recommendations for improving performance appraisal have focused on rater training, instrument design, and policy recommendations. It is now generally accepted that many recommendations in these areas, particularly those based on research in the Measurement School, have reached the limits of their usefulness and that a broader range of issues needs to be addressed if research is to provide improvements in the practice of performance appraisal. At the same time, performance appraisal has been incorporated into performance management systems and needs to be studied in that context to be properly understood.
BIBLIOGRAPHY


CARROLL, S.J. and SCHNEIER, C.E. (1982), Performance appraisal and review systems: The identification, measurement and development of performance in organisations, Scott, Foresman, Glenview, IL.


DeVRIES, D.L., MORRISON, A.M., SHULLMAN, S.L. and GERLACH, M.L. (1986), Performance appraisal on the line, Centre for Creative Leadership, Greensboro, NC.


35


RODGERS, R.C. and HUNTER, J.E. (1989), "The impact of management by objectives on organizational productivity", unpublished manuscript, School of Public Administration, University of Kentucky.


SILVERMAN, S.B. and WEXLEY, K.N. (1984), "Reactions of employees to performance appraisal interviews as a function of their participation in rating scale development", Personnel Psychology, 37, pp. 703-710.


