CHAPTER II

Review of Literature
CHAPTER II

LITERATURE REVIEW

The literature review conducted by the researcher for the purpose of present study is presented in this Chapter in two sections. While this section provides a review of literature on the implications of assessing R&D scientists, the second section deals with the studies made on performance appraisal per se.

2.1. Section A

The principal outputs of research are knowledge and know-how (Hare and Wyatt, 1992), which may be presented in various forms: information in articles and books, innovative prototypes, patents, consultancies or training of personnel. This output is generated by research activities funded by the government or supported by self-financing through sale of technologies. Financial resources within the Organisation are invested in various ways: setting up laboratories, purchasing equipment, books, journals, missions, and consumable materials, offering scholarships and fellowships, etc. Public funding of research can be justified in two ways: the first is the production of a cultural product that educates the public (the discovery of a star, a new animal species, etc.). Although this is important, it is not the type of subject used when supporting public research. The policy used by governments is that scientific research is an investment that gives a return in terms of scientific-technological progress in terms of increased competitiveness, improved standards of living and therefore greater wealth produced by the nation.
Whichever orientation is considered, knowledge is generated by public and private research laboratories. Public laboratories are currently under pressure from the government to increase their efficiency. Moreover, public research bodies and universities are affected in developed countries by a new culture defined as performance-oriented (Ball and Wilkinson, 1994). In fact, in the mid-nineteen eighties, the UK set up a commission to analyse the effectiveness of universities. The commission’s report is known as the Jarratt Report, and it recommended a number of ways of improving the management of the scientific-academic organizations such as (1) an integrated approach to decision making; (2) development and use of a series of performance indicators that consider input and output in order to improve efficiency and permit comparison between various scientific institutes (Jarrat, 1985). Another report known as the Morris Report (1990) recommended universities and colleges to develop a set of performance indicators that would promote the interests of the scientific sector. These indicators, unlike those described in the previous report, referred to scientific fields and not to institutes or laboratories. West (1986) applied a series of performance indicators to the strategic planning process at the University of Strathclyde using a polydirectional matrix within which each department was positioned. The performance indicators for the laboratories were both internal (which fixed the position of the organization on the vertical axis), and external (fixed the position of the organization on the horizontal axis). The planning committee for this method was so firmly convinced of its effectiveness that one of the members voted to close his own department (Ball and Wilkinson, 1994). The DES (Department of Education and Science, 1991) published a report on the use of performance indicators in higher education and research showed that many institutions
that used them improved their organizational management and the attained strategic goals.

The research laboratory that produces goods and services with its input and production processes (of scientific activity) is a function of $P = \text{persons, } G = \text{goods, } O = \text{organizations}$. Mathematically it is possible to say

$$\text{Output of Research Laboratory} = f (\text{Persons, Goods, Organisation ...})$$

The research laboratory is a system, which is not identified by the sum of the three elements, but is the result of their combination according to certain rules that make up the laws of the organisational system. Many studies on the measurement of the performance of research bodies consider only the bibliometric (Narin and Hamilton, 1996), technometric (patents) or productive indices, which only give partial and non-systematic indications of the productivity of the scientific laboratory.

2.1.1. Use and impact of measurement results

Knowledge of measurement areas and how the system operates provides a necessary understanding of the performance measurement context before focusing on the central part of research, namely the use and impact of the results of performance measurement. Before analyzing the literature focused on these issues, it is important to clarify how these notions relate to each other and how they relate to different measurement areas discussed previously. At a first glance, the notions of area, use and impact might seem overlapping or even interchangeable. For example, instead of considering enhanced staff motivation as an impact of performance measurement, the
performance measurement system could be exactly used for that purpose (Kerssens-Van Drongelen and de Weerd-Nederhof, 1999).  

However, as Loch et al. (1996) argue, these notions reflect distinctive and linear phases of the performance management processes. The areas inspire the intentionality of managers to set specific objectives for the measurement process. The use captures how, in practice, metrics are employed in managerial processes and how the results of measurement drive particular actions. The impact, finally, reflects the resulting outcomes ex post. One of the central applied problems in performance management is to minimize the gap between intention and outcome (Loch et al., 1996). The means for doing so resides in the actual managerial levels that represent the use of measurement results.

Literature that focuses explicitly on the use and impact of performance measurements systems is reviewed in the following. Note that, according to the above, one may search for ‘how, in practice, metrics are employed in managerial processes and how the results of measurement drive particular actions’ when reviewing uses, and for ‘resulting outcomes ex post’ when reviewing impacts.

Kerssens-van Drongelen and Bilderbeek (1999), based on a study of 44 companies, identified a range of uses or purposes (as they call them) of measurement that they clustered in four categories with respect to the organizational level, i.e., where the focus of performance measurement is: individual, team, R&D department and company (Table 2.1).
Table 2.1.
Uses (or purposes) of performance measurements at four different organizational levels, adapted from Kerssens-van Drongelen & Bilderbeek, 1999.

<table>
<thead>
<tr>
<th>Purposes of individual performance measurement</th>
<th>Purposes of team performance measurement</th>
<th>Purposes of departmental performance measurement</th>
<th>Purposes of company level performance measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision-making about promotion prospects</td>
<td>Decision-making to dissolve the team</td>
<td>Decision-making about reorganization</td>
<td>Decision-making about reorganization</td>
</tr>
<tr>
<td>Decision-making about salary</td>
<td>Progress control/correction</td>
<td>Correction</td>
<td>Correction</td>
</tr>
<tr>
<td>Decision-making about project participation</td>
<td>Learning/continuous improvement</td>
<td>Learning</td>
<td>Learning</td>
</tr>
<tr>
<td>Correction</td>
<td>Assignment of new projects</td>
<td>Assignment of resources</td>
<td>Assignment of resources</td>
</tr>
<tr>
<td>Assignment of bonuses</td>
<td>Assignment of bonuses</td>
<td>Assignment of new projects</td>
<td></td>
</tr>
</tbody>
</table>

These 19 uses could be regrouped in four categories:

1. use of performance measurement results for Personnel Evaluation, Promotion and Incentives (promotion prospects, salary, project participation, bonuses);
2. use of performance measurement results for Resource Allocation (project participation, forming/dissolving teams, assignment of new projects and of resources);
3. use of performance measurement results for Control/Correction (control, correction, reorganization); and
4. use of performance measurement results for Learning/Continuous Improvement.
These categories correspond closely to the four main functions of R&D performance measurement identified by Loch and Tapper (2002): Evaluation and Incentives, Alignment and Prioritization, Operational Control, and Learning and Improvement. What the category 'alignment and prioritisation' adds compared to 'resource allocation' above is the deployment of strategic intent through measurements, or, as Loch and Tapper put it, 'cascading measures from strategy' (2002, p. 186).

Kerssens-van Drongelen and de Weerd-Nedehof (1999) developed a slightly different but complementary framework, where they also define the notion of function of performance measurement systems as ‘... how the performance measurements are used in the organisation's control and improvement processes’ (p. 401, emphasis added). Based on interviews in nine companies from different industrial sectors, they identified two additional uses that differentiate from the above:

1. use of performance measurement results for diagnosis by management, and
2. use of performance measurement results for communicating objectives, agreements and rules.

The use for diagnosis purposes is concerned with the measurement system providing timely insights into (possible) deviations from objectives and into factors that can affect such deviation (Kerssens-van Drongelen and de Weerd-Nedehof, 1999). This use naturally integrates control and precedes a use for correction and/or for resource allocation. The use of performance indicators in communication and how this is related to motivational purposes has been identified by other scholars as well. Based on the
balanced scorecard literature and an in-depth case study of the process of developing and implementing a comprehensive performance measurement system for a research group, Loch and Tapper (2002)\textsuperscript{11} observed that a fundamental use of R&D performance measures is to focus attention and to communicate the ambition and priorities of the organization. Hence, the communication use is related to alignment/prioritization that precedes resource allocation. In the same vein, Busby and Williamson (2000)\textsuperscript{12} found that the performance measurement system clarifies the goals of the R&D organization and contributes to the development of a common language by encouraging a consensus on these goals.

Based on the above, the uses of performance measurement results to serve as a roadmap for future research: 1. use of performance measurement results for communicating objectives, agreements and rules, thus quantifying and possibly justifying strategy; 2. use of performance measurement results for defining corrective actions based on diagnosis and control; 3) use of performance measurement results for allocating resources in including forming/dissolving teams and assigning projects, thus implementing strategy and/or corrective actions; 4) use of performance measurement results for deciding on individual promotion, salary increase and other incentives in order to facilitate the members of teams to focus on prioritized areas and act towards optimized goal fulfillment; and 5) use of performance measurement results for learning/continuous improvement of the process.

Performance appraisal is the administrative device used to plan and control the assignment of work and to communicate to the employee how well or poorly it is completed. Additionally, if used effectively, performance appraisal can be used as a
motivational tool in that future performance goals can be established for the employee and that knowledge of results can be provided via performance appraisal. Furthermore, the appraisal system can be used to provide the basis of reward such as pay increase and promotion.

Economic compensation for R&D personnel is an important part of motivating R&D scientists and engineers to work more productively. Thus, it is important to design an effective compensation system (Coombs and Gomez-Mejia (1991)\textsuperscript{13} R&D managers must have a fair and effective mechanism to measure R&D performance. An R&D performance measurement system perceived by the R&D personnel as fair and effective is essential for them to feel satisfied with their compensation: when satisfied with their compensation, they feel a commitment to their job, which in turn translates into high R&D performance (Hellriegel et al., 1998)\textsuperscript{14}. In fact, research on R&D performance measurement could shed light on designing an effective R&D compensation system. Much of the research had focused on developing conceptual models (Gomes-Mejia and Welbourne, 1988)\textsuperscript{15}. One may approach the issue from the R&D managers’ perspective, taking into account ideas of the R&D personnel who are most affected by the performance measurement system in the end. From two separate angles, one may ask the primary question: ‘Who should measure the performance of R&D scientists and engineers?’ and ‘What criteria should be used to measure the performance?’ Before reviewing literature on the performance metrics, one needs to justify why it is important to have an effective performance measurement system in the first place. To this question, Cordero (1990)\textsuperscript{16} responded by establishing a substantive relationship between systematic performance measurement systems and R&D performance. That is, the basic premise is
that an effective performance measurement system is a necessary condition for R&D productivity. Firstly, several researchers have focused on the unit of analysis. Collier (1977) was among the first who postulated that project-level metrics be deployed as key performance evaluation criteria, followed by such researchers as Schainblatt (1982), Patterson (1983), and Cooper and Kleinschmidt (1995). In addition, individual-level metrics was suggested as the performance evaluation criterion by Schainblatt (1982) and Wilson et al. (1994), while team-level metrics was suggested by Stahl and Steger (1977) and Moser (1985), and firm level metrics by such scholars as Cordero (1990), Loch et al. (1996), and Werner and Souder (1997).

Regarding the issue of 'who should measure the R&D performance', Whitley and Frost (1971) highlighted the role of the researcher himself and his peers during the evaluation process, concluding that it is necessary to involve multiple evaluators rather than just the R&D project directors and R&D center chiefs. By emphasizing subjective elements such as 'self-esteem', Keller and Holland (1982) underlined the delicate nature of the R&D performance evaluation process. In a similar vein, Werner and Souder proposed multiple evaluators such as the researcher himself, his peers, and external stakeholders. For the issue of 'with what criteria to measure', there are several references, which explored the impact of more qualitative measures on developing a desirable performance measurement system. Barnowe (1975) suggested 'leadership' as a critical performance measurement criterion, while Werner and Souder regarded 'involving diverse evaluators such as the researcher himself, peers, and external auditors as means to utilize 'subjective measures'. Griffin and Page (1993) scored more comprehensive perspectives. Brown and Gobeli (1992) as well as Meyersdorf and Dori
put forth that R&D managers should incorporate the TQM concept into the development activities. Werner and Souder presented an integrated measurement system, which combines both qualitative and quantitative metrics at the same time. Griffin (1997) was among the first who suggested that it is important to use the method of ‘process analyses’ in evaluating R&D performance. Dressler et al. (1999) suggested that CSR (cost savings ratio) be used to measure R&D performance. It is intriguing to note that Hauser and Zettelmeyer (1997) proposed that the performance measurement system should be tailored according to the R&D type in point, e.g., research, development, or engineering. Furthermore, they suggested that the performance metrics should be differentiated across different R&D types.

It is necessary to consider different types of R&D activity. Leifer and Triscari (1987) decomposed forms’ innovative= inventive activities into two, research and Development. Some researchers tried to classify R&D into three; Drongelen and Cook (1997) grouped R&D activities into basic research, applied research, and experimental development, while Hauser (1998) grouped them into research, development and engineering. Others suggested classifications based on more than three groups. For instance, Werner and Souder classified R&D into four levels: basic research, applied research, product development and manufacturing process R&D, while Pappas and Remer (1985) into five levels: basic research, exploratory research, applied research, development and product improvement.

Taking the most common elements from these classifications, Bowon Him and Heungshik Oh (2002) suggested a three layer R&D classification basic, applied and commercial R&D. Table 2.2 summarizes the characteristics of each R&D type.
Table 2.2
Different R&D types

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Basic R&amp;D</th>
<th>Applied R&amp;D</th>
<th>Commercial R&amp;D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Experimental research</td>
<td>Core technology development based on basic R&amp;D. Basis of commercial product</td>
<td>Commercial product development. Evaluated by market success</td>
</tr>
<tr>
<td></td>
<td>Observation of fact requires individual research and concentration</td>
<td>Development of substitutes for Freon refrigerant</td>
<td>Development of energy-saving multi-purpose refrigerator</td>
</tr>
<tr>
<td>Time frame</td>
<td>Over 10 years</td>
<td>5–10 years</td>
<td>1–5 years</td>
</tr>
<tr>
<td>Applicability</td>
<td>High. Generic</td>
<td>Medium</td>
<td>Low. Applicable only to specific products and processes</td>
</tr>
<tr>
<td>Example</td>
<td>Research on the fundamental attributes of carbon dioxide</td>
<td>Development of substitutes for Freon refrigerant</td>
<td>Development of energy-saving multi-purpose refrigerator</td>
</tr>
</tbody>
</table>

For instance, Table 2.2 indicates that the main characteristics of basic R&D are experimental and that researchers, in general, pursue abstract and theoretical findings. In addition, the basic R&D requires a long period of time and has a relatively wide range of applicability in that the findings can be applied to a wide variety of uses. Its time frame as well as applicability is mainly due to the very nature of basic R&D.

First of all, the researcher himself needs to get involved in the evaluation process. It is true that in the past, the person being evaluated would not be able to directly participate in the evaluation process. But more modern thoughts on human behaviors have increasingly emphasized the importance of involving the one being evaluated into the evaluation process itself. It is also necessary to take into account other stakeholders. Peer researchers, subordinate researchers, and project directors (e.g., bosses) are working on the same R&D project. Chief of the R&D center also needs to get involved. All these people are internal stakeholders, i.e., working inside the R&D organization. Finally, it is important to include outside perspectives, i.e., customers’ evaluation of the R&D project.

What are the criteria to be used? Bowon Kim & Heungshik Oh (2002)\textsuperscript{38} considered four
broad categories of performance measurement criteria, market-oriented, R&D project-specific, R&D researcher's technological, and the researcher's behavioral attributes (Table 2.3). Market-oriented criteria include such elements as public relationship building, social activity (including lectures to the public), and contribution to commercial profit. R&D project-specific criteria consist of innate technological complexity and duration of the R&D project. Technological capacity of the researcher measures the individual researcher's technical expertise, publication (papers and books), and number of patents. Finally, the behavioral capacity of the researcher pays attention to the researcher's mentoring of new researchers, leadership, ability to get things done, and efforts for teamwork building.

Table 2.3
Definition of variables

<table>
<thead>
<tr>
<th>Criteria group</th>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Market- oriented</td>
<td>Public relation</td>
<td>Ability of the R&amp;D personnel to communicate and negotiate with external stake-holders such as community, customers, outside institutes, etc.</td>
</tr>
<tr>
<td></td>
<td>Social activity</td>
<td>The R&amp;D personnel's level of social activities such as lectures to the public, community services, public relationship building, etc.</td>
</tr>
<tr>
<td></td>
<td>Commercial profit</td>
<td>The R&amp;D personnel's ability to 'directly' generate profit for/through the R&amp;D project and the R&amp;D center from outside (market / customers)</td>
</tr>
<tr>
<td>R&amp;D Project- specific</td>
<td>Tech. complexity</td>
<td>Technological complexity / requirement (e.g., the level of technical experience, technical knowledge, expertise, etc.) of the R&amp;D project, which the R&amp;D personnel is working on</td>
</tr>
<tr>
<td></td>
<td>R&amp;D project duration</td>
<td>Overall duration of R&amp;D project—how long it takes to complete the R&amp;D project</td>
</tr>
<tr>
<td>R&amp;D Researcher's technological attributes</td>
<td>Technical expertise</td>
<td>The R&amp;D personnel's level of technical expertise and / or sophistication including technical experience, technical knowledge, technological / technical education, etc.</td>
</tr>
<tr>
<td></td>
<td>Technical publication</td>
<td>Quantitative count of the R&amp;D personnel's published papers and books</td>
</tr>
<tr>
<td></td>
<td>Number of patents</td>
<td>Number of patents that can be directly linked with the R&amp;D personnel's contribution</td>
</tr>
<tr>
<td>Researcher's behavioral attributes</td>
<td>Mentoring of junior researchers</td>
<td>The R&amp;D personnel's ability to help / guide / mentor / train junior / young / new R&amp;D personnel</td>
</tr>
</tbody>
</table>
Successful R&D activities are vital for national as well as for an individual firm’s sustainable competitive advantage. Therefore, it is very important for R&D managers to be able to design an ideal R&D performance measurement system. Kim and Oh\(^{38}\) had developed such a system from the perspective of R&D researchers themselves rather than that of their bosses, e.g., directors of R&D project and chiefs of R&D centers. There are two critical dimensions constituting an R&D performance evaluation system: evaluators and criteria. That is, in order to design an ideal performance measurement system, an R&D manager should ask two questions, “Who should evaluate?” and “What criteria should be used?”

Kim and Oh (2002)\(^{38}\) conducted a large-scale survey to over 1,000 R&D researchers in Korea, concluded that a high correlation existed between R&D personnel’s job satisfaction and their satisfaction with the performance measurement system. The researchers wanted an ideal performance measurement system, which takes most of the inputs from the researchers themselves, their peer researchers, subordinates, and outside customers, rather than from R&D project directors and R&D center chiefs. An ideal measurement system should use more of the behavioral criteria such as the researcher’s ability to mobilize the organizational resources to get things done, teamwork building,
and leadership for the R&D organization. They postulated that R&D managers should consider not only the differences due to various R&D types, but also unique organizational characteristics even in the same R&D sector.

2.2. Section B

2.2.1. Issues in Performance Appraisal:

Organisations vary on a wide array of factors (including, for example, size, product market, technology, culture, competitive environment, strategy, and union representation) that are likely to affect performance appraisal practices. Therefore, describing the model setting in which appraisals take place is exceedingly difficult and should be approached cautiously.

The performance appraisal systems in place were designed primarily by personnel specialists with only limited input from the managers who used the system and virtually no input from employees affected by them. In the light of the growing interest in employee and customer participation and involvement (Miller and Monge, 198639; Schweiger and Leana, 198640; Wagner and Gooding, 198741) it was anticipated that more recently-implemented systems might rely on more input from line managers, employees and customers. However, recently implemented systems were no more likely to have them involved in their design than were older systems.

Management-by-objectives (MBO) is the preferred format for assessing executives, managers, and professional employees. Other formats, such as trait-based rating scales are far less common among these employee groups. However, "mixed" formats are common, and rating scales or ranking procedures are often used to
supplement MBO-based approaches. In contrast, MBO is used to a far lesser extent among non-exempt employees (as defined by the Fair Labor Standards Act). For these employees, trait-based rating scales are the norm. A sizeable minority of organizations report using behaviour-based formats. However, pure behaviorally anchored rating scales (BARS), forced-choice scales, or mixed standard scales are very uncommon. Again, the use of “mixed” formats is common, reflecting the multiple purposes appraisals serve in many organizations. Two employee groups are far less likely than others to receive formal evaluations. Executives are less likely to be formally evaluated − perhaps reflecting the difficulty of assessing performance at this level, or the reluctance of executives to submit themselves to the process. Additionally, many hourly employees do not receive formal performance appraisals − reflecting, in part, the organized labor’s distrust of the appraisal process and negotiated agreements, limiting the use of formal appraisals.

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

Quantitative indices are used to supply some performance information in most organizations. Profits, sales and costs are frequently cited as important measures for executives and managers. The acquisition and use of job-specific knowledge is
considered important for professional positions. Finally, attendance, quality and quantity of output are important measures for nonexempt and hourly employees. It is common to spend about seven hours per year assessing the performance of each employee at higher organizational levels and about three hours per year for each employee at lower levels. However, there is considerable variability on this issue. Many organizations report spending less than one hour per appraisal, while a few organizations spend considerable amounts of time. For example, one reported spending between 20 and 40 hours per employee per year on the appraisal process.

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

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

It seems uncommon for managers to be evaluated on how they manage the appraisal process. Basic motivation models suggest that people will tend to behave in ways that maximize their expected payoffs or in ways for which they are reinforced (e.g. Vroom, 1964)\textsuperscript{42}. In spite of this, only about one-quarter of the organizations surveyed attempted to hold raters accountable for how they managed the appraisal process. Among those that did, the most common method for doing so was to include it as a dimension on the rater's own appraisal form.

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

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

Performance appraisal systems typically have five levels to differentiate employee performance. However, even though most organizations report systems with five levels, generally only three levels are used. Both the desired and the actual distributions tend to be top-heavy, with the top “buckets” relatively full and the bottom buckets relatively empty. Also, the actual distribution of performance is generally higher than the distribution desired. That is, even though most organizations expect large percentages of employees to receive the top performance ratings and few, if any, to receive the lowest ratings, their expectations tend to be conservative. It is common for 60 to 70 percent of an organization’s workforce to be rated in the top two performance levels. This could reflect either actual outstanding performance on the part of the organization’s workforce, or it could be indicative of leniency bias. Since the phenomenon is surprisingly constant across organisations, and it is unlikely that all organizations have predominately outstanding employees, the distributions probably reflect the latter. As anecdotal evidence and researchers’ concern about leniency have suggested, it appears that the norm in U.S. industry is to rate employees at the top end of the scale. Skewed performance distributions not only exist, but are common.

Research has indicated that expectations of future performance influence the rating process. Since these expectations are formed on the basis of prior knowledge or beliefs, raters will virtually always have some prior performance expectations. It also seems that the type of appraisal process used would affect the degree to which prior expectations became problematic. For example, in MBO-based systems, the mechanism
through which a manager and subordinate arrives at mutually agreed upon goals requires that each individual form expectations regarding the level of performance that is achievable. Furthermore, since performance is measured against established goals, prior knowledge of job performance can also be expected. Therefore, prior expectations and prior knowledge not only exist, they are salient key features in the appraisal processes used in many organizations. Therefore, performance appraisal practice stands to benefit substantially from this line of research. Additionally, future research should explicitly consider the impact of prior knowledge and expectations under varying conditions of salience created by different appraisal systems. The conditions under which actual appraisals occur also suggest that continued research on memory characteristics should prove to be valuable. This becomes apparent with the realisation that managers report spending only a few hours per year assessing the performance of each employee. This time includes keeping records, completing forms, preparing for the appraisal interview and delivering feedback. This is a relatively small amount of time spread over a long period, particularly when the performance of several employees must be recalled and reported. Clearly demarcating the limits of memory and recall should lead to processes such as more frequent appraisals and systematic documentation, which will reduce reliance on memory.

Several other issues emerge from the cognitive processing literature. First, there has been a heavy reliance on student subjects and laboratory settings. While there is some evidence that laboratory settings may provide results that are as valid and generalised as those obtained in field settings (e.g. Locke, 1986)^43, there is also convincing meta-analytical evidence that in the performance appraisal arena, effect sizes
in paper-people studies are significantly larger than in studies involving observation of behavior (Murphy, Herr, Lockhart and Maguire, 1986). Laboratory studies are often necessary in order to isolate particular effects. However, sterile environments that dilute the richness and complexity of the environment potentially change the phenomenon of interest. The potential effects of situational and contextual variables must be considered.

The task of rating the performance of someone with whom an on-going relationship exists is both conceptually and operationally different from the rating task presented in laboratory settings. Therefore, while continued research on isolated cognitive processes is useful, research agenda should be expanded to include attempts to understand how these cognitive processes are affected by the political, social and affective nature of most rating environments (Dipboye, 1985; Ferris and Judge, 1991).

Research on personal characteristics of raters and ratees is relatively balanced between laboratory and field settings, and between student and employed samples. Given the labor market trends, expected to continue over the coming decade, research age, gender, race, and ethnicity effects seems particularly timely and important. Additionally, the role of women and minorities is expected to increase, particularly among managerial ranks. Finally, the internationalization of the workforce introduces cultural differences regarding job design, performance expectations, and the role of performance feedback. Current research may be culture bound since it assumes a decidedly Western approach to these issues. It may be that employees raised in a traditional Japanese environment, for example, may expect a much less directive approach to appraisal and be offended by the confrontational nature of direct feedback. If
individual difference characteristics such as these change the way ratings are assigned or interpreted, these trends represent significant issues to be addressed.

Researchers and managers appear to have different conceptualizations of accuracy. What does accuracy in performance appraisal imply? Many researchers would suggest that accurate appraisal are those that are both reliable and valid, and conceptually near the true score level of performance. However, managers tend to define accurate appraisals as those that are accepted by employees and allow the identification of relative contribution to organizational effectiveness within the context of the organization and the constraints imposed by the regulatory environment in which it operates. This definition is quite different from one involving deviations from true scores. One potentially fruitfull approach to resolving this disparity is Lord's (1985)\textsuperscript{47} distinction between classification and behavioural accuracy. Padgett and Ilgen (1989)\textsuperscript{48} concluded that behavioural accuracy was both closer to the conceptual criterion and more descriptive of what accuracy seems to mean in organizational settings. However, since conveying behavioural information is seldom the sole purpose of performance appraisal, Murphy (1991)\textsuperscript{49} points out that both types of accuracy have meaning and may be differentially useful depending on the purpose for which the ratings are made (Murphy, Philbin, and Adams, 1989)\textsuperscript{50}. For example, behavioural accuracy may be more important when providing feedback while classification accuracy may be preferable for salary administration decisions (Murphy, 1991)\textsuperscript{49}. Therefore, it appears that additional research on the antecedents, definition, and consequences of accuracy are in order.
It seems that an important point is frequently overlooked in research on rating sources. Rather than focusing on who should rate the performance of others and examining the psychometric properties of various rating sources, perhaps research should examine the propriety of various rating sources under various conditions. That is, when should ratings from alternative sources be used and how should they be integrated with ratings from the immediate manager? It appears that this will become an even more important issue as the nature of work continues to change. For example, the increasing utilization of self-managed work teams may legitimize peer appraisals and reduce reliance on supervisory ratings. Customer-centered appraisals may also offer potential for understanding performance in specific situations. However, appraisals by customers may be problematic since customers may have limited job knowledge and these ratings are prone to selection bias (since only extreme information is likely to be conveyed).

Performance appraisal system may be considered to be a series of decisions which are affected by environmental, organizational and dispositional factors. Research is needed to examine the variety of situational variables that affect appraisal design and administrative choices. For example, what are the determinants of managerial choices in performance appraisal design and administration? What factors lead managers to choose to decentralize policy and administrative responsibilities, to select a forced distribution approach, to use more objective performance indices, to place more weight on second level managers as sources of information, to change the number of levels in the appraisal scale or even to determine the desired performance rating distributions? These types of questions are illustrative of the myriad decisions that go into implementing any performance appraisal system. While MBO is the most common approach for assessing
managers, graphic rating scales are more common among nonexempt employees, and many executives are not subjected to any formal appraisal process. These differences are potentially problematic since the appraisal process appears to become less standardized and systematic at higher organizational levels. It seems that these differences represent potential research opportunities. For example, research could examine whether this practice affects perceptions of procedural and distributive justice both between and within employee groups. An obvious characteristic to consider is the degree to which performance ratings drive pay decisions. Specifically, stronger pay for performance contingencies might be more important under conditions of intense competition than in regulated or cost-plus environments. Also, perhaps administrative use of performance data should be emphasized in a meritocracy, but developmental uses should be emphasized in public sector/civil servant environments. Finally, different types of organizations certainly reward different types of behavior. Olean and Ryes (1984)^51 speculated, for example, that prospectors and defenders would reward very different types of behaviour. Might it also be that particular types of organizations are more willing to set "stretch goals", forgive performance deficiencies, or establish voice mechanisms? Answers to these types of questions might begin to create a taxonomy of "fit" between performance appraisal system characteristics and organizational environments. Such a taxonomy would certainly allow research to more clearly delineate the most effective appraisal practices for particular organizational settings.

Napier and Latham (1986)^52 found that managers perceived no consequences, positive or negative, of conducting performance appraisals. Conversely, Longenecker, Sims and Gioia (1987)^53 reported that because of actual and perceived negative
consequences of accurate appraisal, some managers knowingly make ratings that are inaccurate. However, the rating environment likely affects both rating practices and participant reactions. For example, research has shown that contextual variables (such as participation and rating frequency) affect ratee satisfaction (Dobbins, Cardy, and Platz-Vieno, 1990; Giles and Mossholder, 1990). With these exceptions, recent research has not examined the effects of different appraisal system designs or processes on employee attitudes and behaviours. Some characteristics would appear to influence participant attitudes. For example, research also might address whether the use of three, four or five (or more) rating levels affect employee motivation or work attitudes. Too many levels would supposedly lose the ability to meaningfully differentiate performance. On the other hand, too few levels are likely to create perceptions of inequity. However, what is too many or too few? Consider the proto-typical three-bucket system in which the top and bottom levels are reserved for truly outstanding and truly problematic cases. The middle level is designed to capture 75 to 80 percent of the workforce. If performance is normally distributed, an employee performing one standard deviation below the average would likely receive the same rating as an employee performing one standard deviation above the average. The better performers would likely consider this to be unfair. What are the perceptions of employees at different positions in the distribution? Are the better performers in the middle bucket more likely to withdraw than are the poorer ones? Are they more likely to withdraw than are similar performers in say five-level systems in which their performance would be differentially acknowledged?
Deming (1986)\textsuperscript{56} argues strongly that performance appraisal has serious negative consequences and urges organizations to cease all individual performance appraisals and to evaluate unit or plant level performance instead (Scholtes, 1987)\textsuperscript{57}. This approach, reportedly in wide use in high-involvement, high-commitment facilities, focuses on assisting those whose performance is "out of the system". Deming's notions have received some attention in practitioner performance appraisal literature but no attention from researchers. The possibility of no individual feedback seems difficult to attain. Even without formal individual appraisals, informal appraisals by team leaders and peers seem inevitable and perhaps potentially less systematic and more vulnerable to biases. Accordingly, research might address the effects of informal appraisal processes on employee perceptions, attitudes, and behavior. The effects of skewed performance distributions on pay allocation and employee attitudes also need to be examined. Highly skewed ratings affect the distribution of merit pay increases. In fact, some argue that the size of the merit fund pool affects the ratings skew (Milkovich and Newman, 1991)\textsuperscript{58}. Smaller funds may force managers to give higher ratings which result in smaller average increases for high rated performers. These smaller pay increases coupled with high ratings deliver mixed signals to employers and may affect their attitudes toward merit pay. In this regard, tracking changes in appraisal system decisions over time may be useful. For example, some firms report shifting to forced distributions in an effort to correct their rating skew. Such changes may lower employees' satisfaction with appraisals on the one hand, but on the other improve employees' satisfaction with their pay increases. More generally, it seems desirable to understand why firms make changes in their systems and to examine the effects of these changes. There is need to continue
developing a more comprehensive theory of the rating process. Since Wherry’s work in
the 1950s’ (see the Appendix to Landy and Farr, 1983)^9, the collection of studies on
information processing is the most serious, concentrated attempt to date to better
understands the rating process. In that framework, continued research along those lines is
useful indeed. However, attention must be paid to the potential effects of situational or
contextual variables. Examining appraisal issues in sterile environments may isolate the
effects many researchers wish to investigate, but also limits the generalisability of the
results, and removes the issues from the attention and interests of human resource
decision makers. If research is to inform practice, interaction between researchers and
managers, and application of research results are important.

On the other hand, organisations continue to do things that undermine the
effectiveness of the appraisal process. Little time is spent on the appraisal process; raters
are not systematically trained and are not held accountable. The employee’s role in the
performance process is overlooked as are many potentially valuable sources of
performance information (self, peers, subordinates). While research has done much to
suggest improvements regarding many of the practices noted above, it may ultimately be
the changing nature of work that leads managers to implement practices that research has
legitimized. For example, the trend towards self-managed work teams is diminishing the
traditional supervisor-subordinate relationship. While on the one hand this may
ultimately lead to greater acceptance of peer appraisals, on the other hand it is likely to
force research into new directions as well. The cognitive processes involved in peer
appraisals are likely to differ from those in supervisory ratings due to the differences in
power and social interactions resulting form peer relationships.
Multi-source feedback or 360°-feedback, the process in which subordinates, peers, supervisors and/or customers provide anonymous feedback to recipients, has grown in popularity over the past decade (Waldman and Atwater, 1998). In 1996, Antonioni reported that an estimated 25 percent of companies were using some type of upward or 360-degree feedback process. In fact, several studies have reported improvements in overall performance following 360 or upward feedback interventions (e.g. Atwater, Roush and Fischthal, 1995; Reilly, Smither and Vasilopoulos, 1996). Positive behaviour change and increased self-awareness are not the only outcomes that may result from 360-degree feedback. Recent research has suggested that individuals may react with anger and discouragement when 360-degree feedback is negative or not as positive as expected (Brett and Atwater, 2001). Kulger and DeNisi’s seminal piece on the impact of performance appraisal feedback on performance also raised the issue that performance improvement is not the only outcome that can be expected from performance feedback. In fact, they concluded that in one third of the cases feedback actually resulted in decreased performance. Atwater, Waldman, Atwater and Cartier (2000) found that improvement following an upward feedback intervention only resulted for 50% of the supervisors who received it. It seems unlikely that organizations would want to invest thousands of dollars in a process that helps good managers get better but discourages and angers managers receiving negative feedback. An understanding of why and how recipients react to negative feedback may enhance the likelihood that the 360-feedback process will assist those most in need of the developmental feedback. Salient features of the feedback itself or attitudes of the
feedback recipients may impact how individuals react to their feedback and towards their job or organization following feedback. This study investigates the influence of feedback recipient’s attitudes and feedback format on the feedback recipient’s emotions and reactions.

Recent research on 360-degree feedback suggests that some feedback recipients may not benefit from the process while others demonstrate notable improvements (Atwater, Waldman and Brett, 2002). Part of the reason some individuals do not respond favourably to the feedback may be due to the format in which it is provided. In a summary of research on employee reactions to performance standards, Bobko and Colella (1994) reported that research on performance feedback indicates that specific, descriptive feedback (e.g., how the individual scored), results in perceptions of source credibility and system fairness, more accurate evaluations of expected future success, and increased performance compared to evaluative outcome feedback (e.g., how well the individual scored relative to others). Ilgen and Davis (2000) suggested that the framing of negative feedback may influence how recipients cognitively and affectively respond to feedback. They asserted that the challenge of delivering effective negative feedback is achieving a balance between making recipients accept responsibility for low performance and not lowering the individual’s self-respect.

DeNisi and Kluger’s (2000) research on performance appraisal feedback suggested that feedback that highlights self-other comparisons focuses the recipient’s attention on him or herself rather than on the behaviours or task performance evaluated. This type of comparative feedback, where individuals see themselves as “below average,” causes employees to become alienated and demoralized. This is because most
employees consider their work to be above average (Meyer, 1975)\textsuperscript{71}. DeNisi and Kluger (2000)\textsuperscript{70} discuss feedback intervention theory which suggests that the effectiveness of any feedback depends on the level at which the feedback intervention focuses one's attention. When attention is focused at the task level (e.g., tasks on which the person needs to improve) individuals focus on shrinking the gap between their actual performance and their performance goals. Alternatively, when feedback focuses attention on the self (e.g., when comparisons are made between oneself and others), feedback interventions often produce strong affective reactions that can interfere with task performance. DeNisi and Kluger (2000)\textsuperscript{70} reported that feedback that provided comparative information was associated with performance decline while feedback that provided information about past performance resulted in performance improvement. A typical 360-degree feedback report includes numerical self-ratings, averaged other ratings, and normative data about one's performance ratings compared to others in the organization. Comparing one's own ratings with those received by others may result in increased self-awareness for some individuals, but for others these comparisons may be detrimental because their ego or self-esteem is threatened. Attention to self produces affective reactions that may interfere with task performance and deplete the cognitive resources needed for task performance (Kanfer and Ackerman, 1989)\textsuperscript{72}. DeShon and Alexander (1996)\textsuperscript{73} found that when the discrepancy between a performance standard (similar to a self-rating a 360-degree process) and the actual level of performance (similar to others ratings in a 360-degree process) was large, the recipient was less likely to persist with the task. Kluger and De Nisi (1998)\textsuperscript{74} asserted that feedback interventions will be more effective if they are focused on goals at the task level and less focused on the self.
This literature review indicates that new concepts like 360° Appraisals (multi-rater appraisals) are taking root in assessing the performance of R&D scientists as well as performance appraisals in other organisations. The CSIR performance appraisal system has so far been only hierarchical in nature, i.e. self-appraisal, appraisal by a reporting officer and review by reviewing officer. The systems adopted in Public Sector also indicate the same trend except that mid-term review is conducted in some organisations and performance appraisal interview is conducted in a few organisations. The purpose of the current research is to study the performance appraisal patterns of public funded organisations in general and to study the performance appraisal system of CSIR in particular. The concept of 360° appraisal is absent in the Indian parlance. The researcher proposes to establish a comprehensive performance appraisal system for CSIR. The concept of MBO would also be employed inasmuch as the goals to be achieved are discussed with the assessee beforehand. In essence, the latest trends in performance appraisal research would be used for devising a performance appraisal system for CSIR which includes MBO type, self-appraisal, mid-term review, peer review, review by subordinates, personal interview to discuss strength and weaknesses of an employee in an informal way in order to elicit an effective feedback. To eliminate the bias and also to maintain consistency in the organisation, gradings will also be normalised by an empowered group and the final score would be communicated for the satisfaction of the employee to ensure transparency in the performance appraisal process in its entirety.
REFERENCES


