CHAPTER-II

AN OVERVIEW OF PERFORMANCE APPRAISAL AND LITERATURE REVIEW

2.1 Introduction

The review of literature on employee performance appraisal is divided into five parts covering five aspects: (a) historical overview of performance appraisal, (b) overview of past research and literature, (c) approaches to evaluating performance appraisal, (d) performance appraisal for scientific personnel in research and development organizations and (e) perspective of perception of fairness of and satisfaction with performance appraisal.

2.2 Historical overview of Performance Appraisal

The genesis of performance measurement, its appraisal, its monitoring and finally its management perhaps can be traced to Plato’s theory. Plato argued that men are not born self-sufficient on all alike; hence an organized society in which they are interdependent and specialized according to innate aptitude is both natural and advantageous to all the individuals. The selection of one product over another produced based on quality and craftsmanship would perhaps have been starting point of performance evaluation of one individual by that of another. Kautilya, the prime minister of great emperor of ancient India Chandragupta Maurya, in his treatise “Arthasastra”, emphasizes that the best way for a king to judge the ability of his ministers is to look at their performance. Similarly, Machiavelli exhorts in his classic Text “The Prince” that the ruler should be aware of three types of men. The first type are those who understand themselves, another are those who understand what is explained and third are those who neither understand by themselves nor by any
explanation. Machiavelli rates, as is done in modern rating system, these three types of men as excellent, commendable and the third type as altogether unprofitable.

Wright (2002) claims that performance appraisal can be traced back to the Han Dynasty in 206 BC – 220 AD. Merit examinations were given for selection and promotion decisions as early as the Han Dynasty. A procedure to formally rate members of the Jesuit society was established by Ignatius Loyola. Fairness of raters was questioned since the third century by the Chinese Philosopher Sin Yu who alleged that the Imperial rater of nine grades seldom rates men according to their merits, but always according to his likes and dislikes. In 1648, Dublin evening post of Ireland is said to have evaluated legislators by using a rating scale based upon personal qualities. In his book first published in 1957, McGregor questions the ethical perspectives of man sitting in judgment of another for purposes of controlling rewards.

Performance appraisals in industry were perhaps initiated by Robert Owen in the early 19th century. Owen mentioned performance at his Cotton mills in Scotland through the use “Silent Monitors”. These monitors were essentially blocks of woods painted with different colours on each visible side. The colour of the visible side of the cube was associated with a rating to indicate performance. Here, white indicated ‘excellent’ yellow indicated ‘good’, blue was used to indicate ‘indifferent’ while black indicated ‘bad’. Thus at the day’s end each employee could know not only his own performance, but also that of his colleagues. Performance appraisals were formally implemented in the United States Military in 1813. Army General Cass reported to the war Department on individual ratings of using a global rating with such descriptions as ‘a good natured man’ or ‘a Knave despised by all’ and so on. In the late 1800s, the federal civil service of the United States began giving merit rating. These ratings were primary information about competence, faithfulness and attention of clerks. Due to their large size, hierarchical structure, geographic dispersal and the
necessity to promote the top performers to higher organizational levels, in the late
nineteenth and early twentieth century, performance appraisals were used primarily
by military and government organizations.

2.3 Overview of Past Research and Literature

Popularity of performance appraisal immensely increased among industrial
employees only after the man to man rating system was successfully used during the
World War I (1914-1918). This led to much more academic interest in measurement of
performance. Donald Peterson developed the graphic rating scale in 1922 and from
there on number of innovations was introduced in performance appraisal scale
construction. It was only when industrial Psychologists at Carnegie Mellon University
used trait psychology to develop a man to man rating system for the selection of
salesman that performance appraisal formally began to be used in the U.S industry.
By the World war-II, the forced choice method and critical ranking methods of
appraisal were in place. However, all these tools of appraisal focused on past actions
of people. Supervisors notoriously used such records to justify workforce reduction
and other unrelated purposes. By the early fifties performance appraisal system was
slowly moving out of strictly administrative decisions making paradigm to serve the
bigger purpose of employee development and feedback. Thus, Peter Drucker in his
1954 classic book titled ‘The practice of management’ proposed the concept of
Management by objectives (MBO). In 1954, McGregor, applying MBO principles to
performance management, advocated that instead of appraising employees only on
the parameters of trait, managers and employees should mutually fix short goals on
the basis of which managers can objectively appraise the performance of employees.
Since the goals are specific, measurable, time bound and joined to an action plan, the
entire appraisal process was far more objective, worthwhile and credible. McGregor²
(1957) rightly points out that the rating systems forced the manager into role of judge
thus creating inconsistency with his primary roles of leader and coach which were so much necessary to achieve both the employees and organization’s goals. Smith and Kendall (1963) came up with a rating scale called the Behaviourally Anchored Rating Scale (BARS). This tool of appraising performance was tool with solid psychometric properties in terms of reliability, validity and discriminability and was much closer to real life on the job utility. BARS replaced numerical or adjective anchors which were used in the graphic or trait rating scales, with behavioural examples of actual work behaviour. This development made it possible for supervisors to rate employees on observable behavioural dimensions. Raters cannot provide adequate feedback to develop equitable reward systems unless good performance can be distinguished from bad performance. Subjective criteria are likely to produce errors of human judgment (halo effect, errors of central tendency, leniency etc.) while objective measures can provide an incomplete picture of employee performance because certain unique features of particular jobs and people will be distorted. But subsequently it was found that as rater error went down due to increase in robustness of psychometric properties of scales, there was a simultaneous decrease in rater accuracy i.e. between actual behaviour and subsequent rating of the actual behaviour. Landy and Farr (1980) in their highly critical and influential review of performance appraisal research called for a moratorium on rating format research and attempted to focus research to other area such as understanding the rater and the process in an organizational context. They characterized the abundance of studies in the following categories: ‘rules’ or characteristics of the rater and the ratee; vehicle or rating format and form; the context of the rating including its use and the rating process which dealt with data analysis and rater training. Bernardin and Buckley (1981) proposed written diaries as a tool for performance rating. They suggested that supervisors needed to
maintain written diaries, documenting critical incidents of performance of subordinates that could be used as a basis for performance appraisal ratings at the end of the year.

Prior to early 1980’s, the majority of theoretical and empirical studies focused on improving the psychometric characteristics of the rating instrument in an effort to reduce the subjectivity inherent in performance ratings (Feldman\(^6\) 1981). Due to emphasis on psychometric aspects, the development of a better rating scale format which was valid and reliable received a great deal of attention. Research focusing on rating scale format and development peaked in the 1960’s and 1970’s with the development of several new formats including the Behavioural Observation Scale (BOS), the Behaviourally Anchored Rating Scale (BARS) and mixed standard scale. Other popular and related research topics included training raters to reduce rating errors and improve observational skills and developing performance appraisal practices. According to Arvey and Murphy’s\(^7\) (1998) review of the research, there were literally hundreds of studies between 1950 and 1980 on the different types of rating scales, rating versus ranking and ways of achieving ratings that were objective measures of performance.

The influence of Landy and Farr (1980) and Feldman (1981) resulted in a change of focus away from the rating scale format and rater training to understanding the rater as a decision maker who processes information and social cues. Research in the 1980’s and early 1990’s focused on raters and accuracy of ratings and judgments and the application of knowledge about judgment process in the development of performance appraisal systems. Another contribution is related to the application/use of appraisal ratings. The context in which the ratings were obtained and the beliefs about the use of such rating were found to influence the results. Researchers further argued that the ratings should only be used for those purposes understood by the
raters at the time of the rating. Relative measures of the attitudinal kind may ultimately prove to be better measure and predictors of rating validity than such traditional psychometric variables as leniency, halo and discriminability. A performance appraisal system can be psychometrically sound in design and construction but still wholly ineffective in practice due to resistance or lack of acceptance on the part of users. Thus, the effectiveness of a system is particularly contingent on the attitudes of the system users, both raters and ratees. There was a shift from improving psychometric properties to understanding how the rater processes information about the employees and how this mental processing influences the accuracy of performance appraisal.

Research of the performance appraisal process during the 1980s contributed a number of key ideas to the literature including a heightened awareness of the importance of observation in the appraisal process and how knowledge obtained by raters is utilized. The research of the 1980s also helped to clarify some assumptions about performance appraisal such as the belief that rating errors as commonly defined were evidence for rating errors when in fact the research indicated that there may not be resulting inaccuracies.

Performance management systems started to be used for four distinct purposes - between person appraisals, within person appraisal, system maintenance and documentation (Cleveland et al. 1989). Between persons appraisal is primarily the traditional use of performance appraisal for administrative uses of compensation, recognition, promotion, retention etc. Within person uses of performance management systems is primarily for providing feedback to an individual as regards his strengths and weakness, training needs and job fit. The system maintenance uses of performance management systems is focused at work force planning, determining organizational training needs, evaluating goal achievement, identifying
organizational development needs, assisting in goal identification, evaluating the personnel systems and reinforcing the authority structure. The documentation uses of performance management system is to ensure that safer work is complete so as to meet legal requirements as regards individual employees’ career path progression and also to conduct validation research in performance appraisal tools.

Satisfaction with appraisal is positively correlated with employee participation in the appraisal process, development of action plan to remediate performance weaknesses and the existence of a formal rater training programme as well as closeness of supervision and quality of the leader-subordinate exchange (Dobbins et al. 1990), it is negatively correlated with role ambiguity and role conflict.

Bretz et al. (1992) generalized that research in the late 1980’s and early 1990’s was heavily weighted toward cognitive process issues. Ratee and rater personal characteristics and rating errors and accuracy were also researched. The course of appraisal, appraisal feedback mechanisms, rater training and performance appraisal format were found to be studied frequently. According to them, performance appraisal research has three general streams; first research has focused on developing and honing instruments to more accurately and objectively measure performance. Second research has focused on supervisor and employee characteristics sources of potential and actual bias in performance appraisal ratings. Third research has focused on uses and types of performance appraisal systems within the organizations.

According to Ilgen et al. (1993), the ‘rater process perspective’ include three critical sets of operations: i) acquisition of information about those to be evaluated (ii) organization and storage of information in memory and (iii) retrieval and integration of information in a fashion that leads to the recording of an evaluation of the person
being appraised. Researchers borrowed heavily from basic psychological research in cognitive psychology and social cognition to address the three process described above and to develop theories of performance appraisal process.

Techniques used to help combat errors of human judgment are based on comparison of worker performance or absolute judgment using fixed standards as reference, or description of work behaviour. Rating procedures, rating scales, weighted checklists, forced choice checklists and critical incident techniques are commonly used approaches, none of which guarantee bias-free evaluation. Bias may be minimized through the behavioural description methods. Awareness should be raised among raters so that they can keep abreast with potential sources of judgment error and encouraged to participate in the development and evaluation of the process to remove biases and distortions through subjective evaluation.

Other research included ratee and rater characteristics such as gender, likeability and race. Rater attributes including cognitive style and knowledge of the job to be rated were explained. Rating scale accuracy continued to be studies and the characteristics of the setting in which appraisal occur such as purpose of appraisal, rater training and other factors were investigated.

Research on performance rating accuracy and the development of accuracy criteria was common in the 1980s. Research focused on common Psychometric biases, called rating errors such as leniency, central tendency and halo, with the assumption that these implied a lack of accuracy. It was assumed that decreasing the biases increased accuracy. Researches argued that this assumption was not necessarily true in that bias free ratings were not necessarily more accurate (Murphy and Balzer 1989). Research on accuracy shifted from rater errors to discrepancy between ratings and some standard of performance. Research examining the efficacy
of the different rating scales format generally indicated that ratings were not affected by changes in the rating scale format (Woehr and Miller\textsuperscript{13} 1997).

Recent research into performance appraisal has emphasized process and structural characteristics that influence the attitudes and affective reactions of system participants in addition to psychometric characteristics. Researchers need to consider the rating context before attempting to analyze or evaluate the effectiveness of ratings on rating systems. Research has included measures of employee attitudes towards performance appraisal and systems acceptance and rater and ratee satisfaction in the appraisal process.

Support for the efforts of adequate notice, fair hearing and Judgment based on evidence on the reactions of employees and managers, who as a result generally, perceive greater accuracy and fairness in the appraisal systems and greater satisfaction with appraisals (Taylor et al.\textsuperscript{14} 1995).

Wanguri\textsuperscript{15} (1995) reviewed cross-disciplinary research that was published between 1980 and 1990 as performance appraisals, evaluation, and feedback. Surveying a total of 113 empirical studies, she summarized research on rating instruments, raters and ratees accordingly as follows: (1) organizations use collaborative instruments to evaluate professional and managerial personnel and traditional instruments to evaluate hourly and non-exempt personnel, (2) organizations prefer to use MBO and BOS instead of trait scales, (3) the relative accuracy of BARS is inconclusive, (4) assessment centers should represent only one criterion for managerial promotion, (5) training in the proper use of rating instruments is essential in order to minimize error, (6) multiple rater performance appraisals are more effective than single rater appraisals, (7) the degree of agreement between self rating and others’ ratings is inconclusive, (8) inter-rater agreement is strong between
peer and supervisory ratings, (9) subordinate appraisals of managers are valid measures of supervisory performance, (10) rater characteristics (e.g. organisational tenure, age, gender, style of dress) interact with ratee characteristics to affect performance ratings: (11) managerial ratees’ perceptions of the formal performance appraisals are divided (12) perceived fairness of performance appraisal process is influenced by the superior-subordinate relationship and this perception influences views about compensation and (13) performance feedback is positively correlated with ratee job satisfaction.

Popularized by Kaplan and Norton\textsuperscript{16} (1996) as a means of overcoming the short-termism and narrow tendencies of management accountancy, the balanced score board presents an approach that promises to allow organizations to measure both hard tangible and soft intangible drivers of performance. It is based on the notion that measurement motivates, drawing on the theories such as goal setting and expectancy theory, and aims to move away from the problems created by short term reliance in financial measurement alone. The score board outlines four different perspective viz., financial measures, customer measures, internal business process measures and learning and growth measures under which to identify performance measures and trace organizational performance. Instead of focusing on the definition of measures, European Foundation for Quality management (EFQM) model operates in a different way. Rather than providing a frame work for an organization to define its own strategy, it provides an opportunity to audit current practice against a European standard. The thinking behind the EFQM is that by adhering to the standards (and there are questions to guide an organization to diagnose their current price), higher quality and hence business performance will result. Both the BSC and EFQM approaches influenced by accounting and operational management
perspective to performance management. While these approaches focus on measurable outcomes, performance measurement system solely based on such process can become overly concerned with targets and lose sight of performance.

Beyond 1980s, within person uses of performance system started receiving much more attention primarily to ensure individual and team performance improvement and effective career development process. Thus, the focus got heavily skewed towards robust feedback systems. In the 1980s, multi-source appraisal became popular in industry as an executive development tool. However, as a performance appraisal tool, the 360 degree appraisal gained popularity only in the 1990s primarily because of three key research findings of the Centre for creative leadership in Greensboro, North Carolina, USA (McCarthy and Garavan17 2001). These findings were as follows: a) feedback is an important element of a person’s professional and personal development; b) The most effective executives were learners, in other words effective executives encourage and welcome opportunities for learning and development and c) Most employees operate in ‘feedback – poor’ environments.

Many innovations were done on the basic concept. Today, the 360 degree appraisal tool is quite popular and is being used under multiple variations and names like multi-rater appraisal, multi-source appraisal, full circle appraisal and multi perspective ratings.

The literature indicates a substantial gap between research and practice as apparent in the 1990’s when many studies were conducted in the Laboratory and focused on discrete variables of cognitive processing in appraisal and evaluation. Other empirical research has focused on the performance appraisal process and the factors that influence the communication and behaviours exhibited by the raters and
ratees during evaluation. Most of these studies are in experimental settings and not in the field.

Although objective and subjective performance measures should not be used interchangeably, differences in rating scale formats seem to have little impact on the reliability or validity of performance ratings. The greater the criterion specificity (i.e. behaviours, results or outcomes, as opposed to traits), however, the greater the probability of a fair and accurate appraisals, thereby decreasing the chances of age, race or gender bias. In addition, distributional ratings, which reflect meaningful fluctuations of ratee performance within dimensions, contain substantially less measurement error than more typical global rating scales. Total Performance Effectiveness (TPE) can be used to express the effectiveness of performance in many jobs, particularly in situations where there is not much differentiation among employees in their average performance levels. Performance Distribution Assessment (PDA) is another method of performance appraisal which has revealed significant correlation between the PDAs derived performance measures and objective measures of job performance, differential correlations between ability and the multiple PDA-derived performance measures and equivalent levels of rating accuracy for the PDA and the evaluative measures of typical performance.

A Review on performance appraisal, evaluation and feedback during the decade 1990-2000 shows the following trend in performance appraisal research:

1. Research during the decade reflects more quantitative studies that investigated counter-rational dimensions in the performance appraisal process. Although many studies have approached this process as a highly rational and systematic activity, a number of studies during this decade addressed performance appraisal as a highly potential engagement, where performance is not necessarily the bottom line and where rating inflation occurs regularly.
2. Interaction effects between rater and ratee characteristics were examined. These studies focused on the interaction between rater and ratee, nationality, race and gender.

3. Research during this decade reflected preferences of raters and ratees regarding performance appraisal process. These studies focused on preferences of personnel professionals in the development of a performance appraisal system and on the preferences of employees regarding nontraditional performance appraisal.

4. Communication transaction in the performance appraisal process was focused in the research.

Mann et al.\(^{18}\) (2003) highlights that the focus in performance appraisal research has changed dramatically over the past decade, from the instrument used and the sources of appraisal to the on-going coaching, developing and motivating of employees. A subordinate’s performance increases significantly more in those years in which the supervisor holds appraisal meetings than in the years where such meetings do not take place. Using a rating from one self as a performance appraisal technique has been proven through a meta-analysis to have the least predictive validity and least agreement with other sources. However, those whose self appraisals were aligned with appraisals received from others tended to be high performers. Peer ratings have been found to be among the best predictors of performance likely because peers have more access to relevant job information. Peer feedback is not only consistent with self managing teams and team work; it has also been shown to increase interpersonal effectiveness, task motivation, group cohesion and satisfaction. Peers place greater emphasis on interpersonal performance and
motivation than on task performance. Perhaps the most exciting development in the sources of performance appraisal is 360 degree feedback with over 90% fortune 1000 firms now using this technique. There is a focus on ongoing coaching, developing and motivating employees. To bring about a positive change in behavior, feedback must focus on the behaviour rather than the person, be selective so as not to overwhelm the person, focus on the behaviour that is desired and the way to demonstrate it, and serve as the basis the setting specific high goals. There has been shift from the performance appraisal to the performance coach and a shift from the people and the evaluator to the people developer.

Cyclical year round performance management (i.e. Feedback, analyzing results, setting goals) has been found to effectively increase organization performance. Another tool that has gained prominence among some corporate houses is the key performance measures (KPMs). The primary function of these KPMS is to further enhancement of performance rather than serve ends of control mechanisms for managements. The focus is again on a robust feedback process. All employees know how they contribute to the company without having to ask the boss how they are doing.

For whatever reasons, it would appear that towards the very end of the 20th century, a new template for performance management approaches was emerging that took its lead from accountancy and business process, and quality movements. The main manifestation of a measurement based approach may be seen in the balanced score board (BSC) and the European Foundation for Quality Management (EFQM) model.

While theoretical research on performance appraisal continues to evolve, practical literature has focused primarily on improving performance appraisal process making the review interview more positive, improving the contribution of the employee
to the process, emphasizing the goal setting and applying procedural improvements. Professional Journals are filled with articles discussing performance appraisal practices in various organizations under varying conditions. Case studies and “how to” articles are common.

Compton\(^{19}\) (2005) finds from the longest Australian study on performance management that substantial changes have been made with respect to the purpose and nature of performance management systems and more customized and integrated systems are required to reflect competitive imperatives and to enable closer links between individual, group and organizational objectives and outcomes. Grubb\(^{20}\) (2007) finds that in practice performance appraisal does not produce results intended and expected and it is disliked by many employees and managers alike. Performance appraisal as commonly used, can reduce employee productivity, satisfaction and engagement.

Much of the past research has focused on the individual as related to the act of performance appraisal as opposed to performance appraisal as a system within the larger context of an organization. More recent research has investigated performance appraisal in a more comprehensive and organizational context.

2.4 Approaches to Evaluating Performance Appraisal

Evaluation of proficiency of performance appraisal system is generally suggested as an important step forward for efficient implementation of the system in any organizational set up. There is dearth of comprehensive research on the evaluation of efficacy of the prevalent performance appraisal system in the context of organizations. This is attributed to difficulty in selecting right criteria of evaluation and intricacy of the appraisal system.

Marshall Shaskin\(^{21}\) (1981), in his book “Assessing performance appraisal”, writes “Real participation in mutual goal setting helps develop and support on
organizational climate of increased satisfaction and productivity and an effective performance appraisal system should emphasize superior-subordinate goal setting." These mutual goals should specify what performance criteria and results are needed to be accomplished within a specific time span. They should be based on the key areas that are present in the performance appraisal.

Mohrman and Lawler\textsuperscript{22} (1983) suggest that researchers should concentrate on how performance appraisal systems are perceived by organizational members to improve performance appraisal accuracy. Organizations examine the uses of performance appraisal information to determine if the uses and functions are conducive to accurate performance appraisal.

Researchers make the case for evaluating performance appraisal systems based on a ratio of cost to potential outcomes. Costs include that of system development, system introduction and system maintenance. Major outcomes include meeting the intended goals of the system and achieving organizational acceptance. One of the most difficult aspects of assessing (or creating) a performance appraisal system is to identify a finite set of appropriate goals for the system. They also state the need for informed participants at all levels throughout the organization to know why and how they are to do performance appraisal. Researchers suggest that employee reactions to performance appraisal systems are usually better indicators of the overall viability of a system than the more narrow psychometric indices such as leniency and halo.

In their description of a complete appraisal system, Mohrman et al.\textsuperscript{23} (1989) include the following components: 1) two performance appraisal cycles that deal with immediate feedback and long-term career issues; 2) a decision about who defines performance; 3) how performance will be measured; 4) who will measure performance; and 5) what method will be used to gather performance information; and
6) effective feedback that is timely and correctly delivered by the appropriate person.

Appraisals should be timed so that they coincide with job characteristics and avoid peak periods of activity. The performance appraisal system needs support from top management to generate the requisite commitment from middle managers. An appeal process for employees to question or challenge their evaluation results lends credibility to the appraisal system. The authors state that the following key items are part of an appraisal system: appraisal tools and methods; the degree of fit between other features of the organization and the appraisal system; the system design; the proper introduction of the system; and training of individual system users. The authors further state the performance appraisal process must be designed to match the organization's goals and the type of work that is performed. They believe that one of the most critical factors in effective performance appraisal is clearly defining the purpose of the appraisal system. Possibilities include monetary compensation, career planning, documentation of staffing changes, work load evaluation, counseling and development and training.

Giles and Mossholder (1990) argue that while the context in which appraisal occurs has been designated as a source of considerable influence in the appraisal process, relatively little research has been conducted on the environmental issues. The researchers attempted to extend the development of measures that assess contextual aspects of performance appraisal; to investigate relationships between system contextual variables and employee reactions to performance appraisal; and, to assess the extent to which system contextual variables were related to employee satisfaction. The study confirmed that commonly used reaction scales of fairness, satisfaction, perceived utility and perceived accuracy did indeed represent appraisal reactions.
Murphy and Cleveland (1991) advise that the effectiveness of all human resource systems including performance appraisal need to be evaluated. Murphy and Cleveland maintain that the psychometric indices and rater error measures most often used to evaluate ratings are not adequate criteria for evaluation of performance appraisal system. As an alternative they suggest: 1) developing information on employees from a variety of sources and maintaining adequate performance documentation for review; 2) developing methods for measuring the accuracy of ratings; 3) determine the aspects of accuracy that are most relevant to the various uses of performance appraisal; 4) determining the conditions under which so-called rater errors are beneficial versus harmful; 5) developing practical methods for establishing indifference curves among qualitatively different outcomes of appraisal; 6) developing methods for monitoring changes in the value of an organization's members that will necessitate changes in the appraisal system; 7) determining whether reaction criteria are important and 8) determining whether utility estimates provide useful and credible information. The authors further state that performance appraisals are most likely to be perceived by employees as accurate and fair when the following criteria are met: 1) Appraisals are conducted frequently; 2) There is a formal system of appraisal; 3) Supervisors have a high degree of job knowledge; 4) Ratees have an opportunity to appeal ratings; 5) Performance dimensions are seen to be highly relevant; 6) Action plans are formed for dealing with present weaknesses; and 7) The organizational climate is cooperative rather than competitive.

Research is only beginning to address how context affects employees, perception of appraisal, reactions to appraisal, outcomes and how appraisal purposes (administrative vs. developmental) moderate these relationships. The limited research considering organizational context has focused on system design and characteristics,
system management, and other important performance appraisal issues including fairness and justice issues.

Reactions may even contribute to the validity of a system (Ostroff26, 1993). Cardy and Dobbins27 (1994) suggest, “With dissatisfaction and feelings of unfairness in process and inequity in evaluations, any performance appraisal system will be doomed to failure”.

Thomas and Bretz28 (1994) conclude that performance appraisal continues to be a vexing human resource challenge that the academic research world has not adequately addressed. The focus of academic research on appraisal accuracy, rating errors, or an understanding of the cognitive processes used in the appraisal process are not considered by practicing managers to be major organizational concerns. The authors called for a transition from laboratory studies into the organizational world but realized the lack of access to organizational settings continues to hamper research.

Researchers have suggested that reaction to performance appraisal is critical to the acceptance and use of a performance appraisal system. Murphy and Cleveland29 (1995) stated, “Reaction criteria are almost always relevant, and an unfavorable reaction may doom the carefully constructed appraisal system.” They referred to employee reaction to appraisals as one class of neglected criteria that might be considered in evaluating performance appraisal systems.

Summary of the research points towards the following five areas as measures of an efficacious performance appraisal system: 1. Determines pay; explains and communicates pay decisions, 2. Provides the subordinate with development information and support, 3. Fosters mutual task definition and planning of future work goals, 4. Documents and recognizes subordinate’s performance and 5. Allows the subordinate to provide feedback about feelings, supervision and definition of work. Other variables that may influence performance system
effectiveness include the type of performance standards employed, the frequency of evaluation; the presence of written administrative procedures; and existence of an appeal process.

Tziner et al.\textsuperscript{30} (1997) measured political considerations in performance appraisal to determine the extent to which distortions in ratings were present. Their study investigated evidence that rating inaccuracy has more to do with deliberate volitional distortion of ratings than lack of training or ability. Deliberate distortion of ratings includes raters' conscious efforts to produce ratings that will achieve personal goals such as avoiding negative consequences; avoiding confrontations or bad feelings with employees; or portraying the image of a caring boss.

Martin and Bartol\textsuperscript{31} (1998) discussed the need to monitor a performance appraisal system to keep it responsive to the needs of the organization. The major actions required to maintain a performance appraisal system include three major categories: controlling the system; monitoring the system; and furnishing feedback to those who use the system. Control of the system includes the more technical aspects of the system such as rating techniques, rating periods, rater training, and development of performance standards. Monitoring the system can include a review of the quality of performance standards; evaluation of the actual conduct of the appraisal process and interview; and, analysis of the intended, perceived and actual use of the system. Other factors in monitoring the system include review of the actual quality of ratings to check for rater biases, inconsistencies, rating inflation and investigation for any adverse impact as a result of the system. The third primary area to monitor is that of the amount and quality of feedback generated as part of the performance appraisal process.

Tziner and Murphy\textsuperscript{32} (1999) studied the attitudes of managers towards performance appraisal and their organizations. Raters who showed low levels of
confidence with the system were more likely to rate employees unusually high and fail to discriminate well among ratees. Raters who showed higher levels of attitudinal commitment or who perceived more risks associated with distorting ratings tended to give lower ratings and discriminate more between raters and/or dimensions.

According to Keeping and Levy\textsuperscript{33} (2000) employee reactions toward performance appraisal may be considered important for a number of reasons. First, reactions are of great interest to practitioners. Second, while reactions have been theoretically linked to determinants of performance appraisal success and acceptance they have been overlooked in the research. These issues are both within the context of the gap between research and practice that has been noted in the performance appraisal literature by a number of researchers. Keeping and Levy examined the measurement of performance appraisal reactions. They investigated how well commonly used reaction scales, representative of those used in the field, measured the substantial constructs of satisfaction. They found that these scales did a "favorable" job of measuring appraisal reactions. In addition, they found that the data also fit a higher order appraisal reactions model. Among the reactions investigated were satisfaction (with the system and session), fairness (procedural and distributive justice) perceived utility and perceived accuracy.

Tziner et al.\textsuperscript{34} (2001) report that attitudes and beliefs toward the organization and about the appraisal system affect how ratings are done and how feedback is handled. These attitudes and beliefs have an influence on the accuracy and usefulness of ratings. Their finding showed that beliefs about the performance appraisal system and rater orientation toward the system explained tendencies to give higher versus lower ratings and discriminate between ratees and rating dimensions.

Employees want to be measured against objective standards that all parties involved in the process understand. If the employees are rewarded and recognized for
outstanding performance, they are more likely to work with the goal of high level achievement.

As per Howard (2003), the new interest in creating high-performance organization suggests a need to reconsider how employee performance is viewed. The 'meets expectation' rating now has to shift to stretch goals and competencies that contribute to success in this environment. The role of the manager in raising the bar and in creating a supportive culture also must be considered. Manuel et al. (2004) report on the basis of prior research that training raters to recognize errors will increase rater accuracy and that employee participation in goal setting is more effective than assigning goals. Theory based research suggests ways to help raters recognize expected performance and enable employees to self regulate their pursuit of goals. Feedback to the employee certainly improves efficiency to both the organization and the employee while allowing both to prosper.

Further, employees must be included in the entire performance appraisal system process in order to be successful. These individuals should be involved in this process, participating in all facets including writing job descriptions as well as identifying necessary goals and standards for the process. It has been determined that personnel that are intimately involved in the process will trust the appraisal system and be committed to their own development.

Prowse and Prowse (2009) reviewed the literature on the development of appraisals. They argue that the critical area of management development that was identified as a critical success factor in appraisals has been ignored in the later literature evaluating the effectiveness of performance appraisals. The review identifies the lack of theoretical development in appraisal and argues for the psychological approach analysis and a more critical realization of appraisal, re-evaluating the
challenge to remove subjectivity and bias in judgment of appraisal. Further evaluation of key interpersonal skills is required for appraisal systems to develop performance.

2.5 Performance Appraisal for Scientific Personnel in Research and Development (R&D) organizations

The R&D Management literature indicates that the management of research scientists is different from that of other employees. Among the factors used by organizations to evaluate the performance of scientific or technical staff are number of publications, quality of publications, citations by other authors, number and quality of technical reports, number of papers presented at conferences, patents disclosed or granted, quality of work as viewed by supervisor, peer rating (both inside and outside the Laboratory), reputation in the scientific/technical community, overall certification to organization’s goal, demonstrated creativity or original approach and awards or honors received from scientific societies. There are several validity and reliability problems associated with these measures such as the unequal quality of publications and the Halo effect on subjective ratings which have to overcome or neutralize to ensure a valid evaluation.

One factor that affects both objective and subjective measures of R&D performance is Merton’s concept of the “Mathew Effect”. Merton (1968) observes that those who have achieved considerable recognition are given even greater recognition in the future, while those who have not made a name for themselves find that recognition tends to be withheld. In effect, a misallocation of credit for accomplishment occurs whereby earlier accomplishments are underrated, while later ones are overrated. In the case of subjective ratings, the prior reputations of the ratee may overwhelm other consideration in the evaluation process or may intimidate the rater.
In a review of papers concerned with measuring the performance of researchers, Edward and McCarray\textsuperscript{39} (1973) state that the basic drive for accurate evaluation methods stems from the need to distinguish between an above average scientist and an average one. They pose the questions, should a scientist be judged on the basis of contribution to science or to their employing organizations or to both? Is it possible to evaluate a scientist completely on objective factors or must one consider subjective evaluations as well? They conclude that scientific output is multidimensional and cannot be satisfactorily measured by any one criterion alone.

Five basic categories of obstacles that have been identified to evaluating or measuring the productivity of scientists and engineers in R&D (Ruch\textsuperscript{40}, 1980):

1) Difficulty of defining the output or contribution that is made by a knowledge worker;
2) Overcoming the tendency to measure activities (e.g. number of papers or reports produced) rather than results or impact of the scientific/technical output;
3) Matching of inputs to outputs or impact within a reasonable time frame. Output or impact may occur years later;
4) Quality dimension in the measure; and
5) Concept of effectiveness as well as efficiency in the productivity measure.

As Keller and Holland\textsuperscript{41} (1982) point out in the search for valid measures of R&D professional performance, the pendulum often swings between a preference for objective indicators such as number of publications or patents and subjective performance ratings by peers, supervisors or by the individual in question. They conclude that, however, neither objective nor subjective performance measures have an inherent superiority.

Domsch et al.\textsuperscript{42} (1983) reviewed the current state of knowledge on peer assessment of performance with special reference to industrial R & D organizations. They are of the view that peer review can add value to various techniques of
assessment (nomination, ranking and rating) and certain obvious objectives to peer rating applying also to supervisory review. Their conclusion is that peer review is a useful adjunct to appraisal but only if it is completely acceptable to all involved.

Gratton\(^43\) (1987) recommends the use of a career aspirants programme based on exposure to the process operating in an assessment Centre. This is a systematic procedure using an array of personnel evaluation techniques through which a person is made aware of the demands of a management job and managers made aware of the likelihood of that person’s meeting them. In contrast, good scientists adopt an innovative, unconstrained, independent approach to their work, name high self esteem, are not much interested in people and above all, prize technical soundness. Persons likely to make good R&D manager differ qualitatively from good researchers.

Badawy\(^44\) (1988) in his review of literature on managing scientists and engineers finds that there are criteria for effective performance appraisals. There are many evaluation tools and appraising performance of scientists is difficult.

Jain and Triandis\(^45\) (1990) have pointed out that since there are differences in the goals and aspirations of engineers and scientists, the concept of employee contribution needs to be differentiated for engineers and Scientists. R&D work is complex and doing any one thing well is unlikely to provide a clue to the total performance. Thus, rather than observing an individual’s specific performance, the supervisor is much more likely to observe large choice of performance such as the presentation of a research plan. On completion of the project, usually there are products of groups rather than individuals. It is then difficult to know how much the particular scientist has contributed to the group product. The authors further highlights that as many R&D organizations are non profit enterprises, output measures are more complex. Output measures for a research organization can be subjective or objective,
quantitative or non quantitative, discrete or scalar and can include some measure of quality. While the measurement of quality requires extra effort and, at times human judgment, this dimension of output should not be ignored. Since R&D organizations have multiple objectives and their output are often incommensurate, the output measures are usually non quantitative and subjective. Quantitative measures for output elements are usually in different units, thus defying precise comparison between different quantitative outputs. It might be feasible to combine a multi dimensional array of indicators into aggregate units, which could then provide trends, indicators and patterns of the individual (and organizational) output measures. One suggested categorization of output measures includes the following: a) Process measures (related to activities carried out in an organization useful for the measurement of the current short run performance); b) Result measures (stated in measurable terms, end-oriented) and c) Social indicators (stated in broad terms) related to overall objectives of the organisation rather than specific activities, useful for strategic planning.

There are several metrics useful for helping appraise the performance of scientists and engineers (Cordero46, 1990, Werner and Souder47 1997). Some of these metrics are quantitative: a publication is a sign of innovation, the number of times the publication is cited is a measure of the quality of publication, or patent is a sign that innovation has marked potential, the number of times or patent is cited and the revenues generated by the patent are a measure of the importance of patent. Some of these are qualitative: self evaluation and evaluation by experts, managers, peers and customers that rank individual on one or more dimensions of performance and evaluate the extent the individuals accomplish previously agreed upon objectives. Because individual metrics are imperfect, multiple measures are typically used.
Domm et al.\textsuperscript{48} (1990) while writing on R&D managers’ expectations state that managers evaluating the performance of technical professionals use multiple data sources to arrive at a single conclusion. While both objective and subjective standards are used, complex technical tasks foster the use of judgment and the use of rater cognitive set or expectations within the evaluation process. Managers of professional employees of 18 companies were assessed to determine attributes associated with the most and least-preferred technical professionals on 16 variables (scales) viz. temperament, communication, motivation, gregariousness, attitude, loyalty, intellectual capacity, responsibility, technical expertise, interpersonal skills, supervision. Consistent expectations were found. Cognitive set provides expectancy model of desired professional behaviour found to be consonant among managers of technical professionals. As personal expectations impact the formal performance evaluation process, it would seem beneficial to convey this expectancy model to technical professionals and to research the degree to which factors within the model override the use of formal performance measures.

Drucker\textsuperscript{49} (1999) outlines six major factors that determine knowledge worker productivity: 1) Knowledge worker productivity demands that we ask the question “What is the task?” 2) It demands that we impose the responsibility for their productivity on the individual knowledge workers themselves. Knowledge workers have to manage themselves. They have to have autonomy; 3) Continuing innovation has to be part of the work, the task and the responsibility of knowledge workers; 4) Knowledge work requires continuous learning on the part of the knowledge worker; 5) Productivity of the knowledge worker is not- at least not primarily- a matter of quantity of output. Quality is at least as important; 6) finally, knowledge worker productivity requires that the knowledge worker requires both seen and treated as an asset rather than a ‘cost’.
Grote\textsuperscript{50} (2000) cited the case of the U.S. Air force Research Laboratory in Dayton, Ohio. The Laboratory managers conducted annual reviews of their 3200 scientists and engineers. They found that nearly all the appraisals were positive - not a single person had been rated ‘marginal’ thus resulting in general ‘performance inflation’. Clearly, such uniformly glowing appraisals are useless in evaluating the relative merits of staff members. The scientists of the lab were put in charge of creating the new system. Traditional systems ask supervisors and subordinates to agree on a list of job elements at the start of the year and to base the end of year appraisal on how well those elements were performed. But that approach tends to automatically generate high rating since, in doling out jobs supervisors take into account workers past performance and duly assign tasks they are fairly certain the workers can complete successfully. The scientists proposed a radically different approach. Assume that all staff members are performing well and evaluate them on the contribution their particular jobs make to the mission of the organization. The laboratory now encourages scientists and engineers to take on responsibility as they can handle. They naturally gravitate to the toughest jobs they can completely perform and their pay is based on the value of those jobs to the laboratory.

Kim and Heungshik\textsuperscript{51} (2002) put further the following lessons based on a large scale survey covering over 1200 R&D Scientists and engineers in Korean R&D Organisations. Although most of the R&D personnel prefer their compensation based on performance, they indicate that lack of fair performance evaluation system could be the biggest obstacle toward implementing such a compensation scheme. They also suggest that a fair performance evaluation should utilize more behavioral and quantitative measures such as leadership and mentoring for younger researchers, and bottom-up (e.g. R&D researchers’ evaluation of their own bosses, say R&D managers) as well as horizontal (e.g. peers and/or colleagues) evaluation schemes.

Please purchase PDF Split-Merge on www.verypdf.com to remove this watermark.
The major reason for a different management approach is the many unique features found in an R&D environment. Some of these are (Clarke 2002): a) the uncertainty associated with R&D activities; b) the difficulty in assessing the contribution on impact of research results on the company/organization and on the advancement of science/technology in general; c) the rapid changes in science and technology that result in an ongoing battle by researchers to stave off technological obsolescence in both people and equipment; and d) the values, expectations and attitudes of research scientists and engineers that differ in many aspects from those of other professionals (e.g. Need for autonomy and peer recognition).

Farris and Cordero (2002) reviewed literature on managing Scientists and engineers after 1988. They have reported that traditionally functional managers and senior R&D professionals appraise the performance of scientists and engineers. When these scientists are assigned to project teams, however, 360 degree feedback often occurs. Successful organizations approach evaluation of performance in a formal and systematic way.

Katz (2005) while writing on motivating technical professional views that although R&D professionals may complain the most about inadequate communication and feedback about job performance and career opportunities, it is not the most powerful task dimension for establishing high levels of work motivation. The most critical dimension for increasing motivation is task significance. One of the problems in managing and motivating R&D professionals is that there are at least two ways of looking at multidimensional task dimensions. While the organization orientation is to work on projects that are important to the organization, professional’s orientation is to work on projects that are exciting within the profession. R&D professionals like to have operational autonomy and learn and develop new skills and ability to become a
contributing member at the profession rather than that of organization. They look forward to establishing ‘professional identity’ towards a successful technical career.

MacCoby (2006) puts forth that there is no analogue to knowledge work anything like the Toyota model is to car manufacturing. Every attempt to make it better is or should be an experiment. Less complex knowledge work can be dramatically improved with new tools. As the knowledge work assumes more complexity, the productivity depends more on the qualities of individual workers. MacCoby in his article cites Thomas Davenport who has found that the most effective knowledge workers are the best learners. A R&D organization is a part of socio-technical systems that has a purpose and business strategy. The more the system’s parts, including people, are aligned with this purpose, the more effective the system will be. Davenport recognizes that the best way to get knowledge workers to buy into new policies and process is to involve them in designing their own work.

Badawy (2007) reports on appraising performance of technical professionals for the last 50 years that technical professionals generally believe R&D cannot be controlled in the traditional managerial scale; it can only be monitored, measured and evaluated. They further believe that only other technical professionals can monitor, measure and evaluate R &D performance properly. Difficulty in establishing goals in R&D, standards of performance because of creative nature of the activity and the fact that it usually lacks precedent and labour intensive character of R&D are the problems faced by technical managers in measuring productivity of R&D. There is very little agreement as to what constituents scientific output and as to what measures should be used to reflect the output. The instruments used to obtain performance scores range from gross overall performance ratings to rather complex indicators of the quantity of written output and the ability and creativity of the scientists output. Research has clearly revealed that scientific output is multidimensional and cannot be
satisfactorily measured by one criterion alone. The author additionally observes that in spite of the difficulty of appraising performance and productivity of technical professionals, R&D performance should somehow be related to the financial performance of the company. There is no doubt that appraising a researcher’s performance will be more subjective than that of a design engineer, for example, whose performance can be directly compared to often people performing similar tasks. However, by breaking down the researcher’s activities into several parts, the supervisor and researcher hopefully can agree as the researcher’s performance in each part. Badawy further observes that there has been little agreement on operational definitions of scientific performance, let alone what components to include or who should measure it. In measuring the individual’s scientific performance, the research available utilizes one or more of the following measures: overall performance, quantity of written output, quality of output, and creativity of output. Criteria used to measure scientific performance include productivity in written work, recent reports, originality of written work, professional society membership, judgment of actual work output, creativity ratings by high level supervisors, likeableness as a member of a research team, visibility, recognition for organizational contributions, status-seeking tendencies, current organizational status and contract monitoring load.

Armstrong\textsuperscript{57} (2008) in his book ‘performance management’ states that measuring performance is relatively easy for those who are responsible for achieving quantified targets, for example sales. It is more difficult in the case of knowledge workers, for example scientists. But this difficulty is alleviated if a distinction is made between the two forms of results – outputs and outcomes. An output is a result that can be measured quantifiably, while an outcome is a visible effect that is the result of effort that cannot necessarily be measured in quantifiable terms. But all jobs produce outcomes even if they are not quantified. It is therefore often necessary to measure
performance by reference to what outcomes have been attained in comparison with what outcomes were expected and the outcomes may be expressed as quantitative terms as a standard on level of performance to be attained.

2.6 Perspective of perception of fairness of and satisfaction with Performance Appraisal

There has been paradigm shift in the focus and direction of research on approaches towards evaluation of efficacy of performance appraisal. Employee perceptions of fairness interlinked with their satisfaction of the appraisal process have shown to be pervasive determinant factors of evaluation of performance appraisal efficacy. This development has led researchers and practitioners to take a more comprehensive view of performance appraisal system efficacy and evaluation of systems which include these factors. One common theme of recent research is that attitudes of the system's users toward the process determine to a large degree the ultimate effectiveness of a performance appraisal system.

Fairness of performance appraisal has been studied by a number of researchers over time. Study of perception of fairness is based on the original social justice theory proposed by Adam (1963). It suggested that people perceived social exchanges to be fair when they felt that their contributions were in balance with their rewards. Early research focused on this perceived fairness of resource allocation decisions such as the level of one's pay or departmental budget allocation. Derived from "equity" theory, this became known as distributive justice because it involved the allocation or distribution of resources. Subsequent research indicated that people will accept a certain amount of unfairness in distribution if they perceive that the process by which the distribution decisions were made is fair. Procedural justice is the term used to describe this phenomenon.
Significant research has been conducted in the areas of procedural and distributive justice. Much of the research relies on a series of studies by Thibaut and Walker\(^{59}\) (1975) that investigated individuals' reactions to various dispute resolution techniques. The primary focus of Thibaut and Walker's work emphasized the amount of influence an individual had in the decisions that were made and the process used to make the decisions. The opportunity to present information relevant to a decision enhances judgments of the fairness of the decision making process. Thibaut and Walker termed this as the “process control effect” while Folger\(^{60}\) (1987) referred to it as the “voice” effect. The voice or process control effect may be the best documented phenomenon in procedural justice research. Both pre-decision and post-decision voice led to higher fairness judgments than no voice. Pre-decision voice was found to lead to higher fairness perceptions.

Landy et al.\(^{61}\) (1978) studied employee perceptions of the fairness and accuracy of a performance appraisal system. The researchers found that frequency of evaluation, identification of goals to eliminate weaknesses, and supervisory knowledge of a subordinate’s level of performance and job duties were significantly related to perceptions of fairness and accuracy of performance appraisal. Their results confirmed traditionally held perceptions that performance appraisal should be done as frequently as possible, that the supervisor should work with the subordinate to agree on responsibilities; and, that the supervisor should devote sufficient time to observe and evaluate employee’s performance.

Leventhal\(^{62}\) (1980) extended the discussion of distributive and procedural justice beyond the process to include specific distributive mechanisms and procedural factors other than process control.

In a study of 367 Washington state government employees, Lovrich et al\(^{63}\)
(1980), found that both ratees (58 percent) and raters (71 percent) believed that participative performance appraisal was a fairer way of conducting appraisals than non-participative methods. They also found that, if given a choice, raters and ratees would choose participative performance appraisal over a non-participative type of system.

A major problem for organizational leaders is that the performance appraisal process and the performance evaluation system are often perceived as both inaccurate and unfair. A fair evaluation is one that contains certain procedural elements regardless of the outcomes of the evaluations themselves.

Greenberg\textsuperscript{64} (1986) studied 217 private sector middle managers and asked them on an open-ended questionnaire what single factor made a recent performance evaluation fair or unfair. Factor analysis of the results indicated that soliciting employee input, two-way performance interview communication, and the ability to challenge or rebut the performance ratings account for a significant proportion of the variance in perceived efficacy of the performance appraisal system.

Greenberg was one of the first researchers to apply organizational justice theory to performance evaluation. He posed the basic research question as to what makes a performance appraisal appear to be fair. Further, he considered whether it is what one receives or how the decision is made, or both, that makes performance appraisals seem fair. Beginning with the two distinct concepts of procedural and distributive justice he proposed seven categories that contributed to perceptions of fairness. Five procedural categories included supervisors soliciting input prior to evaluation and use of the input during evaluation, two-way communication between supervisor and subordinate during the appraisal interview, the ability of an employee to dispute or challenge a rating, rater characteristics such as
consistency in applying standards, and rater familiarity with the work being rated. The occurrence of performance based ratings and pay or promotion outcomes based on the ratings were included in the distributive category. Greenberg’s work showed when certain conditions were in place employees was more likely to accept a performance appraisal system and believe that their performance was rated fairly. Identifiable processes within the appraisal system influenced perceptions of appraisal fairness and this influence was independent of the perceptions that the actual rating was favorable. The conditions which helped predict the perception of fairness in the process included frequency of performance feedback, supervisor familiarity with work performance, the opportunity of employees to express their feelings during a performance review and the setting of new performance goals.

Domsch et al.65 (1986) conducted a study among R & D executives assuming that the general confidence variable is of critical importance for the effectiveness of performance appraisal systems. They intended to develop multi-items measure of the general confidence attitude and to identify correlates of this attitude in the light of the prior literature in order to facilitate the successful introduction and use of appraisal systems in an R & D context. Furthermore results indicate that agreement with different appraisal purposes is not always positively related to general confidence in formal performance appraisal procedures.

Beyond distributive and procedural justice, the third form of justice known as interactional justice proposes that the quality of interpersonal treatment received during the enactment of organizational processes and distribution of organizational outcomes is an important contributor to fairness perceptions. Bies and Moag66 (1986) defined interactional justice as the fairness of the interpersonal treatment that one receives at the hands of an authority figure. Bies67 (1987) extended this idea by
adding the concept of social justice to the factor. Early studies of interactional justice focused on the social accounts or explanations that agents (most often leaders) gave for their decisions and actions. The study of interactional justice focuses on how formal agents of the organization treat those who are subject to their authority, decisions, and actions (Cobb et al. 1995). How leaders enact procedures and treat their followers and why followers react the way they do to leaders acting fairly or unfairly are topics that have received increasing attention. Leaders often come to personify the organization for many of their followers. As such, subordinates are likely to assess the fairness of an organization’s procedures by the treatment they receive from their leaders. Another research has shown that fair treatment by one’s leaders communicates that subordinates have higher standing in the organization even when they face disappointing outcomes. Described as "interactional" justice, this concept has been included as an interpersonal aspect of procedural justice and also as a distinct construct along with procedural and distributive justice. These researchers contend that interactional justice can be understood as separate from procedural justice on the grounds that it represents the enactment of procedures rather than the development of the procedures themselves.

Cropanzano and Folger (1989) attempted to integrate distributive and procedural forms of justice in a referent cognition theory. The goal of the theory is to describe the role that decision-making procedures play in shaping perception of unfair treatment. The theory predicts that people will react positively to an unfair outcome if the procedures used to determine the outcomes were fair and that they will react negatively if they perceive the procedures as being unfair.

The concepts of procedural and distributive justice are relatively well accepted in the study of organizational justice. The third component of justice, the
interpersonal, social or interactional factor, is acknowledged, but is not integrated consistently in the formulation of justice models. Researchers have proposed a variety of models ranging from the two-factor distributive and procedural factor model excluding interactional type justice to two and three factor models incorporating interactional justice as part of procedural justice or as a stand-alone component.

Folger et al.\textsuperscript{70} (1992) used a "due process" metaphor to apply the concept of justice to performance appraisal. Three essential factors were used to describe nine elements of a procedurally fair system. The three factors included: adequate notice, fair hearing and judgment based on evidence. Adequate notice requires that organizations set, publish, distribute and explain standards and criteria to employees before the actual evaluation and rating. It also includes employee input in developing performance standards and allows employees to question why and how objectives should be met as well as specifying timely performance feedback on a regular and recurring basis.

The factor of fair hearing requires a face-to-face evaluation meeting or interview with the rater and a performance evaluation based on adequate opportunity to observe and evaluate employee behavior and work product. The right of the employee to provide a self-assessment or other input to the appraisal and to challenge the rating is also included. Judgment based on evidence requires raters to apply standards honestly and consistently across employees and to do so without bias or pressure. The opportunity for the employee to question, discuss or appeal the rating is also indicated.

The empirical work by Greenberg as mentioned earlier in the review clarify the distinction and independent contribution of distributive and procedural forms of justice to other organizational attitudes and behaviors. The due process model proposed by
Folger, Konovsky and Cropanzano and the subsequent empirical work show that attention to performance appraisal processes has an impact on several organizational outcomes.

In their review of performance appraisal research, Bretz et al.\textsuperscript{71} (1992) indicate that the most important performance appraisal issue faced by organizations is the perceived fairness of the performance review and the performance appraisal system. Their findings suggest that most employees perceive their performance appraisal system as neither accurate nor fair. The appraisal process can become a source of extreme dissatisfaction when employees believe the system is biased, political or irrelevant.

Greenberg\textsuperscript{72} (1993) emphasized the need to consider more fully the social determinants of fairness that were not recognized by the prevailing emphasis on the structural aspects of outcome of distributions and procedures. He proposed taxonomy of justice classes formed by cross cutting the two commonly accepted categories of justice, procedural and distributive, with two focal determinants, social and structural. The distinction between social and structural determinants is based on the immediate focus of the just action. Structural determinants reflect the situation whereby justice is sought by focusing on the environmental context in which the interaction occurs. Structural determinants ensure fairness by structuring a decision-making context. The social determinants of justice focus on the treatment of individuals and help ensure fairness by focusing on the interpersonal treatment one receives. Greenberg's four proposed classes of justice include: systemic (structural-procedural); configural (structural-distributive); informational (social-procedural); and, interpersonal (social-distributive).

The importance of the social aspects of organizational systems such as
performance appraisal cannot be under-estimated (Greenberg, 1993). This may be particularly true for performance appraisal where the inter-personal relationships and perceptions of raters and ratees are significant to the results of the process. Some researchers have grouped structural and social determinants without making a distinction between them. An integrated framework is needed that can more clearly distinguish between the various perceptions of organizational justice. Such a distinction is important to researchers in the area of justice as it applies to systems such as performance appraisal as well as to practitioners seeking to understand the actual use and dynamics of performance appraisal in an organization.

As in the traditional models of organizational justice, the distributive justice perceptions in the four-factor model concern outcome allocations while the procedural justice perceptions concern how allocation decisions are made. The structural components determine the “decision making context,” for processes and outcomes, while the social components determine the quality of interactions during the communication of processes and outcomes (Greenberg, 1993).

Subsequent work by Taylor et al.73 (1995) showed that the due process model is consistent with the procedural justice concept. According to these researchers, employees involved in a due process performance appraisal system displayed more favorable reactions regarding the perceived system fairness, appraisal accuracy, attitudes towards the system, and intention to remain with the organization. Managers also responded positively reporting greater ability to resolve work problems, satisfaction with the system and less distortion of appraisal results to further their own self interests.

Korsgaard and Roberson74 (1995) examined subordinate voice (the practice of allowing individuals who are affected by a decision to present information relevant to
the decision) in creating positive attitudes in the performance appraisal context. Instrumental and non-instrumental voices were studied. Instrumental voice is described as the perception of indirect control over decisions when direct control is impossible (Thibaut & Walker 1975). Non-instrumental voice refers to the idea that “voice” is valued intrinsically regardless of whether the impact influences the decision. Their findings indicate that perceptions of instrumental and non-instrumental voice were independently and comparably predictive of appraisal satisfaction. Trust in manager was related only to non-instrumental voice. The difference in the impact of voice components on satisfaction versus trust supports the notion that instrumental voice is more important to reactions to allocation decision than to attitudes towards management.

Other researchers reported similar results in which voice affected perceptions of procedural justice, distributive justice and agreement with the decision. These authors concluded that voice affects distributive justice only when it is considered to be influential whereas voice only had to be considered by the decision maker to affect procedural justice.

Roberts and Reed75 (1996) found evidence of a positive relationship between satisfaction and acceptance of performance appraisal outcomes with employee perceptions that their supervisors encouraged participation, assisted in goal setting and provided frequent feedback.

Cobb and Frey76 (1996) studied the effects of procedurally fair leadership and payment outcomes on subordinate reactions to the supervisor. Subordinate’s reactions were measured for perceptions of supervisory fairness (both procedural and distributive) and the subordinate’s relationships with the supervisor. The results indicated that procedurally fair leadership was linked to subordinate perceptions of
leader fairness and on their relationships with the leaders. Subordinates discerned
differences in leadership behaviors that enact procedural fairness. These behaviors
affected subordinate assessment of supervisor fairness and relationships with the
supervisor. The researchers found some evidence that unfair behavior can have
negative effects on favorable outcomes. When leaders act unfairly their decisions are
seen as unfair even when subordinates benefit from them.

Continued investigation of organizational justice related to performance
appraisal has provided additional insight into perceptions of fairness and employee
satisfaction. Tang and SarsField-Baldwin\textsuperscript{(1996)} found a relationship between
distributive justice and personal level job satisfaction and between procedural justice
and organizational commitment. These researchers developed scales for procedural
and distributive justice and used them to predict satisfaction with pay, promotion,
supervision and their performance appraisal. A 22 item scale reflecting aspects of
procedural justice formed five factors labeled: fairness; two-way communication; trust
in supervisor; clarity of performance appraisal process; and, understanding the
performance appraisal process.

Skarlicki and Folger\textsuperscript{(1997)} found that at high levels of interactional justice
the two-way interaction of distributive and procedural justice was not significant
when studying retaliation in the work place. This result implied to the researchers
that when supervisors show adequate sensitivity and concern towards employees,
treating them with dignity and respect, those employees seem somewhat willing to
tolerate the combination of unfair distributions and procedures. A supervisor
personifies the organization for the employees; being able to count on the goodwill of
a well-meaning supervisor (perceived interactional justice) makes up for unfavorable
procedures combined with the unfairness of a particular outcome. Procedural justice
and interactional justice are capable of functioning as substitutes for each other.

Whether considered as part of procedural justice or as an independent construct, interactional justice can be thought of as having at least two components (Cropanzano and Greenberg\textsuperscript{79} 1997). The first subpart is interpersonal sensitivity which prescribes that treatment should be polite and respectful. The second subpart of interactional justice includes explanations or social accounts. Individuals are more tolerant of an unfavorable outcome when an adequate justification is provided (Shapiro et al.\textsuperscript{80} 1994).

Utilizing an organizational justice perspective, Giles et al.\textsuperscript{81} (1997), in a cross organizational study, developed procedural fairness dimensions for the appraisal review session and two contextual domains of the performance appraisal process. The contextual domains were the structure, policies and support characterizing between supervisors and subordinates. In addition, relationships of the three procedural variable sets and their respective individual variables with three fairness (two distributive justices and one global) criteria were assessed. Procedures from the session and system domains were found to be most useful for predicting the fairness criteria.

Cobb et al.\textsuperscript{82} (1997) studied whether, and to what extent, workers see either formal policies and procedures or the organizational agents (their supervisors) who apply them as the source most responsible for the procedural fairness they receive in their performance evaluation. Results indicated that workers perceived shared, yet independent, responsibility for delivery of procedural justice between supervisors and formal policies. The formal policies and procedures related to the structural aspect of justice and the supervisor or agent to the interactional form. Differences in the perception of the source of justice were also found between the groups of exempt
and non-exempt workers. Non-exempt workers perceived formal policies as more responsible for procedural fairness than did the exempt workers. Some support was found for structural dominance of policies and procedures as the source of procedural justice as opposed to the supervisor in the non-exempt group. Overall, the study provided support for argument that the way agents enact and apply policies is as important to the worker’s perception of justice as the formal policies and procedures themselves.

Gosselin et al.$^83$ (1997) conducted study with research questions on appraisal source, feedback issues and performance management process. Results indicated that subjects most trusted their immediate supervisor as an accurate source of their appraisal and they preferred having their prior knowledge of their supervisor’s expectation, receiving ongoing informal feedback throughout appraisal period and receiving formal appraisal at least twice a year. Subjects expressed preferences for developmentally oriented appraisals that were based mostly on the results of work. Preferences were not moderated by ratee experience or gender.

Ahmed$^84$ (1999) investigated the measure of effectiveness that a state agency uses to assess its performance appraisal function. Some of the criteria for assessment as suggested by the respondents included impact on employee motivation, employee satisfaction with the system, employee's perception regarding fairness and objectivity, and the degree to which it provides adequate and valuable feedback.

Gabis and Ihrke$^86$ (2000) reported that leadership credibility of immediate supervisors is significantly associated with whether employees perceive performance appraisal systems as procedurally fair and instrumentally just and appropriate. Their study of county government professionals explored this issue as well as related
issues of job burnout, job satisfaction, manager innovation and cooperation between organizational units. Boswell and Boudreau\textsuperscript{86} (2000) found a significant positive relation between employee attitudes and procedurally just performance appraisals and underscored the importance employees place on fairness.

The literature as well as the direction of recent research indicates that employee reactions towards performance appraisal variables such as fairness and satisfaction will continue to play a role in the evaluation of performance appraisal systems. Organizational justice theory can also be used to help explain the perceptions of fairness of performance appraisal as related to performance appraisal system efficacy.

Cropanzano and Ambrose\textsuperscript{87} (2001) have suggested that procedural and distributive justice may not be distinct constructs as traditionally conceptualized. They offer some evidence that these factors overlap and can affect one another. Individuals can make references about procedural justice based on distributive justice and outcomes information and vice versa.

Bartol et al.\textsuperscript{88} (2001) investigated the impact of alternative appraisal categories available for rating employee performance (rating segmentation) on motivation and perceptions of fairness. The researchers found that the rating system and the performance rating itself affected perceptions of distributive justice.

Erdogan et al.\textsuperscript{89} (2001) argue that in the performance appraisal context, procedural justice can be conceptualized as two-dimensional: system procedural justice and rater procedural justice. Their study indicates that the two factors are independent. Different components of due process were related to different dimensions of procedural justice. Knowledge of performance appraisal criteria and validity of appraisal criteria are related to system procedural justice whereas fair
hearing and performance feedback are related to rater procedural justice. The authors suggest that understanding the source of perceived injustice can help organizations improve overall justice perceptions by focusing improvement efforts on the appropriate source of either rater procedural justice or system procedural justice. Employees expect the organization to develop appropriate performance criteria and communicate these to them. However, it is the role of the supervisor to conduct a fair hearing through performance appraisal and provide feedback.

Leung et al.⁹⁰ (2001) found that fair interpersonal treatment by the supervisor elicits positive attitudinal reactions from recipients towards both the supervisor and the organization. Fair interpersonal treatment had both direct and indirect paths to attitudes towards the supervisor while the effects of just formal procedures were primarily directed to the organization.

Thurston⁹¹ (2001) developed ten scales to reflect Greenberg’s (1993) four-factor taxonomy of justice. These scales were allocated to each factor according to Thurston’s (2001) interpretation of Greenberg’s theory. The social dimensions of appraisal practices can be represented by interpersonal and informational justice perceptions. Perceptions of the way that the rater treats the person being evaluated, such as with respect and sensitivity, concerns interpersonal justice (social-distributive). Informational justice (social-procedural) reflects fairness perceptions based on the clarification of performance expectations and standards, feedback received, and explanation and justification of decisions. An adequate explanation (informational) clarifying performance expectations or a rating is considered fair based on the interactional/social component.

Brockner⁹² (2002) reviewed studies on the effects of outcome favorability and procedural fairness on people’s support for decisions, decision makers and the
organizations. The interactions found indicated that high procedural fairness reduced the effect of an outcome's favorability or people's support, relative to when procedural fairness was low. Brockner et al.\textsuperscript{93} (1997) has earlier suggested that it is not the procedural fairness that interacts with the outcome favorability but rather it is the degree of trust resulting from procedural fairness of others that interacts with outcome favorability to influence employee support.

The relationships between justice perceptions and reactions to performance appraisal have been theorized and found to exist by a number of researchers. Respectful treatment, fair processes, relevant criteria, participation, and information flow all have a positive association with affective and behavioral responses to performance appraisal. Fairness and justice are clearly important concepts in many organizational processes including performance appraisal. Researchers have suggested that additional research is needed to distinguish between the many types of perceptions involved.

Walsh\textsuperscript{94} (2003) conducted study of employee perception of fairness of performance appraisal in two large Government organizations based on four-factor model of organizational justice (Greenberg,1993) as operationalised by Thurston (2001). The relationships of these perceptions to employee reactions indicating satisfaction with key components of performance appraisal were investigated. This study indicated substantial association each of configural justice scale, interpersonal justice scale, informational justice scale and systemic justice scale with the dimensions of satisfaction.

Upon review of over 300 articles on performance appraisal research over the last 10 years, Levy and Williams\textsuperscript{95} (2004) find that research has broadened the traditional conceptualization of performance appraisal effectiveness to include and
emphasize ratee reactions. The influence that the feedback environment or feedback culture has on performance appraisal outcome is an especially recent focus that seems to have both theoretical and applied implications. There appears to be a reasonably large set of distal variables such as technology, human resource strategies and economic conditions that are potentially important for understanding the appraisal process.

Belanger et al.96 (2006) suggest that perceptions of procedural justice are high if there are standards to insure the results of monitoring are accurate and that the organization has appeal procedure to correct unreasonable outcomes. Kavanagh et al.97 (2007), in their study on fairness among public sector employees, find that participation in performance appraisal, attitudes towards the supervisor and knowledge of the appraisal process are positively and significantly associated with employee perceptions of performance fairness.

Narcisse and Harcourt98 (2008) report on the basis of case study that distributive, procedural and interactional justice factors identified in the existing literature influence employee perceptions of fairness in their appraisal. Results suggest that employees also consider four additional justice factors, as yet not formally recognized in the justice literature, one distributive-the consistency in reward distribution and three procedural- appraisal frequency, job relevant criteria, and rater and ratee training.

2.7 Conclusion

It is obvious to note from the above research studies that many authors have had contributed substantially to performance appraisal system highlighting both traditional and modern method of assessing the effectiveness of performance of employees of various levels. But none of these studies highlighting the research
focused on scientific personnel working in Defence Research and Development Organisation (DRDO). Hence, the Researcher identified this as “Research Gap” in the area of Performance Appraisal System. In order to fulfill the gap, the researcher had chosen the topic “Performance Appraisal System for Scientific Personnel in Defence Research and Development Organisation – An Empirical Study” as his research area and hence, this study.
REFERENCES


