

Research Article

Psychometric Errors in Performance Appraisal of Employees analysis from Various Interrelated Perspectives

Ajaz Ahmad Bhat¹, Noorul Amin²

¹Nursing Officer, SKIMS Soura M.Sc Mental Health Nursing.

²Nursing Officer, SKIMS Soura M.Sc CFT, M.Sc Community Health Nursing.

DOI: <https://doi.org/10.24321/2348.2141.202306>

I N F O

Corresponding Author:

Noorul Amin, SKIMS Soura M.Sc CFT, M.Sc Community Health Nursing.

E-mail Id:

noorul.amin@skims.ac.in

How to cite this article:

Bhat AA, Amin N. Psychometric Errors in Performance Appraisal of Employees analysis from Various Interrelated Perspectives. Trends Nurs Adm Edu. 2023;12(2):30-40.

Date of Submission: 2023-08-08

Date of Acceptance: 2023-09-12

A B S T R A C T

Humans are the greatest assets and play a key role in achievement of organizational purposes and goals, it is essential that those superior to other individuals in organizations know what factors motivate humans to reach better performance. If organizational management or leadership takes this important issue into consideration, it should put a great deal of emphasis on HR development in goal setting and attainment of its organizational goals, strategies, and policies.

An important managerial controlling responsibility is determining how well employees carry out the duties of their assigned jobs. This is done through performance appraisals, in which work performance is reviewed. Performance appraisals let employees know the level of their job performance as well as any expectations that the organization may have of them. Performance appraisals also generate information for salary adjustments, promotions, transfers, disciplinary actions, and terminations.

None of the manager's actions is as personal as appraising the work performance of others. Indeed, next to disciplining an employee, performance evaluation is likely the most dreaded task a manager faces. Because work is an important part of one's identity, people are very sensitive to opinions about how they perform. When a supervisor becomes an 'expert' on a worker's performance, employees will often become resistant or even defensive. However, creating a positive environment, involving the employee in the process and discussing performance goals can help mitigate some of the challenges of performance evaluations.

Indeed, when used correctly, performance appraisal can be one of the greatest tools an organization has to develop and motivate staff. It can also encourage staff and increase retention and productivity; however, in the hands of an inept or inexperienced manager, the appraisal process may significantly discourage and demotivate workers.

In addition, because a manager's opinions and judgments are used for far-reaching decisions regarding the employee's work life, they must be determined in an objective, systematic, and formalized manner as

possible. Using a formal system of performance review reduces, but does not eliminate, the appraisal's subjectivity. In addition, the more professional a group of employees is, the more complex and sensitive the evaluation process becomes. The skilled leader-manager who uses a formalized system for the appraisal is better able to build a team approach to professional work.

Keywords: Performance Appraisal, Checklists, Self-Appraisal, Peer Review

Performance Appraisal

The literature continues to burgeon with multiplying terms which refer to Performance evaluation, namely merit rating, appraisal, evaluation, performance report, performance appraisal, personnel evaluation, audit, and survey (Haji Sharif, 1992). Performance appraisal refers to a hierarchy of formal actions aimed at rating the performance of employees in a given time period. It involves evaluating all visible actions in the light of value criteria for decision making purposes. Evaluation of personnel competence is the systematic and disciplined measurement of individuals' work with regard to their accomplishment of assigned responsibilities and duties and determining their potential for growth and development. It is a process through which employee's performance is formally appraised in specific time intervals (Saadat, 1996).

Through evaluating employee performance managers can not only make decisions about employee transfer and redundancies through identification of extra forces, but also can decide about employee promotion, progress, and appointment. Managers identify skill deficiencies, strengths, and weaknesses of employees through PE, and then plan training programs based on the results of such evaluation. PE is also regarded as a tool for evaluation of training programs of organizations and help managers identifies the most effective programs.

Additionally, Employee PE can act as a basis for reward allocation. Performance-based allocation of rewards brings about satisfaction of qualified employees and encourages them to have a longer tenure in the organization. Obviously, being rewarded based on one's performance is appealing to qualified employees. Another major goal in rating performance is providing feedback about outcomes of employees' work. Such feedback should be provided immediately after the event to provide necessary energy and motifs for sustenance of behavior.

Performance Rating and Psychometric Tools

Since the 1920s, many appraisal tools have been developed, all of which have been popular at different times. Since the early 1990s, The Joint Commission has been advocating the

use of an employee's job description as the standard for performance appraisal. It is important, however, to make sure the job description is current and truly reflects the work the employee is assigned to do.

The effectiveness of a performance appraisal system is only as good as the tools used to create those assessments. An effective competence assessment tool should allow the manager to focus on the priority measures of performance. The following is an overview of some of the appraisal tools commonly used in health-care organizations.

Trait Rating Scales

A trait rating scale is a method of rating a person against a set standard, which may be the job description, desired behaviors, or personal traits. The trait rating scale has been one of the most widely used of the many available appraisal methods. Rating personal traits and behaviors is the oldest type of rating scale. Many experts argue, however, that the quality or quantity of the work performed is a more accurate performance appraisal method than the employee's personal traits and that trait evaluation invites subjectivity. Rating scales are also subject to central tendency and halo- and horns-effect errors and thus are not used as often today as they were in the past.

Instead, many organizations use two other rating methods, namely, the job dimension scale and the behaviourally anchored rating scale (BARS).

Job Dimension Scales

Job dimension scales require that a rating scale be constructed for each job classification. The rating factors are taken from the context of the written job description. Although job dimension scales share some of the same weaknesses as trait scales, they do focus on job requirements rather than on ambiguous terms such as "quantity of work."

Behaviourally Anchored Rating Scales

BARS, sometimes called behavioral expectation scales, overcome some of the weaknesses inherent in other rating systems. As in the job dimension method, the BARS technique requires that a separate rating form be developed for each job classification. Then, as in the job dimension rating scales, employees in specific positions work with management to delineate key areas of responsibility. However, in BARS, many specific examples are defined for each area of responsibility; these examples are given various degrees of importance by ranking them from 1 to 9. If the highest ranked example of a job dimension is being met, it is less important than a lower ranked example that is not.

Appraisal tools firmly grounded in desired behaviors can be used to improve performance and keep employees focused on the vision and mission of the organization.

However, because separate BARS are 645 needed for each job, the greatest disadvantage in using this tool with large numbers of employees is the time and expense. Additionally, BARS are primarily applicable to physically observable skills rather than to conceptual skills. Yet, this is an effective tool because it focuses on specific behaviors, allows employees to know exactly what is expected of them, and reduces rating errors. Although all rating scales are prone to weaknesses and interpersonal bias, they do have some advantages. Many may be purchased, and although they must be individualized to the organization, there is little need for expensive worker hours to develop them. Rating scales also force the rater to look at more than one dimension of work performance, which eliminates some bias.

Checklists

There are Several types of Checklist Appraisal Tools

The weighted scale, the most frequently used checklist, is composed of many behavioral statements that represent desirable job behaviors. Each of these behavior statements has a weighted score attached to it. Employees receive an overall performance appraisal score based on behaviors or attributes. Often, merit raises are tied to the total point score (i.e., the employee needs to reach a certain score to receive an increase in pay).

Another type of checklist, the forced checklist, requires the supervisor to select an undesirable and a desirable behavior for each employee. Both desirable and undesirable behaviors have quantitative values, and the employee again ends up with a total score on which certain employment decisions are made.

Another type of checklist is the simple checklist. The simple checklist comprises numerous words or phrases describing various employee behaviors or traits. These descriptors are often clustered to represent different aspects of one dimension of behavior, such as assertiveness or interpersonal skills. The rater is asked to check all those that describe the employee on each checklist.

A major weakness of all checklists is that there are no set performance standards. In addition, specific components of behavior are not addressed. Checklists do, however, focus on a variety of job-related behaviors and avoid some of the bias inherent in the trait rating scales.

Essays

The essay appraisal method is often referred to as the free-form review. The appraiser describes in narrative form an employee's strengths and areas where improvement or growth is needed. Although this method can be unstructured, it usually calls for certain items to be addressed. This technique does appropriately force the

appraiser to focus on positive aspects of the employee's performance. However, a greater opportunity for personal bias undoubtedly exists. In addition, it is time consuming, and some appraisers simply write better than others. Many organizations combine various types of appraisals to improve the quality of their review processes. Because the essay method does not require exhaustive development, it can quickly be adapted as an adjunct to any type of structured format. This gives the organization the ability to decrease bias and focus on employee strengths.

Self-Appraisals

Employees are increasingly being asked to submit written summaries of their work-related accomplishments and productivity as part of the self-appraisal process. Self-appraisal is not easy, however, for many employees. Flescher (2016) notes that many employees ask, "What is the point of this?" "What am I supposed to say?" and "Is whatever I put down going to be used against me?" Flescher argues that employees should view the opportunity to provide a self-assessment as part of the performance appraisal process positively because managers can't possibly remember as well as you can everything you accomplished throughout the year. "By asking you to provide input into your own employee evaluation, it reminds your boss about all the good things you achieved".

Management by Objectives

MBO is an excellent tool for determining an individual employee's progress because it incorporates both the employee's assessments and the organization. Although infrequently used in health care, MBO is an excellent method to appraise the performance of the employee in a manner that promotes individual growth and excellence.

The following steps delineate how MBO can be used effectively in performance appraisal:

1. The employee and supervisor meet and agree on the principal duties and responsibilities of the employee's job. This is done as soon as possible after beginning employment.
2. The employee sets short-term goals and target dates in cooperation with the supervisor or manager, and the manager guides the process so that it relates to the position's duties. It is important that the subordinate's goals not be in conflict with the goals of the organization. In setting these goals, the manager must remember that one's values and beliefs simply reflect a single set of options among many. This is especially true in working with a multicultural staff. Professional expectations and values can vary greatly among cultures, and the manager must be careful to resist judgmental reactions and allow for cultural differences in goal setting.

3. Both parties agree on the criteria that will be used for measuring and evaluating the accomplishment of goals. In addition, a time frame is set for completing the objectives, which depends on the nature of the work being planned. Common time frames used in health-care organizations vary from 1 month to 1 year.
4. Regularly, but more than once a year, the employee and supervisor meet to discuss progress. At these meetings, some modifications can be made to the original goals if both parties agree. Major obstacles that block completion of objectives within the stipulated time frame are identified. In addition, the resources and support needed from others are identified.
5. The manager's role is supportive, assisting the employee to reach goals by coaching and counselling.
6. During the appraisal process, the manager determines whether the employee has met the goals.
7. The entire process focuses on outcomes and results and not on personal traits.

One of the many advantages of MBO is that the method creates a vested interest in the employee to accomplish goals because employees are able to set their own goals. Additionally, defensive feelings are minimized, and a spirit of teamwork prevails.

MBO as a performance appraisal method, however, also has its disadvantages. Highly directive and authoritarian managers find it difficult to lead employees in this manner. Also, the marginal employee frequently attempts to set easily attainable goals. In addition, set objectives may hinder innovation, and they are time consuming to create and challenging to maintain. In addition, although collaboration is intended to be the key to success, the desire to help colleagues may suffer when individual goals are the entire focus of a department.

Peer Review

When peers rather than supervisors carry out monitoring and assessing work performance, it is referred to as peer review. Most likely, the manager's review of the employee is not complete unless some type of peer review data is gathered. Peer review provides feedback that can promote growth. It can also provide learning opportunities for the peer reviewers.

Peer review can be carried out in several ways. The process may require the reviewers to share the results only with the person being reviewed, or the results may be shared with the employee's supervisor and the employee. The review would never be shared only with the employee's supervisor. The results may or may not be used for personnel decisions. The number of observations, number of reviewers, qualification and classification of the peer reviewer, and procedure need to be developed for each

organization. Peer review has the potential to increase the accuracy of performance appraisal. It can also provide many opportunities for increased professionalism and learning.

The 360-Degree Evaluation

An adaptation of peer review, and a relatively new addition to performance appraisal tools, the 360-degree evaluation, includes an assessment by all individuals within the sphere of influence of the individual being appraised. Simply put, it is a mechanism for evaluating someone's performance based on feedback from everyone with whom the individual comes in contact-supervisors, co-workers, partners, subordinates, the general public.

Getting feedback from multiple individuals provides a broader, more accurate perspective of the employee's work performance. This divergent thinking suggests that involving additional individuals in the appraisal process provides unique and valuable perspectives that might otherwise not be considered. When done properly, 360 is highly effective as a development tool. The feedback process gives people an opportunity to provide anonymous feedback to a co-worker that they might otherwise be uncomfortable giving. Feedback recipients gain insight into how others perceive them and have an opportunity to adjust behaviors and develop skills that will enable them to excel at their jobs.

In most 360-degree evaluations, somewhere between 8 and 12 people fill out an anonymous online feedback form that asks questions covering a broad range of workplace competencies. The feedback forms include questions that are measured on a rating scale and also ask raters to provide written comments.

The most challenging aspect of the 360-degree evaluation typically is the evaluators' concerns about confidentiality. When implementing this type of evaluation, it's best to assure other employees that what they share will remain strictly confidential. Likewise, the evaluator should explain to each employee that his or her performance will be evaluated by many people, including those who know their work best.

Psychometric Errors In Performance Rating

Evaluation errors and biases may occur in judgment observations or information process stages. These affect the appropriateness and accuracy of PE. In the context of PE, organizations are assumed to communicate performance standards or goals, discuss how these standards can be met, explain criteria for evaluation, and provide timely feedback (Folger & Cropanzano, 1998). These elements, respectively, correspond to goal setting (Cederblom, 1982; Klein, Snell, & Wexley, 1987), and job relatedness of criteria used to evaluate performance (Cederblom, 1982; Klein et al., 1987). Therefore, existence of an integrated and culture-based

evaluation system as well as organizational strategies, policies, and goals is inevitable. Bias in performance appraisal is problematic since it makes it difficult to make appropriate personnel decisions, such as promotions (Moers, 2005). In fact, the empirical results indicate that performance measure subjectivity is positively related to PE bias. According to Prendergast and Topel (1993), superiors in organizations have incentives to bias the PE. Moers (2005) also found that using subjectivity in PE lead to evaluations that make it difficult to differentiate among subordinates and may result in problems in personnel decisions and future incentives.

Halo Error

As conceptualized in the literature, halo effect is the tendency to rate an employee uniformly high or low in other traits if he is extra-ordinarily high or low in one particular trait. For example, if a worker has few absences, his supervisor might give him a high rating in all other areas of work. Put differently, halo is a tendency to let our assessment of an individual on one trait influence our evaluation of that person on other specific traits. Managers often do this when they have a generally good relationship with the person they are rating and do not want to be too harsh or when they really like an employee and, thus, allow their personal feelings about this employee to influence their performance ratings. This is a very common type of error and is also one that is very difficult to correct.

It is imperative to mention that in contrary to horn error, halo effect is letting a single strength of an employee determine the overall rating. In horn error, however, the individual's performance is completely appraised on the basis of a single negative quality or feature perceived. This results in an overall lower rating than may be warranted. For instance, on the basis of the fact that an employee is not formally dressed up in the office, managers might judge that he may be casual at work too.

Leniency and Severity Errors

The incentives of superiors to bias the performance rating of subordinates stems from the psychological cost of communicating poor performance, favouritism, and preferences for equity in rewards (Prendergast & Topel, 1993). Bias in PE is problematic because it has both direct and indirect costs. The direct costs are associated with higher compensation costs than those warranted by true performance of subordinates. The indirect costs are related to difficulty of making important personnel decisions based on performance ratings and the impact of incentives on motivation.

Just as some professors are known as easy A's, some managers tend to give relatively high ratings to virtually everyone under their supervision depending upon their own

standards, values, and physical and mental makeup at the time of appraisal. Here, the situation is ripe for the abuse of inactive employees. Past studies have empirically shown that leniency error is a critical issue where evaluation-related decisions have a bearing on official decision making, such as rewards and promotion. The opposite happens when managers believe in the tyranny of exact assessment, considering more particularly the drawbacks of the individual and thus making the assessment excessively severe. The former raters might best be described as being generally easy or lenient, while the latter may be classified as being hard or severe in their judgments or ratings. It is worth mentioning that managers also give a good evaluation hoping a poorly performing employee will "grow into" it. However, if everyone is to be rated high, the system has not done anything to differentiate among the employees. There are evidences that raters who feel accountable may exert more care when they rate others (Rosenbaum, Lehman, & Holcom, 1993). In a similar vein, Mero and Motowidlo (1995) found that accountability was negatively associated with rating errors in downward appraisals. Haeggberg and Chen (1999) found similar results in upward appraisals.

Both leniency and severity errors are commonplace in graphical evaluations which use figures to represent evaluation results and where evaluation does not require any written supporting statements. Rankings, in which each ratee's performance is ranked in comparison with other ratees, is a better alternative. In order to rank the ratees, the rater does not have to assign them grades and have concerns over assigning higher or lower grades. Rather, he should discriminate them according to their relative level of performance. Obviously, leniency and severity errors are not an issue in rankings (Dessler, 1997). Both errors skew the performance appraisal results and render a system ineffective and should, thus, be avoided. The following diagram illustrates the distributions of ratings one might get from two different raters, one who is overly lenient and the other who is overly severe. There are evidences that discretion in PE gives rise to a number of problems (Prendergast & Topel, 1993). Moers (2005) found that superiors give more lenient performance ratings when they have discretion in PE due to subjectivity per se or the use of multiple objective PEs. According to Vance, Winne, and Wright (1983), most of the variance in halo and leniency is attributable to the behaviors of raters, rather than the work of the ratees.

Contrast

Superiors, as evaluators, are usually required to rate the performance of a large number of employees. Contrast error occurs when the manager compares an employee's performance to other employees instead of the company

standard. When employees are ranked in comparison, someone must end up at the bottom, even if they are exceeding the company standard. The problem is not the employee. It is the goal or standard that has been set. Moreover, there is an order effect, i.e. individuals who are rated first are rated higher than those evaluated last. If the time gap between the two evaluations is large, the effect is larger. This issue is of great importance in evaluations aimed at electing or employing individuals.

Proximity Error

Proximity error comes about from the way in which the various items have been placed or ordered on the rating form. Sometimes referred to as an order effect, this error illustrates the influence that surrounding items have on the rating one gives a person on a particular item. For example, if the preceding item was a trait on which the individual was given a very favourable rating, the rater may tend to let the favourable response carry over to the next item on the list. There is always the possibility of the reverse reaction occurring. If the preceding items have been on traits in which the worker was generally rated high, the supervisor may give a rather unfavourable score when the rater gets to a trait in which the worker truly deserves only a "moderate" ranking, simply because of the contrast effect of the preceding items.

Central Tendency

This is the raters' tendency to avoid making extreme judgments of employee performance resulting in rating all employees in the middle part of a scale without any consideration of their actual performance. For instance, when the score range is 1 to 7, supervisors tend to avoid assigning scores on either extremes, i.e. 1, 2, 6, or 7. This error may exist even in graphical evaluation scales. It is an attitude to rate people as neither high nor low. Thus, a manager might follow the middle path when he is not comfortable with conflict and avoids low marks to avoid dealing with behavioral issues. This gives the impression that there are no very good or very poor performers on the dimensions being rated. To avoid this problem, supervisors are recommended to use rankings because rankings prevent them from putting many in the centre.

Spill-Over Effect or Past-Record Anchoring

Spill-over effect happens when the present performance is evaluated much on the basis of past performance. If the results of previous PEs of employees were positive, managers will keep rating them overly highly, regardless of their current performance. For example, the person who was a good performer in distant past is assured to be okay at present also. It has been observed that even if employee performance is really low according to current evaluation criteria, managers seldom assign them a mark

lower than one point below their previous evaluation score. For instance, in cases where the employee had previously obtained a score of 29, he would not get a score below 28 even if his actual score is much below. In Jawahar's (2006) study, previous performance was significantly associated with rater criticism ($B = -21, p < .001$) and subsequent performance ($B = .68, p < .001$).

Recency Error

Recency effect is the rater's tendency to allow more recent incidents, either effective or ineffective, of employee behavior to have too much bearing on evaluation of performance. Thus, the employee's most recent behavior becomes the primary focus of the review. This can be extreme on both ends of the spectrum. An example is being critical of an employee who is usually on time but shows up one hour late for work the day before his or her performance appraisal. It is for this reason that keeping accurate records of performance throughout the year to refer back to during performance appraisal time is so critical.

Equity and fairness concerns are particularly crucial factors for understanding how employees react to a particular performance management system. The fairness of both the evaluation and the rewards an employee receives influence employee performance (Akerlof & Yellen, 1988). Equity theory holds that employees will compare evaluations and rewards they receive with those doing similar work. Such comparison can affect both motivation and responses of those who see themselves as not rewarded to the same extent as others performing the same task at the same level. Such perceived inequitable evaluation and pay decisions can have a negative impact on an individual's motivation to perform in the future. A second type of relevant employee perception is fairness of the process used to determine the rewards. These processes are regarded as fair when they are consistent and accurate and lack any kind of bias that favors one group over another (Burney, Henle, & Widener, 2007). Finally, there are evidences that employees' perceived lack of fairness undermines organizational commitment and organization citizenship behaviors (Schminke, Cropanzano, & Rupp, 2002).

Personal Bias

There are evidences that performance appraisals might be influenced by rater's attitudes toward the ratee, i.e. liking and trust in the rater and perceived quality of their relationship. It stems from the way a supervisor feels about each of his subordinate employees and whether he likes or dislikes them. Eighty percent of managers admit that the fact that they like or dislike an employee affects their appraisals. Personal bias may come from various sources, to name just a few, information obtained from colleagues,

considerations of faith and thinking, and social and family background. It includes stereotyping and hostility errors both of which take the rater away from reality, which is the actual behavior of the ratee. Common examples include showing bias based on race or gender. Although rater effect is irrelevant to the evaluation of employee performance of a given task, the fact that a rater likes or dislikes a particular ratee can influence his ratings and determine his evaluation of ratee's performance (Cardy & Dobbins, 1986; Decotiis & Petit, 1978; Dipboye, 1985). Rator effect may affect what the raters observe (Isen, Shalker, Clark, & Karp, 1978); it may influence the attributions raters make about ratees' behaviors (Feldman, 1981); and it may impact the information raters receive from memory at the time of evaluation (DeNisi, Cafferty, & Meglino, 1984; Srull & Wyer, 1980). Tsui and Barry (1986) found that positive affect is associated with higher ratings, i.e. high leniency and negative affect are related to lower ratings, that is, high severity in the types of PE targeted. Surprisingly, Antonio found an interaction between the amount of time raters observe ratees and the extent to which they like them. The influence of rater effect on ratings increases by enhancing the raters' observation of rates.

Apparently, raters are so strongly affected by positive or negative effect that increased observation. This refers to noticing specific behaviors that conform to their affect or retrieving affect consistent with information at the time of evaluation. However, this does not mean that such evaluations should be discounted. Rather, raters should be trained to control their emotional reactions toward ratees (Cardy & Dobbins, 1986). This type of training should not be a one-time event. Rather, it should be an on-going process focused on improving the way in which performance appraisers observe and rate others. Spence and Keeping (2013) argue that it is difficult, if not impossible, for a manager to have an intention for appraisal without considering the performance of whom he is evaluating.

Rater Attitudes and Values

Value system of appraisers has direct influence on their PE. Managers claim that they evaluate on the basis of employees' current performance. However, evidences show that managers' value system is the best predictor of their method of evaluation. Accordingly, employees who give managers a sense of accomplishment are more highly rated. Manager attitudes toward accurate rating is defined as overall favourability or lack of favourability of rendering an accurate performance rating for a particular subordinate, within a particular context, and at a given time. This encompasses general attitudes about rating accurately in the organization and more specific attitudes about rating a particular subordinate for a particular period of time in an organization (Spence & Keeping, 2013). For

instance, if a manager believes that a particular employee will be demoralized and that the organization will not benefit from the evaluation, he will have a negative attitude toward accurate evaluation of the employee. Here, both the value and experience of providing an accurate rating will be negative (Spence & Keeping, 2013).

Fame

An implicit analysis revealed that employees might waste time by fraternizing in the workplace, such as forming work groups and group memberships, instead of doing true work. A group member might be very skilled in creating a positive image in the group but evade his responsibilities and meanwhile pretend to be important to the accomplishment of group tasks.

Employee Appearance

The dominant narrative that emerges from the literature is that employee appearance impacts performance appraisals and employees' salary amount. There are also evidences that attractive appearance brings about 2000 to 2600 salary increase. This indicates that salary of an attractive employee would be approximately 13000 dollars higher than salary of an employee with modest appearance. This is especially rampant among men in male jobs as well as in the elderly.

Discrimination between Insider and Outsider Employees

Members of insider groups are supported and trusted by the supervisor and, thus, receive challenging tasks. Members of outsider group, however, are treated as strangers and are, consequently, assigned unimportant and trivial tasks. The literature shows that insider group members receive higher ranks while, according to evidences, they do not necessarily perform better than others.

Performance Evaluation and Psychometric Errors- Finding a Way Forward

It is argued that employee performance impacts organizational success (Gardner, Wright, & Moynihan, 2011; Wright, Dunford, & Snell, 2001). Accordingly, improvement of employee performance brings about better organizational performance and efficiency. Shepherd, Carley, and Stuart (2009) found that PE impacts four career areas, namely promotion in academic rank (87%), tenure decisions (86%), compensation decisions (82%), and retention of untenured faculty (81%).

An individual's attitude toward a behavior is a core determinant of whether or not that behavior will occur. An attitude represents an individual's overall evaluation of the favorability of a particular behavior determined by the assessment of the consequences of that behavior via two dimensions. One dimension, i.e. instrumental, concerns

whether the An individual's attitude toward a behavior is a core determinant of whether or not that behavior will occur. An attitude represents an individual's overall evaluation of the favorability of a particular behavior determined by the assessment of the consequences of that behavior via two dimensions. One dimension, i.e. instrumental, concerns whether the Obviously, inaccurate performance appraisal will result in employee dissatisfaction (Jawahar, 2010). Nathan, Mohrman, and Milliman (1991) and Jawahar (2010) found that jobrelatedness of criteria was associated with ratee's perceptions of quality of feedback ($r = .49, p < .001$).

Job knowledge, job relatedness of evaluation, goal setting, and suggesting ways to improve performance are proved to be the most important antecedents that affect ratee's reactions. Although it is true that lenient ratings contribute to employee satisfaction with feedback (Jawahar, 2006), past research has demonstrated that some factors, namely rater's knowledge of the ratee's job (Ilgen, Fisher, & Taylor, 1979), relatedness of evaluation (Nathan et al., 1991), giving ratees the opportunity to participate in feedback (Cawley, Keeping, & Levy, 1998), setting goals (Giles & Mossholder, 1990), suggesting ways to improve performance (Keaveny, Inderrieden & Allen, 1987), and evaluating performance in distributive and procedurally fair manner (Jawahar, 2006) are positively related to ratee's level of satisfaction. These further highlight the important role of rater in the appraisal process.

Ford et al. (2011) also found that repeated occupational commitment enhances task performance through surface acting. This highlights that human resource practitioners and managers should examine the factors that influence such commitment (Ford et al., 2011).

PE performed based on the actual performance of an individual can serve to inform a variety of HR decisions, namely pay or recognition, training, promotion and demotion, and adverse action and termination. In order to solve this problem, many researchers have suggested coaching employees on how to perform, rather than merely telling them what they need to do. Coaching, as in athletic issues, involves giving on-going feedback on performance. It is greatly adopted by managers who believe that coaching is more likely to lead to positive organizational outcomes than a "tell and sell" PE (Ford et al., 2011). Some positive outcomes of coaching are that specific goals are set, the relationship between what the person is doing and the outcomes he can expect is clarified, good performance is praised, and the individual is inspired to take action that will result in an improvement in job performance (Ford et al., 2011). For many reasons including legal, coaching as a means for providing performance feedback is replacing or being added to conventional PE hoped to improve not only the performance of employees, but also that

of organizations. Many managers who recognize the importance of coaching based on credible performance data suggest a 360 degree review process. An advantage is that since 360 degree data are gathered from multiple sources, namely boss, peers, customers, direct reports, and other employees, the employee who is evaluated is more likely to believe that performance assessment is not just a biased opinion of an uninformed supervisor, rather it is a collective and, hopefully, convergent opinion by all who have contact with the employee. Generally, 360-degree data are regarded as more objective.

Another practice performed by organizations to affect such performance is pay-for-performance (Gerhart & Trevor, 1996; Gerhart, Rynes, & Fulmer, 2009). It is the pay that changes with some measure of individual or organizational performance (Milkovich, Newman, & Gerhart, 2011). In key psychological perspectives, scholars argue that pay-for-performance (PFP) will affect performance through mechanisms such as instrumentality, i.e. the perceived link between performance and pay (Vroom, 1964) and meeting obligations by the employer (Robinson & Rousseau, 1994). Likewise, in fundamental economic perspectives, pay is considered to exert an influence on employee behavior through creation of transactional norms (Coase, 1937; Williamson, 1981), overcoming of monitoring challenges (Jensen & Meckling, 1976), and the motivating organizationally desirable behaviors. Despite widespread use of PFP, supportive PFP-relevant theories, and a metaanalytic finding that individual-level PFP has a weak but positive relationship with past employee performance (Jenkins, Mitra, Gupta, & Shaw, 1998), PFP efficacy is questioned by both academic researchers and practitioners (e.g., Pink, 2009). There are also strong concerns that PFP is negatively associated with performance (Ariely, Gneezy, Loewenstein, & Mazar, 2009), is not motivating (Pfeffer, 1998), decreases intrinsic motivation (Deci & Ryan, 1980), motivates undesirable employee competition (Deming, 1986), and is difficult to effectively implement (Lawler, 2000).

Yet, another alternative which can help overcome limitations of traditional performance appraisals is mystery shopping. It has two advantages, namely timeliness and unpredictability as to when feedback will be given (Ford et al., 2011). Employees expect to be rated based on their success in fulfilling their individual performance objectives regardless of their grade or position or other non-performance-related criteria. Overall, there are considerable evidences that the less the risk of bias and errors in PE, the more accurate employee PE system, the greater employee satisfaction, and the better their performance and efficiency.

Conclusion

To put it in a nutshell, PE is a managerial mechanism

to obtain organizational development and efficiency. Nowadays, it is also considered a huge competitive advantage. Evaluation is an important tool in HRM whose appropriate accomplishment aids organizations in reaching their goals and satisfying employee interests. In the appraisal process, the manager evaluates employee behavior and performance through measuring and comparing it with predetermined criteria. Then, the manager records the results and informs the employee of them. However, it is rather difficult to perform accurate and appropriate evaluations because evaluation involves judgment about behavior and performance of individuals. It is important to evaluate with as little bias and errors as possible. To accomplish this, a combination of different error-free methods can be implemented. A great majority of HRM theoreticians maintain that prior to any decision making about promotion, salary increase, transfer, nomination, and redundancy, organizations need to obtain bias- and error-free information regarding performance of employees. Managers are assumed to rate behavior according to suitable performance criteria.

In sum, the overwhelming consensus among researchers suggests that the less the errors and bias in performance appraisal, the better employee performance and satisfaction and efficiency of the evaluation system. As with any empirical study, this study has its limitations. According to Moers (2005), the data do not allow us to examine the behavior of individual superiors. Therefore, the analysis assumes that all superiors behave in an identical way. Although psychological research holds that superiors are generally inclined to bias their performance ratings, it might be that super-specific characteristics lead to such behavior. Thus, a more detailed analysis of superiors' behavior in evaluation process would perhaps alter our results. Next, the present research merely explores the impact of subjective performance, while the literature recommends that multiple performance measures should be used because no single performance measure is complete. There are evidences that overall effect of a mixed PE is greater than a solely objective and precise evaluation for people high in autonomous motivation.

Some proposals have been made in the literature for improving PE systems and for minimizing different forms of bias in the PE processes. However, these suggestions may not eliminate bias completely, but help reduce it. Firstly, regarding bias due to organizational position, employees should trust there is as much a chance for a grade 11 to get a level 3, 4, or 5 as for a grade 14 or 15 to receive one. Secondly, HR managers and organizational leaders should try to improve the credibility and fairness of the system through addressing bias that exists in their PEs processes. For example, some agencies use pass/fail systems which

result in limited performance conversations with employees and have no way to connect results to pay for performance, and, thus, increased training for evaluators.

References

1. Akerlof, G. A., & Yellen, J. L. (1988). Fairness and unemployment. *American Economic Review of Public Personnel Administration*, 78(2), 44–49.
2. Cardy, R. L., & Dobbins, G. H. (1986). Affect and appraisal accuracy: Liking as an integral dimension in evaluating performance. *Journal of Applied Psychology*, 71(4), 672–678.
3. Cawley, B. D., Keeping, L. M., & Levy, P. E. (1998). Participation in the performance appraisal process and employee reactions: A meta-analytic review of filed investigations. *Journal of Applied Psychology*, 83(4), 615–633.
4. Cederblom, D. (1982). The performance appraisal interview: A review, implications, and suggestions. *Academy of Management Review*, 7(2), 219–227.
5. Decotiis, T., & Petit, A. (1978). The performance appraisal process: A model and some testable propositions. *Academy of Management Review*, 3(3), 635–646.
6. Deming, W. (1986). *Out of the crisis: Quality, productivity and competitive position*. Cambridge, MA: MIT Press.
7. Feldman, J. M. (1981). Beyond attribution theory: Cognitive processes in performance appraisal. *Journal of Applied Psychology*, 66, 127–148.
8. Folger, R., & Cropanzano, R. (1998). *Organizational justice and human resource management*. Thousand Oaks, CA: SAGE.
9. Ford, R. C., Latham, G. P., & Lennox, G. (2011). Mystery shoppers: A new tool for coaching employee performance improvement. *Organizational Dynamics*, 40(3), 157–164.
10. Gardner, T. M., Wright, P. M., & Moynihan, L. M. (2011). The impact of motivation, empowerment, and skill enhancing practices on aggregate voluntary turnover: The mediating effect of collective affective commitment. *Personal Psychology*, 64(2), 315–350.
11. Gerhart, B., Rynes, S. L., & Fulmer, I. S. (2009). Pay and performance: Individuals, groups, and executives. *Academy of Management Annals*, 3(1), 251–315.
12. Gerhart, B., & Trevor, C. O. (1996). Employment variability under different managerial compensation systems. *Academy of Management Journal*, 39(6), 1692–1712.
13. Giles, W. F., & Mossholder, K. W. (1990). Employee reactions to contextual and session components of performance appraisal. *Journal of Applied Psychology*, 75, 371–377.

14. Haeggberg, R. L., & Chen, P. Y. (1999). Can accountability assure the accuracy of upward appraisals? Paper presented at the Fourteenth Annual Conference for the Society for Industrial and Organizational Psychology, Atlanta.
15. Haji Sharif, M. (1992). Evaluating HR management (1st ed.). Tehran: Abyek Cement Company Publications.
16. Ilgen, D. R., Fisher, C. D., & Taylor, S. M. (1979). Consequences of individual feedback on behavior in organizations. *Journal of Applied Psychology*, 64(4), 349–371.
17. Isen, A. M., Shalke, T. E., Clark, M. S., & Karp, L. (1978). Affect, accessibility of material in memory, and behavior: A cognitive loop? *Journal of Personality & Social Psychology*, 36, 1–12.
18. Keaveny, T. J., Inderrieden, E. J., & Allen, R. E. (1987). An integrated perspective of performance appraisal interviews. *Psychological Reports*, 61(2), 639–646.
19. Klein, H. K., Snell, S. A., & Wexley, K. N. (1987). Systems model of the performance appraisal interview process. *Industrial Relations*, 26(3), 267–280.
20. Javidmehr & Ebrahimpour / *International Journal of Organizational Leadership* 4(2015) 286-302
21. Jawahar, I. M. (2006). Correlates of satisfaction with performance appraisal feedback. *Journal of Labor Research*, 27, 213–234.
22. Jawahar, I. M. (2010). The mediating role of appraisal feedback reactions on the relationship between rater feedback-related behaviors and ratee. *Group & Organization Management*, 35(4), 492–526.
23. Jenkins, G. D., Jr., Mitra, A., Gupta, N., & Shaw, J. D. (1998). Are financial incentives related to performance? A metaanalytic review of empirical research. *Journal of Applied Psychology*, 83(5), 777–787.
24. Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), 305–360.
25. Latham, G. P., Almost, J., Mann, S., & Moore, C. (2005). New developments in performance management. *Organizational Dynamics*, 34(1), 77–87.
26. Lawler, E. E., III. (2000). *Rewarding excellence: Pay strategies for the new economy*. San Francisco: Jossey-Bass.
27. Mero, N. P., & Motowidlo, S. J. (1995). Effects of rater accountability on the accuracy and the favorability of performance ratings. *Journal of Applied Psychology*, 80(4), 517–524.
28. Meyer, H. H. (1991). A solution to the performance appraisal feedback enigma. *Academy of Management Executive*, 5(1), 68–76.
29. Milkovich, G. T., Newman, J. M., & Gerhart, B. (2011). *Compensation*. New York: McGraw-Hill Irwin.
30. Moers, F. (2005). Discretion and bias in performance evaluation: The impact of diversity and subjectivity. *Accounting, Organizations, & Society*, 30(1), 67–80.
31. Nathan, R. D., Mohrman, A.M., & Milliman, J. (1991). Interpersonal relations as a context for the effects of appraisal interviews on performance and satisfaction: A longitudinal study. *Academy of Management Journal*, 34(2), 352–369.
32. Pfeffer, J. (1998). *The human equation: Building profits by putting people first*. Boston: Harvard Business Press.
33. Pink, D. (2009). *Drive: The surprising truth about what motivates us*. New York: Riverhead Books.
34. Pooyan, A., & Eberhardt, B. J. (1989). Correlates of performance appraisal satisfaction among supervisory and nonsupervisory employees. *Journal of Business Research*, 19(3), 215–226.
35. Prendergast, C., & Topel, R. (1993). Discretion and bias in performance evaluation. *European Economic Review*, 37(2), 355–365.
36. Robinson, S. L., & Rousseau, D. M. (1994). Violating the psychological contract: Not the exception but the norm. *Journal of Organizational Behavior*, 15(3), 245–259.
37. Rosenbaum, A. E., Lehman, W. E. K., & Holcom, M. L. (1993). Effects of personal accountability and rating consequences on performance evaluation. Paper presented at the Eighth Annual Conference of the Society for Industrial and Organizational Psychology, San Francisco.
38. Saadat, E. (1996). *Human resource management* (10th ed.). Tehran: SAMT Publications.
39. Schermerhorn, J. R., Jr., Hunt, J. G., & Osborn, R. N. (2008). *Organizational behavior*. Hoboken, NJ: John Wiley.
40. Schminke, M., Cropanzano, R., & Rupp, D. E. (2002). Organization structure and fairness perceptions: The moderating effects of organizational level. *Organizational Behavior & Human Decision Processes*, 89(1), 881–905.
41. Shepherd, C. D., Carley, S. S., & Stuart, R. S. (2009). An exploratory investigation of the periodic performance evaluation processes for marketing faculty. *Journal of Marketing Education*, 31(2), 143–153.
42. Spence, J. R., & Keeping, L. M. (2013). The road to performance ratings is paved with intentions: A framework for understanding managers' intentions

- when rating employee performance. *Organizational Psychology Review*, 3(4), 360–383.
43. Srull, T. K., & Wyer, R. S., Jr. (1980). Category accessibility and social perception: Some implications for the study of person memory and interpersonal judgments. *Journal of Personality & Social Psychology*, 38(6), 841–856.
 44. Tsui, A. S., & Barry, B. (1986). Interpersonal affect and rating errors. *Academy of Management Journal*, 29(3), 586–599.
 45. Vance, R. J., Winne, P. S., & Wright, S. E. (1983). A longitudinal examination of rater and ratee effects in performance ratings. *Personnel Psychology*, 36, 609–621.
 46. Williamson, O. E. (1981). The economics of organization: The transaction cost approach. *American Journal of Sociology*, 87(3), 548–577.
 47. Wright, P. M., Dunford, B. B., & Snell, S. A. (2001). Human resources and the resource-based view of the firm. *Journal of Management*, 27(6), 701–721