"Manifestation of idiosyncratic rater effect in employee performance appraisal"

AUTHORS	Athmeeya Hunuganahalli Paramesh Vishal Samartha Rajesha Thekkekutt Mathukutti ⓘ https://orcid.org/0000-0001-6214-0509 Ⓡ https://publons.com/researcher/1710646/rajesha-t-m/ Iqbal Thonse Hawaldar ⓘ https://orcid.org/0000-0001-7181-2493 Ⓡ https://publons.com/researcher/1456475/iqbal-thonse-hawaldar/

NUMBER OF REFERENCES	NUMBER OF FIGURES	NUMBER OF TABLES
35	0	5

BUSINESS PERSPECTIVES

Athmeeya Hunuganahalli Paramesh,
MBA, Sahyadri College of Engineering
& Management, Mangalore, India.

Vishal Samartha, Ph.D., Professor,
Director of MBA, Sahyadri College
of Engineering & Management,
Mangalore, India.

Rajesha Thekkekutt Mathukutti, MBA,
Institutional Assessment Administrator,
Accreditation and Quality Assurance
Centre, Kingdom University, Bahrain.

Iqbal Thonse Hawaldar, Ph.D.,
Professor, College of Business
Administration, Kingdom University,
Bahrain. (Corresponding author)

Athmeeya Hunuganahalli Paramesh (India), Vishal Samartha (India),
Rajesha Thekkekutt Mathukutti (Bahrain), Iqbal Thonse Hawaldar (Bahrain)

MANIFESTATION OF IDIOSYNCRATIC RATER EFFECT IN EMPLOYEE PERFORMANCE APPRAISAL

Abstract

Performance appraisal is the bedrock of talent management and has received much attention from scholars and researchers alike in their pursuit to develop accurate, objective, and robust Performance Management Systems (PMS). Through survey questionnaire the present study examines the prevalence of idiosyncratic rater biases on the performance appraisal systems and evaluates the measure of its impact. The correlations between the personality traits and the similarities of the raters' workplace characteristics with the raters' performance ratings are also determined. The study has provided empirical evidence of the manifestation of idiosyncratic rater bias in the company under study. The idiosyncratic rater tendencies showed a significant impact on performance ratings. It was seen that about one-third of the variations in the ratings were resultant of the idiosyncratic factors, such as similarities in the personality traits and workplace identities. It is also found that there exists a positive correlation between the similarities in the identities, as well as the personality traits of the raters and the ratees, and the way the rating awarded by the rater.

Keywords　　human capital, performance assessment, idiosyncratic bias, rater judgments, job performance, performance appraisal

JEL Classification　　E24, M10, M12, M19

INTRODUCTION

Employees are the "greatest asset" of any organization. In the knowledge-based economy of the 21st century, the human capital's quality is the determinant of the long-term viability and competitive advantage of any organization. In the modern world, the workforce's effectiveness governs the competency of any firm, indicating that employees are the greatest asset for any organization (Samartha, Rajesha, Hawaldar, & Souzal, 2019). Performance evaluations have significant importance in every organization, predominantly to guide managerial decisions that link employee performance with rewards and punishments, increments, promotions, or dismissals (Rynes, Gerhart, & Parks, 2005). Traditional performance appraisals measure employee performance in a specific period, which may not work well in the current business world, where skilled and talented employees are at a premium (Javad & Sumod, 2016). To be effective, the rating must reflect accurate job performance, i.e., it requires a high degree of correlation between the ratings and the accurate levels of performance (Austin & Villanova, 1992). However, the general inference drawn from the existing literature is that the actual work performance has a constructive but less than optimum consequence on ratings (Scullen, Mount, & Goff, 2000). The source of this variance labeled "idiosyncratic rater effects" (Hoffman, Lance, Bynum, & Gentry, 2010) are not random measurement errors but are known to be associated with the rater one

way or the other. The present study is conducted to identify, evaluate, and measure the prevalence of idiosyncratic rater effect in performance ratings of the employees by undertaking and examining the influence of the similarities in the raters' personality traits and the ratees on idiosyncratic rater effect. The correlation between workplace characteristics exhibited by the raters and the ratees and the performance ratings are also determined. The emerging interdisciplinary studies of business management and applied psychology have suggested that performance evaluations are saturated with rater-associated idiosyncratic variance, thereby reducing the performance evaluations (PE) from being the representations of the employee's actual performance. The discussion on the components of performance ratings as identified by various theories on ratings and the components of variances influencing the ratings of idiosyncratic rater effect is a large part and is presented in the subsequent paragraphs.

1. LITERATURE REVIEW

1.1. Factors influencing performance ratings

In their theory, Wherry and Bartlett (1982) indicate that three broad kinds of aspects impact performance evaluation: the ratee's actual work performance, rater's perception of performance, and evaluation error. The ratee's actual job performance was hypothesized as the accurate measure of ratee's performance, and the various rater biases and measurement error were the sources of variances in job ratings. These variances in job ratings were found to have two distinct characteristics wherein the rater biases were found to be specific to the rater under observation and the measurement error being random. Scullen, Mount, and Goff (2000) expanded and conceptualized the five factors that influence performance ratings as follows:

a) ratee's overall performance;
b) ratee's performance on a specific performance dimension;
c) rater's idiosyncratic evaluating tendency;
d) rater's organizational perspective (i.e., self, peer, supervisor, etc.); and
e) random evaluation error.

Thus, idiosyncratic rater bias or rating tendencies were distinguished separately from other variances, and efforts were made to understand the influence of the same on the performance appraisals. Contrary to the popular perceptions that the performance ratings are reflections of rater's actual job performance, the research suggests that there exists a moderate correlation amongst objective and rating measures such as quality and quantity, which ranges between 0.10 and 0.40 (Bommer, Johnson, Rich, Podsakoff, & Mack, 1995). Scullen, Mount, and Goff (2000) studied the factor influencing the performance rating and found that the idiosyncratic rater effect accounted for more than 50 percent of the rating variance. The combined effects of general and dimensional ratee's performance (21% and 25%) were less than half the size of the idiosyncratic rater effects. The influence of random measurement error on the variance in performance ratings was studied by Viswesvaran, Ones, and Schmidt (1996). It was found that managers commit approximately 19% of the performance rating variances due to transient error, random measurement error, and other unidentified aspects. These studies brought into focus the substantial influence of the idiosyncratic rater effects on the variances in the performance ratings.

1.2. Idiosyncratic rater effect and its components

Probing further on the components of the large variance noticed in the performance ratings, the researchers agree that the rater perspective and performance dimensions are largely influenced by rater's bias (Castilla & Benard, 2010). Mount, Judge, Scullen, Sytsma, and Hezlett (1998) found that 72 percent of the job performance rating variance is due to IRE. According to Hoffman, Lance, Bynum, and Gentry (2010), idiosyncratic rater effects are too large to be neglected. The systematic variance in job performance rating is different across the raters. To understand the phenomenon of idiosyncratic rater effect, Buckingham (2015) cites the example of a manager being required to rate his colleague on a quality such as "potential". He states that the manager's idiosyncrasies, such as how the manager defines "potential," how much of it he

thinks he has and how tough a rater he usually is likely to influence the performance ratings rather than the actual performance of the ratee. All the above studies concluded that even though the performance evaluation measures the employee, most of the time, the rating reveals more about the rater rather than the ratee (Buckingham, 2015).

Other findings of performance rating research (Ullal, Hawaldar, & Samartha, 2019; Samartha, Rajesha, Hawaldar, & Souza, 2019; Cleveland & Murphy, 1992) also suggest that performance appraisals are far from an objective exercise as raters are found to pursue different goals such as motivating subordinates, maintaining good interpersonal relationship among the peers when completing the performance appraisal while evaluation of their subordinates is relatively a minor concern for the raters. The other well-documented errors/effects, which are known to impact the variations in performance appraisals such as halo error, proximity error, leniency and severity errors, contrast error, central tendency, past-record anchoring, recency error, rater attitudes, personal bias and values discrimination between inside and outside employees, employee appearance, etc. (Javidmehr & Ebrahimpour, 2015), are also reasons that contribute towards the idiosyncratic bias of the raters on the performance appraisal of the candidates. Scullen, Mount, and Judge (2003) described the manifestations of rater effect by way of biases such as halo and leniency that have received widespread attention of the researchers in recent times compared to other cognitive biases. Halo error refers to the tendency of raters to allow an overall impression of a ratee to influence judgments along several quasi-independent dimensions (King, Hunter, & Schmidt, 1980), and leniency error is the rater's tendency to assign ratings that are generally either higher or lower than are acceptable by the ratees' actual performance. Hoffman, Lance, Bynum, and Gentry (2010) described idiosyncratic rater bias as a systematic effect that is common only to an individual rater but refrained from offering specifics. O'Neill, McLarnon, and Carswell (2015) identified the variance components in performance ratings based on generalizability theory (Cronbach et al., 1972), broadly classified the performance ratings' variances two discrete categories, i.e., idiosyncratic rater components and interrater-reliable components. The idiosyncratic rater components were further classified into four-folded rater-related variances identified by O'Neill, McLarnon, and Carswell (2015):

1) variances due to rating differences concerning individual rater irrespective of the ratee or the dimension;
2) rating differences on a particular dimension, regardless of the ratee;
3) rating differences on a particular ratee, regardless of the dimension; and
4) unexplained variance.

Despite having large literature on the idiosyncratic rater effect, the business world appears to be completely unaware of it. More so in the context of Indian businesses and industry as estimates suggest that more than one-third of US companies are modifying their performance appraisal systems. Technology giants such as Dell, Microsoft, Adobe, IBM, and Juniper Systems have led the way (Javad & Sumod, 2016). Despite consistent reports that idiosyncratic rater effect as the main variance component of work performance ratings (Ullal, Hawaldar, & Samartha, 2019; Hawaldar et al. 2016; Conway, 1996), this variance's distinctions are not well explored. Scullen, Mount, and Goff (2000), while undertaking a comprehensive study on the latent structure of ratings and acknowledging the IRE to be the major source of variances, recommend investigations into the nature and causes of the idiosyncratic variations in the performance ratings. The role of cognitive biases such as halo effect, leniency error in influencing performance ratings is well documented. However, there is a need to look beyond the cognitive biases and look at the raters' psychological makeup influencing idiosyncratic rater effects. The studies in organizational psychology have correlated "Big Five" personality traits with job performance criterion (Barrick & Mount, 1991) and have established a positive correlation with one of the dimensions. This study goes a step further and attempts to connect the big five personality studies with the performance ratings. It also tries to investigate the connection between the similarities and dissimilarities in one's approach to work, referred to as workplace characteristics, and how it influences how the person evaluates others based on this worldview.

2. RESEARCH METHODOLOGY

The study is partly descriptive and partly diagnostic. The primary data are collected through a survey method using a well-structured questionnaire. The first part of the questionnaire was designed to capture the individuals' workplace characteristics, with each question presenting two contradicting responses to a given scenario. The Cronbach alpha value at 0.77 indicates that the internal consistency of the data measured by the first part of the questionnaire could be characterized as "fairly high", i.e., between 0.76 and 0.95 (Thai et al. 2016). The survey questions on workplace characteristics are extracted from the previous studies by Doornbos, Simons, and Denessen (2008), Ben-Ner, McCall, Stephane, and Wang (2009), which are designed to capture personal (Q4, Q5, and Q7), relational (Q6, Q8, and Q9), and work (Q1, Q2, Q3, and Q10) characteristics of the workplace.

The second part of the questionnaire was designed to measure the personality trait of the individual participants. The "Big Five" personality test or more famously known as the five-factor model test, which has emerged to be the most acknowledged test to measure personality traits by psychologists worldwide, was used (Antonioni & Park, 2001). Four questions on the five personality traits such as neuroticism, extraversion, conscientiousness, agreeableness, and openness to experience were asked to the participants. The questionnaire was designed in a structured format, including open-ended, closed-ended, rating preferences, and 5 points 'Likert scale' questions. The Cronbach alpha value at 0.87 indicates that the internal consistency of the data measured by the second part of the questionnaire could be characterized as "reliable", i.e., between 0.84 and 0.90 (Thai et al. 2016). Secondary data were collected through the previous year's performance rating from the managers/team leads reported upon. The company surveyed is a mid-size IT firm with around 130 employees; 66 percent are male, and 34 percent are female respondents. 27% of the respondents are senior-level, 35% of the respondents are middle level, and 38% are entry-level employees. The next step in the process was to measure the idiosyncratic rater effect in the performance appraisal systems by comparing the primary and the secondary data. The entire data was segregated into different teams of the previous year. Of the 11 managers in the previous year, one had moved out; thus, the data were arranged into ten sets, with all the managers and their erstwhile teams segregated. The study intended to observe the manifestations of idiosyncratic rater effect in each of the ten managers/team leads who carried out the performance appraisal to their reporters. The primary data regarding workplace characteristics and the "Big Five" personality traits yielded the manager/team leads' personality map. The data were used to determine the matching between the manager/team-leads with the reporters to find out the similarities in the style of functioning and the behavioral traits. In order to verify the existence idiosyncratic biases similarities or dissimilarities between the rater and the ratee based on the information gathered the following hypothesis is tested to find out if any discernible pattern in the ratings could be seen, which shows the manifestations of the rater effect.

H_0: The ratings reveal more about the rater than they do about the ratee.

H_1: There is no significant influence of rater's personality on performance rating.

Multiple options method was used to study the workplace characteristic, and 5-point Likert scale was used to study the "Big Five" personality traits with 1 = Strongly Disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly Agree. The researcher has analyzed and interpreted the data using techniques like the Chi-square test and regression analysis.

3. RESULTS

3.1. Capturing the idiosyncratic biases of managers/team leads

The theoretical basis for understanding rater-variances is drawn from two streams of research on the topic. Ben-Ner, McCall, Stephane, and Wang (2009) studied the relationship between identity and behavior of individuals and found significant bias in favor of persons of same/similar (Self) identity over the persons of different identity (Other). The first part of the primary data collected by the

method of a survey was designed in line with the above findings were classified as measuring the workplace identities under the broader umbrella of "workplace characteristics" of both the raters and ratees as manifested in the day-to-day work environment. The factors that were considered for this purpose are as below.

Table 1. Workplace Characteristics Factors

Factor 1	Following established procedure vs. experimentation
Factor 2	Multi-tasking vs. singular focus
Factor 3	Perfection with some delay vs. fast output
Factor 4	Being organized desirable or over-rated
Factor 5	Self-discipline a must vs. deliverable matter the most
Factor 6	Assertiveness vs. listening to others
Factor 7	Getting things done vs. getting the basics right
Factor 8	Soft skills vs. technical competence
Factor 9	Description of character
Factor 10	People by nature are lazy vs. people by nature are work-loving

Secondly, based on reviews on the association between personality and organizational behavior (Barrick, Mount, & Judge, 2001; Shaffer & Postlethwaite, 2013; Hawaldar et al. 2016), the authors adopted the "Big Five" model for a greater understanding of the personality bedrock of ratings (Ceschi, Costantini, Scalco, Charkhabi, & Sartori, 2016). This study's key premise is that matching certain personality dimensions of the "Big Five" are likely to influence idiosyncratic rater-related variances in the ratings one assigns to others. The second part of the primary data was aimed at measuring the individuals' personality traits by using the "Big Five" personality test. The personality trait of the individual rater and the ratee were constructed using the Likert scale on the five factors listed in Table 2.

Table 2. Big Five Personality Traits

Factor 1	Neuroticism	4 variables measuring the factor
Factor 2	Extraversion	4 variables measuring the factor
Factor 3	Conscientiousness	4 variables measuring the factor
Factor 4	Agreeableness	4 variables measuring the factor
Factor 5	Openness to experience	4 variables measuring the factor

The primary data regarding the workplace characteristics and the big five personality traits yielded the manager/team leads' personality map. The data were used to determine the matching between the manager/team leads with the reporters to find out the similarities in the style of functioning and the behavioral traits. The hypothesis – "the ratings reveal more about the rater than they do about the ratee" – was attempted to be tested by first learning about the similarities or dissimilarities between the rater and the ratee based on the information gathered by the primary data and then by comparing the ratings of the employees of the previous year to find out if any discernible pattern in the ratings could be seen, which shows the manifestations of the rater effect.

Table 3. Regression analysis of idiosyncratic rater effect team-wise

Source: Authors' calculations.

Teams	"Big Five" matching and previous year rating		Workplace characteristic and previous year rating	
	R-value	R^2 value	R-value	R^2 value
Team A	0.8713	0.7591	0.7651	0.5854
Team B	0.5524	0.3051	0.8086	0.6538
Team C	0.7828	0.6127	0.8350	0.6972
Team D	0.8234	0.6780	0.7527	0.5665
Team E	0.0964	0.0093	−0.4436	0.1968
Team F	0.7039	0.4955	0.8854	0.7840
Team G	0.5844	0.3415	0.4875	0.2377
Team H	0.2054	0.0422	0.0699	0.0049
Team I	0.7079	0.5011	0.4975	0.2475
Team J	0.6485	0.4205	0.4288	0.1839
Average R^2 value		0.4165	−	0.3764

Here the matching in the personality traits and the workplace characteristics formed the independent variables individually, and the previous year ratings were taken as the dependent variable. The similarities in the employee and his manager's personality traits showed a dependence of 41.65% of how the employee's performance was rated. Also, the matchings in the workplace showed a dependence of 37.64% in ratings. The R^2 values observed are quite significant since the research tries to analyze human behavior. It is typically seen that in fields such as psychology, social sciences, and humanities, the R^2 values of .12 or below indicate low, between .13 and .25 values indicate medium, .26 or above and above values indicate high effect size. (Cohen, 1992). The regression analysis of the "Big Five" personality trait matching showed a high correlation in teams A, C, D, and I ($R^2 > 0.50$) and moderate correlation in teams B, F, G,

and J (R^2 between 0.20 and 0.50 and low correlation in teams E and H ($R^2 < 0.20$). The regression analysis of the workplace characteristics matching showed a high correlation in teams A, B, C, D, and F ($R^2 > 0.50$) and moderate correlation in teams G and I (R^2 between 0.20 and 0.50). However, this also showed no correlation in team H ($R = 0.06$) and a negative correlation in team E ($R = -0.44$). The regression analysis thereby shows that the raters (managers/team-leads) in 8 out of 10 teams have shown a tendency to rate the employees who match their style of the functioning or the ones who have similar personality traits high compared to others, thereby confirming the existence of idiosyncratic rater effect in the performance appraisal system of the organization under study. Thus, it is concluded that there is no significant impact of the similarities in the personality traits and the workplace characteristics between rater and ratee on the performance appraisals.

The *p*-value was obtained using CHITEST function on observed and expected values in MS Excel.

$$p = 0.000137901.$$

Chi-square test was used to test the hypothesis that "there is no significant impact of the similarities in the personality traits and the workplace characteristics between rater and ratee on the performance appraisals". Since the *p*-value is less than

0.05 conclude that there is a significant impact of similarities in the personality traits and the workplace characteristics between rater and ratee on the performance appraisals.

4. DISCUSSION

The research in psychometrics and organizational behavior points to various means to reduce the effect of idiosyncratic rater bias. The replacement of annual performance appraisals with continuous evaluation methods, which are undertaken around the year, has shown to enhance performance evaluation quality. Introducing more flexibility in performance appraisals by reducing the reliance on the annual appraisal and making the appraisals them project-specific, informal, and multi-dimensional is hereby suggested (Javidmehr & Ebrahimpour, 2015). Reductions in rater-related variances were detected when evaluations are used for developmental purposes rather than managerial decisions (Greguras & Robie, 1998). Further, greater lengths of acquaintanceship time are shown to improve the quality of performance ratings. This is because the circumstances required providing a precise rating of other individuals, adequate observational opportunities, and proper weighting of performance-related cues will intensify with augmented acquaintanceship time (Hammond, 1955). With a small acquaintance-

Table 4. "Observed frequencies" of matching vis-à-vis performance ratings

Observed values	Total matchings ("Big Five" and characteristics)									
Adjusted performance ratings	4	5	6	7	8	9	10	11	12	Grand total
6	5	4	4	3	2		2	2		22
7	1	7	6	5	8	3	3	1		34
8		1	2	5	3	7	5	7		30
9							2	5	2	9
Grand total	6	12	12	13	13	10	12	15	2	95

Table 5. "Expected frequencies" of matching vis-à-vis performance ratings

Source: Field survey.

Expected values	Total matchings ("Big Five" and characteristics)									
Adjusted performance ratings	4	5	6	7	8	9	10	11	12	Grand total
6	1.389	2.779	2.779	3.011	3.011	2.316	2.779	3.474	0.463	22
7	2.147	4.295	4.295	4.653	4.653	3.579	4.295	5.368	0.716	34
8	1.895	3.789	3.789	4.105	4.105	3.158	3.789	4.737	0.632	30
9	0.568	1.137	1.137	1.232	1.232	0.947	1.137	1.421	0.189	9
Grand total	6	12	12	13	13	10	12	15	2	95
Proportion	0.0632	0.1263	0.1263	0.1368	0.1368	0.1053	0.1263	0.1579	0.0211	

ship period, the rater may tend to defer general evaluation heuristics such as leniency or central tendency, which would be manifested as idiosyncratic rater variance (Murphy & Deshon, 2000). Defining measurable work goals and metrics is another area that must be studied to condense the idiosyncratic rater effect. This would give a more accurate picture of the employee's past achievements in terms of tangibles and the intangibles regarding the value judgments, and the team leaders could ascertain the employee's prospect by asking them to respond to four future-focused statements about each employee (Buckingham, 2015):

1. Would I award the highest increase in bonus or the highest amount of compensation to the individual if it were my money? [to measure the performance of the individual during the relevant time of review – 5-point scale answer]

2. Would I always prefer the individual on my team? [to measure if the person is a good team player and if he works well with others – 5-point scale answer]

3. This individual is at risk for weak performance [measures the threat of low performance to the team and the customer – yes-or-no answer].

4. This individual is ready for promotion today [measures potential on a yes-or-no basis].

Lastly, the Number of Ratees Rated (NRR) is associated with variance component magnitudes (O'Neill, Goffin, & Gellatly, 2012). Information processing theory suggests that many ratees may be problematic because of raters' definite cognitive resources and fatigue (Lord & Maher, 1990).

CONCLUSION

This study's major contribution is to augment the understanding of idiosyncratic rater biases, which are known to plague our present-day performance management systems. The study has provided empirical evidence of the manifestation of idiosyncratic rater bias in the company under study. The idiosyncratic rater tendencies showed a significant impact on performance ratings. It was seen that about one-third of the variations in the ratings were resultant of the idiosyncratic factors, such as similarities in the personality traits and workplace identities. This observation is both qualitatively and quantitatively in line with the findings of the studies conducted by O'Neill, McLarnon, and Carswell (2015) and Scullen, Mount, and Goff (2000). Further, the data also fortify the second objective of the study that there exists a positive correlation between the similarities in the identities, as well as the personality traits of the raters and the rates, and the way the rating awarded by the rater, which is an extension of the studies conducted by Ben-Ner, McCall, Stephane, and Wang (2009) and Barrick and Mount (1991). The research hypothesizes the linkages between the psychological reasons between the idiosyncratic rater effects such as "behavioral traits" and "identities" that have been unchartered territory in this research field. However, there is a need to integrate the same with the broader framework of the studies conducted in the field of cognitive biases (Javidmehr & Ebrahimpour, 2015) and mathematical models based on G-theory (Cronbach, Gleser, Nanda, & Rajaratnam, 1972). The study mainly focuses on measuring the idiosyncratic biases impacting the performance appraisal. The research calls for an increase in awareness of the perforation of such biases on appraisal systems, especially in the persons who are the appraisers in an organization. The study only gives indicative steps for reducing the biases and does not put forth a concrete action plan for removing the same. The research conducted on the topic has concluded that it is impossible to design a performance appraisal system free of biases. The emerging research and the pioneering methods adopted by the industry leaders call for an overhauling of the traditional approach to performance appraisal beyond this research's scope of the study. In light of the findings, there is a need to renew research to find ways to reduce idiosyncratic rater effects and understand the same in the socio-cultural context of Indian businesses.

AUTHOR CONTRIBUTIONS

Conceptualization: Athmeeya Hunuganahalli Paramesh, Vishal Samartha, Iqbal Thonse Hawaldar.
Data curation: Athmeeya Hunuganahalli Paramesh, Rajesha Thekkekutt Mathukutti.
Formal analysis: Vishal Samartha, Iqbal Thonse Hawaldar.
Investigation: Athmeeya Hunuganahalli Paramesh, Vishal Samartha, Rajesha Thekkekutt Mathukutti.
Methodology: Athmeeya Hunuganahalli Paramesh, Rajesha Thekkekutt Mathukutti, Iqbal Thonse Hawaldar.
Project administration: Vishal Samartha, Iqbal Thonse Hawaldar.
Supervision: Vishal Samartha.
Validation: Rajesha Thekkekutt Mathukutti, Iqbal Thonse Hawaldar.
Visualization: Rajesha Thekkekutt Mathukutti, Iqbal Thonse Hawaldar.
Writing – original draft: Athmeeya Hunuganahalli Paramesh, Vishal Samartha, Rajesha Thekkekutt Mathukutti.
Writing – review & editing: Iqbal Thonse Hawaldar.

REFERENCES

1. Antonioni, D., & Park, H. (2001). The relationship between rater affect and three sources of 360-degree feedback ratings. *Journal of Management*, 27(4), 479-495. https://doi.org/10.1177/014920630102700405

2. Ben-Ner, A., McCall, B. P., Stephane, M., & Wang, H. (2009). Identity and in-group/out-group differentiation in work and giving behaviors: Experimental evidence. *Journal of Economic Behavior and Organization, 71*(1), 153-170. https://doi.org/10.1016/j.jebo.2009.05.007

3. Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personal Psychology, 44*(1), 1-26. https://doi.org/10.1111/j.1744-6570.1991.tb00688.x

4. Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: what do we know and where do we go next? *International Journal Selection Assessment, 9*(1-2), 9-30. https://doi.org/10.1111/1468-2389.00160

5. Bommer, W. H., Johnson, J. L., Rich, G. A., Podsakoff, P. M., & Mackenzie, S. B. (1995). On the interchangeability of objective and subjective measures of employee performance: A meta-analysis. *Personnel Psychology, 48*(3), 587-605. https://doi.org/10.1111/j.1744-6570.1995.tb01772.x

6. Buckingham, M. (2015). Most HR data is bad data. *Harvard Business Review*. Retrieved from https://hbr.org/2015/02/most-hr-data-is-bad-data

7. Castilla, E. J., & Benard, S. (2010). The paradox of meritocracy in organizations. *Administrative Science Quarterly, 55*(4), 543-676. https://doi.org/10.2189/asqu.2010.55.4.543

8. Ceschi, A. C., Costantini, A., Scalco, A., Charkhabi, M., & Sartori, R. (2016). The relationship between the Big Five personality traits and job performance in business workers and employees' perception. *International Journal of Business Research, 16*(2), 63-76. http://dx.doi.org/10.18374/IJBR-16-2.5

9. Cleveland, J., & Murphy, K. (1992). Analyzing performance appraisal as goal-directed behavior. *Research in Personnel and Human Resources Management, 10,* 121-185.

10. Conway, J. M. (1996). Analysis and design of multitrait-multirater performance appraisal studies. *Journal of Management, 22*(1), 139-162. https://doi.org/10.1177/014920639602200106

11. Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles.* New York: Wiley.

12. Doornbos, A. J., Simons, R.-J., & Denessen, E. (2008). Relations between characteristics of workplace practices and types of informal work-related learning: A survey study among Dutch Police. *Human Resource Development Quarterly, 19*(2), 129-151. https://doi.org/10.1002/hrdq.1231

13. Greguras, G. J., & Robie, C. (1998). A new look at within-source interrater reliability of 360-degree feedback ratings. *Journal of Applied Psychology, 83*(6), 960-968. https://doi.org/10.1037/0021-9010.83.6.960

14. Hammond, K. R. (1955). Probabilistic functioning and the clinical method. *Psychological Review, 62*(4), 255-262. https://doi.org/10.1037/h0046845

15. Hawaldar, I.T., Lokesha, Biso, S.SA., & Joseph, N.R. (2016). Factors affecting leaders' behaviour: A study of Bahrain banking sector. *British Journal of Economics, Finance and Management Sciences, 12(1),* 11-21, http://www.ajournal.co.uk/EFpdfs/EFvolume12(1)/EFVol.12%20(1)%20Article%202.pdf

16. Hoffman, B., Lance, C. E., Bynum, B., & Gentry, W. A. (2010). Rater Source Effects are alive and well after all. *Personnel Psychology, 63*(1), 119-151. https://doi.org/10.1111/j.1744-6570.2009.01164.x

17. Javad, S., & Sumod, S.D. (2016). It's time to bring performance appraisal into the twenty-first century. *Human Resource Management International Digest, 23*(7), 23-26. Retrieved from https://www.deepdyve.com/lp/emerald-publishing/it-s-time-to-bring-performance-appraisal-into-the-twenty-first-century-hn2LdlA4YI

18. Javidmehr, M., & Ebrahimpour, M. (2015). Performance appraisal bias and errors: The influences and consequences. *International Journal of Organizational Leadership, 4,* 286-302. Retrieved from https://ijol.cikd.ca/article_60464_e664300a56883a0e-6b8a49a754f0bd81.pdf

19. King, L. M., Hunter, J. E., & Schmidt, F. L. (1980). Halo in a multidimensional forced-choice performance evaluation scale. *Journal of Applied Psychology, 65*(5), 507-516. https://doi.org/10.1037/0021-9010.65.5.507

20. Lance, C. E. (1994). Test of a latent structure of performance ratings derived from Wherry's (1952) theory of ratings. *Journal of Management, 20*(4), 757-771. https://doi.org/10.1016/0149-2063(94)90029-9

21. Lord, R. G., & Maher, K. J. (1990). Alternative information processing models and their implications for theory, research, and practice. *Academy of Management Review, 15*(1), 9-28. https://doi.org/10.5465/amr.1990.4308219

22. Mount, M. K., Judge, T. A., Scullen, S. E., Sytsma, M. R., & Hezlett, S. A. (1998). Trait, rater, and level effects in 360-degree performance ratings. *Personnel Psychology, 51*(3), 557-576. https://doi.org/10.1111/j.1744-6570.1998.tb00251.x

23. Murphy, K. R., & Deshon, R. (2000). Interrater correlations do not estimate the reliability of job performance ratings. *Personnel Psychology, 53*(4), 873-900. https://doi.org/10.1111/j.1744-6570.2000.tb02421.x

24. O'Neill, T. A., Goffin, R. D., & Gellatly, I. R. (2012). The use of random coefficient modeling for understanding and predicting job performance ratings: An application with field data. *Organizational Research Methods, 15*(3), 436-462. https://doi.org/10.1177/1094428112438699

25. O'Neill, T. A., McLarnon, M. J., & Carswell, J. J. (2015). Variance Components of Job Performance Ratings. *Human Performance, 28*(1), 66-91. https://doi.org/10.1080/08959285.2014.974756

26. Rynes, S., Gerhart, B., & Parks, L. (2005). Personnel Psychology: Performance Evaluation and Pay. *Annual Review of Psychology, 56*(1), 571-600. https://doi.org/10.1146/annurev.psych.56.091103.070254

27. Saadat, E. (1996). *Human resource management* (10th ed.). Tehran: SAMT Publications.

28. Samartha, V., Rajesha, T. M., Hawaldar, I. T., & Souza, L. J. D. (2019). Application of Lev and Schwartz Compensation Model On the Accounting Practices of MCF Limited. *International Journal of Scientific & Technology Research, 8*(8), 356-360. Retrieved from https://www.ijstr.org/final-print/aug2019/Application-Of-Lev-And-Schwartz-Compensation-Model-On-The-Accounting-Practices-Of-Mcf-Limited.pdf

29. Scullen, S. E., Mount, M. K., & Goff, M. (2000). Understanding the latent structure of performance ratings. *Journal of Applied Psychology, 85*(6), 956-970. https://doi.org/10.1037/0021-9010.85.6.956

30. Scullen, S. E., Mount, M. K., & Judge, T. A. (2003). Evidence of the construct validity of developmental ratings of managerial performance. *Journal of Applied Psychology, 88*(1), 50-66. https://doi.org/10.1037/0021-9010.88.1.50

31. Shaffer, J. A., & Postlethwaite, B. E. (2013). The validity of conscientiousness for predicting job performance: a meta-analytic test of two hypotheses. *International Journal of Selection Assessment, 21*(2), 183-199. https://doi.org/10.1111/ijsa.12028

32. Thai, N., Taber-Thomas, B. C., & Pérez-Edgar, K. E. (2016). Neural correlates of attention biases, behavioral inhibition, and social anxiety in children: An ERP study. *Developmental cognitive neuroscience, 19,* 200-210. https://doi.org/10.1016/j.dcn.2016.03.008

33. Ullal, M. S., Hawaldar, I. T., & Samartha, V. (2019). Employee relations in micro, small and medium enterprises in India. *International Journal of Scientific & Technology Research, 8*(8), 117-121. Retrieved from https://www.ijstr.org/final-print/aug2019/Employee-Relations-In-Micro-Small-And-Medium-Enterprises-In-India.pdf

34. Viswesvaran, C. O., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology, 81*(5), 557-574. https://doi.org/10.1037/0021-9010.81.5.557

35. Wherry, R. J., & Bartlett, C. J. (1982). The control of bias in ratings: a theory of rating. *Personnel Psychology, 35*(3), 521-551. https://doi.org/10.1111/j.1744-6570.1982.tb02208.x