

EMPLOYEES APPRAISAL: A FIELD RESEARCH OF VARIABLE AND CONSTANT ERRORS ON PERFORMANCE RATING

Dr. John Bouris

Professor

Technological Educational Institute (TEI) of Athens

School of business and economics

Department of Business Administration

e-mail: jbouris@teiath.gr

ABSTRACT

The purpose of this study was to explore, in a public interest Financial Institution, the relationships among several possible sources of errors and bias as they are investigated in the performance appraisal process.

The errors and bias are categorized into two main factors: a) the constant errors or latent tendency toward a 'groupthinking' appraisal b) the variable errors dealing with either discrepancies among raters or longitudinal disagreements on the same subject.

The significance of the constant errors (i.e. halo effects, leniency or strictness bias), on the distortion the performance appraisals drive us to formulate and investigate the following hypotheses:

1. the leniency or strictness evaluation is highly depended on the job the employee assigned
2. the line personnel is evaluated more strictness than the staff personnel
3. there is an heterogeneous performance evaluation procedure between staff and line personnel

Accepting the alternative hypothesis, a group of 'underdeveloped' relationships is raised among reliability, validity, acceptability, dissatisfaction, objectivity fairness and consistency, paternalism, manipulations and non-transparency of human resource policies.

The importance of these relationships arises from the facts that performance ratings are often used (a) to aid in the determination of promotions, assign human resource to organization positions, merit increases, lateral transfers, and (b) to guide the organization in its estimate of training, educational and developmental needs and changes in hiring and career pathing policy.

The sample in this study were N=300 employees whose performance evaluated for a year period 2010 under 'Criteria-Based Performance Reviews'. The employees (ratees) were equivalently splitted in staff and line personnel derived from a public interest financial institution which organizational structure consist of two tiers : (a) the central or strategic (steering committee) and administrative schema (b) the peripheral or operational one.

The performance appraisal form contained eight criteria (ie. quality of work, quantity of work, initiative, reliability, interpersonal relationship, dependability, knowhow, and administrative competency).

The over all performance rated on eighth point scale ratings from 1=unsatisfactory.....8=exceptional.

For data processing, transforming data collected into a form appropriate to analysis and hypothesis testing, we use the software: S.P.S.S. statistics v. 17 release 17.0.0.

In summary, Human Resource Decisions based on Performance Appraisal Models have significant human consequences and should embrace a moral balance between equity and effectiveness.

Keywords: Multi criteria Decision Making Analysis, Performance Appraisals

1. Theoretical Background

Several problems undermine performance appraisal. In particular our field research has identified rating leniency or strictness as one of the most troublesome of evaluating errors. Specifically, raters have the tendency to rate ratees either high or low [8]. H. John Bernadin [1] have related the personality traits of the raters with performance appraisal rating and concluded that a conscientiousness personality tends to be too strict with respect to agreeable one which is more lenient.

Studies [2] & [3] suggest ways of reducing the problems of leniency or strictness in performance appraisals such as : a) rank employees b) conduct a job analysis to establish criteria and standards for successful performance c) administer and score appraisals in a standardized fashion d) use clearly defined job performance dimensions.

Common appraisal problems mainly arise because appraisal often says more about the appraiser than about the appraisee. [7]. This is a powerful reason for using multiple raters. In other words the performance appraisals may be conducted by: the immediate supervisor and the peers and the rating committees

In our field study, the so far implied knowledge concerning the relationship between leniency or strictness and employees assignments is hypothesized and tested.

To approach our goal, we testing the following hypothesis:

- 1: the leniency or strictness evaluation is highly depended on the job the employee assigned
- 2: the line personnel is evaluated more strictness than the staff personnel
- 3 : there is an heterogeneous performance procedure between staff and line personnel

Accepting the alternative hypothesis, a 'latent' factors launches: the paternalism and non-transparent manipulation on human resource policies .

Despite recent exhortations to scholars in performance appraisal , rating paternalism trends on human resource polices has received little attentions. Therefore, little is known about the extent of bias due to paternalism subject to 'political' manipulation on both the fairness and integrity of performance procedures.

Paternalism indicates a more intricate relationship between the involved parties than an organizational hierarchy requires. In such a relationship, there are two parties first, a patron who protects, helps, cares, and guides; second, a subordinate who is loyal and deferent to the patron. So between rater and ratee there is a reciprocal relationship [6].

Paternalism indicates managers take a personal interest in their subordinates' off -the-job lives and attempt to promote ratees' personal welfare [4]

Studies [5] distinguish between two forms of paternalism: authoritative and benevolent. Authoritative paternalism includes emphasis on duty and lacks sincere generosity on the part of the superior. Benevolent paternalism emphasizes the subordinate's loyalty and the superior's generous concern for that subordinate.

In paternalistic cultures, people in authority assume the role of parents and consider it an obligation to provide protection to others under their care. Subordinates, in turn, reciprocate such care and protection of the paternal authority by showing

loyalty, deference, and compliance. In a paternalistic relationship, the subordinates voluntarily depends on the manager. Paternalism implies voluntary compliance; therefore, subordinates who experience high levels of trust, obligation, and respect in their relations with a manager (i.e., high quality rater-ratee exchange) may be more likely to accept the manager's authority as a father figure.

Researchers [9] & [10] trying to monitoring and eliminating the paternalism and enhance ethics in human resource polices. develop the Performance Appraisal Fairness Scale

The original questionnaire has 56 items. The questions are grouped under the following categories:

- Setting performance expectations (six items, including: "the performance planning and review process makes sure that my performance expectations measure what I really do for the organization" and "the expectations set during performance planning session reflect the most important factors in my job")
- Rater confidence (five items, including: "My organization makes sure that I am assigned a rater who is qualified to evaluate my work" and "My organization ensures that I am assigned a rater who knows what I am supposed to be doing")
- Clarifying expectations (six items, including: "My rater clearly explains to me what he or she expects for my performance" and "My rater explains how I can improve my performance")
- Providing feedback (six items, including: "My rater frequently lets me know how I am doing" and "My rater reviews with me my progress towards my goals")

<https://sites.google.com/site/icqqmeas2015>

- Accuracy of rating (five items, including: “My performance rating is based on how well I do my work” and “My performance rating reflects how much work I do”)
- Explaining rating decisions (five items, including: “My rater gives me clear and real examples to justify his or her rating of my work” and “My rater helps me to understand the process used to evaluate and rate my performance”)
- Seeking appeals (six items, including: “I have ways to appeal a performance rating that I think is biased or inaccurate” and “I know I can get a fair review of my performance rating if I request one”)
- Concern over ratings (seven items, including: “My rater gives me the rating that I earn even when it might upset me” and “The rating I get is a result of my rater applying performance rating standards consistently across employees”)

2. Methods and procedures

Sample: Ratees in this study were 300 employees in a public interest Financial Institution. We use the stratified sampling methods by grouping the population into two homogeneous group: a) the line personnel b) the staff personnel.

In each group we recruited 150 employees having the same demographic and administrative characteristics. So, a high degree of homogeneity between and within the two groups has succeeded.

In other words, to achieve comparability between the two groups we select each case based on the following variables:

a) **age** (1=31-40; 2=41-50; 3=51-60; 4=60+) b) **gender** (1=male; 2=female) c) **position** (1=operational; 2=administrative) d) **years of total employment** e) **tenure** in the organization (1=junior; 2=senior employee) f) tenure in the job

By recruiting each case and matching randomly in each group based on the above variables, we succeed the ceteris paribus-all the others remain stable hypothesis. It means that neither demographic nor system-organization variables affects the human resources polices. So an latent variable has been revealed which is the parenamism.

Rating instrument

The performance appraisal form contained eight criteria (ie. quality of work, quantity of work, initiative, reliability, interpersonal relationship, dependability, knowhow, and administrative competency)

The over all performance rated on eighth point scale ratings from 1=unsatisfactory.....8= exceptional

Procedure

1. Testing the hypothesis concerning the difference between means of performance evaluation

Ho : There is **no** evidence of significant difference in evaluation scores between groups

$$\mu_1 = \mu_2$$

Ha1: There is evidence of significant difference in evaluation scores between groups

$$\mu_1 \neq \mu_2$$

Based on **Table 1**, since $P < 0,025$ the evaluation scores between groups differ significant at the level of significance $\alpha=0,05$ (2-tailed alternative). **We accept the alternative hypothesis**

2. Testing the hypothesis concerning the difference between variances between groups

Ho : There is equality of variances between groups¹

$$\text{Var}(\text{group}_1) = \text{Var}(\text{group}_2)$$

Ha2 : There is evidence of significant inequality of variances between groups

$$\text{Var}(\text{group}_1) \neq \text{Var}(\text{group}_2)$$

Based on **Table 1**, since $P > 0,025$ the evaluation scores between groups differ significant at the level of significance $\alpha=0,05$ **We accept the Null hypothesis**

¹ Levene's test of equality of variances

Table 1: Independent Samples Test

Levene's Test for Equality of Variances			Test for Equality of Means					
							95% Confidence Interval of the Difference	
	F	Sig.	t	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Equal variances assumed	,901	,351	6,252	,000	5,40000	,86373	3,63073	7,16927
Equal variances not assumed			6,252	,000	5,40000	,86373	3,62328	7,17672

Based on the values of Table 2 we observe the following:

1. For the Line personnel

- a. the coefficient of skewness >0 indicates a positively skewed distribution (skewed to the Right)
 - b. the line personnel (lower part of Figure 1) is evaluated stricter than the staff personnel
- The raters tent to rate the line personnel consistently low, problem know as strictness

2. For the Staff personnel

- a. the coefficient of skewness <0 indicates a negatively skewed distribution (skewed to the left)
 - b. the Staff personnel (upper part of Figure 1) is evaluated lenient as compare to Line one.
- The raters tent to rate the Staff personnel consistently high, problem know as leniency

Table 2 : DESCRIPTIVE STATISTIC

		Statistic	Std. Error	
Line Personnel	Mean	11,5333	,69773	
	95% Confidence Interval for Mean	Lower Bound	10,0369	
		Upper Bound	13,0298	
	5% Trimmed Mean	11,5093		
	Median	11,5000		
	Variance	7,302		
	Std. Deviation	2,70229		
	Minimum	6,50		
	Maximum	17,00		
	Range	10,50		
	Interquartile Range	4,00		
	Skewness	,196		
Kurtosis	,115	1,121		
Staff Personnel	Mean	16,9333	,50912	
	95% Confidence Interval for Mean	Lower Bound	15,8414	
		Upper Bound	18,0253	
	5% Trimmed Mean	16,9815		
	Median	17,0000		
	Variance	3,888		
	Std. Deviation	1,97183		
	Minimum	13,00		
	Maximum	20,00		
	Range	7,00		
	Interquartile Range	2,50		
	Skewness	-,377		
Kurtosis	-,232	1,121		

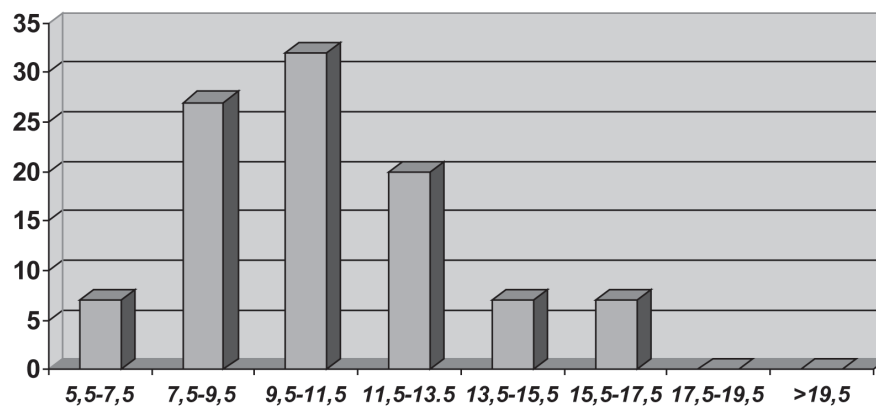
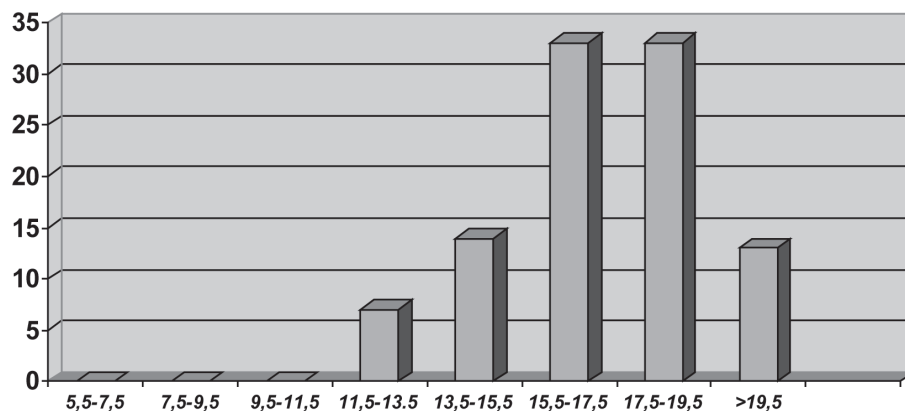


Figure 1: Negatively & Positively skewed distribution between Staff & Line Personnel

3. Results and main themes

1. The 'propitious niche' of performance appraisal are Fairness and integrity —The first challenge for the management is to make sure the subordinate views the appraisal as fair. However, there is evidence that in hierarchical type oriented organization, the management doesn't focus on the developmental aspect of the appraisal, but instead use it for political purposes.
2. The skewness distribution in scores due to technical problems can affect the fairness of the process . These are the followings:
 - a. Unclear Standards—Ambiguous traits and degrees of merit can result in an unfair appraisal.
 - b. Halo Effect—the influence of a rater's general impression on ratings of specific qualities.
 - c. Central Tendency—where supervisors stick to the middle of the rating scales, thus rating everyone average.
 - d. Leniency or Strictness—Supervisors have the tendency to rate everyone either high or low.
 - e. Bias—the tendency to allow individual differences such as age, sex, education to affect the appraisal ratings employees
3. The validity of appraisal can be achieved by using different potential raters to appraise a person's performance. The advantages of using several raters (either a rating committee or a combination of peer, supervisor, and subordinate ratings) are that the ratings tend to be more valid than those of one individual rater.
4. The appraisal must focus on develop plans to correct performance deficiencies and must not limited to administrative transactions (promotions, demotions, transferring, firing)
5. A fair performance appraisal process itself must be deployed in three steps: 1) define the job, 2) appraise performance, and 3) provide feedback. Defining the job means making sure that the rater and ratee agree on his or her duties and job standards. Appraising performance means comparing the employee's actual performance to the standards that have been set. Finally, the performance appraisal process requires one or more feedback sessions. In these sessions, the manager and employee discuss the employee's performance and progress and make plans for any development required.

4. Conclusion

The study has some implications regarding drawing the attention on the crucial role of performance appraisal fairness. Managers will develop appropriate strategies for enhancing the influence of their appraisal systems on work attitudes and outcomes.

Despite lots of attention, money, and effort, performance appraisals remain an area with which few managers or employees are satisfied. The following questions may be worth further investigations.

Is it just that we don't have a good enough system yet, is there an intrinsic problem with performance appraisals, or is it just human nature to dislike them?

To conclude, the effectiveness of the appraisal is so high quality, as it is high the Board of Directors' "Gnothi Seauton (Greek) Know thyself» towards organization endogenous problems.

It is hoped that this article can give a starting point and assumptions to be further researched by other scholars and practitioners on the main themes of how managerial attitudes, values, behaviours, and efficacy in a paternalistic managerial environment affects the fairness of appraisal.

References

- [1] Bernardin, H. John (2000), 'Conscientiousness and Agreeableness as Predictors of Rating Leniency', *Journal of Applied Psychology*, 85, 232-234.
- [2] Clinton, Wingrove, (2003) "Developing an Effective Blend of Process and Technology in the New Era of Performance Management," *Compensation & Benefits Review* Jan.-Feb.: 25-30
- [3] Gary, Gregures et al. (2003), "A Field Study of the Effects of Rating Purpose on the Quality of Multiscore Ratings" *Personnel Psychology* 56: 1-21
- [4] Lam, S.S.K., Schaubroeck, J., & Aryee, S. (2002). Relationship between organizational justice and employee work outcomes: A cross-national study. *Journal of Organizational Behaviour*, 23(1), 1-18.
- [5] Levine, H.Z. (1986). Performance appraisal at work. *Personnel Journal*, 6, 63-71.
- [6] London Manuel, Mone Edward M., Scott John C. (winter 2004) 'Performance management and assessment: Methods for improved rater accuracy and employee goal setting' *Human Resource Management* 43 : 319-336
- [7] Mead, R. (1994). *International management: Cross cultural dimensions*. Oxford, England: Blackwell.
- [8] Posthuma, Richard (2011) 'Twenty Best Practices for just Employee Performance Reviews' *Compensation & Benefits Review* Jan.-Feb.: 146-164
- [9] Uichol Kim (Editor), Harry C. Triandis (Editor) et.al. , *Individualism and Collectivism: Theory, Method, and Applications (Cross Cultural Research and Methodology)* Paperback – July 15, 1994 () (No. 18, pp. 251-266). Thousands Oaks, CA: Sage.
- [10] Walsh, M.B. (2003). *Perceived fairness of and satisfaction with employee performance appraisal* (Unpublished doctoral dissertation). Louisiana State University and Agricultural and Mechanical College, Baton Rouge.