

Performance Appraisal Biases In A Public Organization: An Emprical Study

Sait Gürbüz*

Onur Dikmenli.**

Abstract: The objective of the present investigation is to find out the existence of performance appraisal errors or biases in a public organization in Turkey. Magnitude of six types of appraisal errors was studied: the halo effect, the horn effect, the recency effect, the error of strictness, the leniency error and similarity effect. After the theoretical framework was provided, attitudes of the personnel towards to the six appraisal biases tried to be determined in the context of the research hypotheses using sample data collected from 150 public personnel, who are the rater and the ratees in the system. According to the results of the study, personnel who work in the organization think that six performance appraisal errors or biases are present in the public performance appraisal system. Attitudes of the public persons concerning the appraisal errors are significantly varied according to their status, but their ages. Implications for the appraisal errors are discussed, limitations of the study are revealed, and future research directions offered.

Key Words: Performance Appraisal, Performance Evaluation, Performance Appraisal Biases, Public Organizations

I. Introduction

In today's competitive business world, it is understood that organizations can only compete with their rivals by innovating and organizations can be innovative by managing their human resources well. The human resource system can become more effective by having a valid and accurate appraisal system used for rating performances of the employees. Unfortunately, the number of the organizations using an effective performance appraisal system is limited (Yalçın, 2002).

* Yzb. Sait Gürbüz.

** Onur Dikmenli, İstanbul Üniversitesi İşletme Fakültesi, Yönetim ve Organizasyon Bilim Dalı'nda araştırma görevlisidir.

The number of public organizations using a good performance appraising system is narrowly bordered. Because public organizations have been still under the influence of the central authority today even if they are managed autonomously (in developing countries such as Turkey, this autonomy is open to discussion). In other words, there is no comprehension in public organizations such as “making profit” and “paying the highest return to the stakeholders” as there is in private organizations (Boland and Fowler, 2000: 417). Due to these characteristics of the public organizations, without any forcible external power, we can say that performance measurements can make limited change at the comprehension of public organization being good for its own condition (Corbett and Kenny, 2001).

Performance appraisal has occupied the attention of researchers in human resource management (HRM) for many years. In USA, performance appraisal was used for the first time in 1915 by public organizations. In Turkey, performance appraisal was used for the first time in 1948 at Karabük Iron and Steel Factory, afterwards at public organizations like Sümerbank, Mechanical and Chemical Industry Corporation, Turkish State Railways and after 1960 at some private organizations (Örücü and et al, 2003).

The performance appraising systems, used a century ago were in primitive conditions for contemporary management. Within the last century, improved systems were developed so as to evaluate performance of the employees fairly and correctly. Today, most of the organizations have relied upon some forms of performance appraisal systems to decide the pay raises, promotions, training needs and individual improvement of the employees (Landy and Farr, 1980; Wexley and Latham, 1981; Cleveland, Murphy and Williams, 1989).

In order for a performance system to be perceived as fair, it must be free of bias. It has been known that appraisal errors (e.g.: the halo effect, leniency effect) can harm perceptions of pay system fairness by confusing the relationship between true performance differences (Miceli, Jung, Near and Greenberger, 1991). In Turkey, there are some researches concerning the performance appraisal (Yücel, 1999; Erdoğan and Medihan, 2002; Akın, 2002; Tak, 2003). However, little or no research has been conducted about performance appraisal biases in terms of public organizations perspective.

The objective of this paper is to find out the existence of possible performance appraisal biases (the halo effect, the horn effect, the recency effect, the error of strictness, the leniency error and similarity effect) in the current performance appraisal system of a public organization in Turkey. The employees, whose performances are being appraised and the raters, who are appraising the performances are accepted as the users of the current performance appraisal system in the surveyed or-

ganization. After the introduction, the importance of the performance appraisal, comparison of public and private organizations and recent studies from Turkey, concepts of the performance appraisal, process and methods of performance appraisal, biases in performance appraisal are clarified at the second part of the study. At the final part of the study; the hypotheses are tested, the method and the measures are explained, the findings are discussed, limitations of the study are revealed, and future research directions offered.

II. Literature Review

A. The Importance of Performance Appraisal

Over the past sixty years performance appraisal has become very popular in organizational life so that almost every company has an appraisal system. Researches show that there are many approaches for evaluating employee behavior and performance relative to job tasks and/or organizational culture. As a result, various applications of performance appraisals have left managers in a state of confusion and frustration with the employee evaluation process. In fact, no one seems satisfied with the system they gain the results it produces. As Grote said; "Performance appraisal systems are a lot like seat belts, most people believe they are necessary, but they don't like to use them" (Grote, 1996: 214)

Appraisal process is a continuing process for employees. The employees, after they are hired, will be subjected to evaluation process periodically. So employees' current position, improvement, awards and career plans are determined by these evaluation activities. The literature on performance appraisal has evolved from an early emphasis on the person, through the focus on the job and the person. Early performance appraisal methods were fairly simple and involved ranking and comparing individuals with other people (e.g., simple ranking methods) (Milkovich and Boudreau, 1997: 287). However, these early person-based measurement systems often exhibited a number of problems. As a result, a transition to job-related performance assessments has been made. Thus, performance measurement was modified from being person oriented to behavior oriented, the emphasis being on those tasks or behaviors that were associated with a given job (Welbourne, Johnson and Amir, 1998: 543).

It is unavoidable to make appraisal biases when the ratees and raters are human beings (Bakan and Kelleroğlu, 2003), and affect is irrevocable during the process. Perhaps it is impossible for manager to set aside their feelings when participating performance appraisal. For improvements of the system to be done in the shortest

time and at the best manner, it is significant that these biases are minimized and the system is reviewed frequently. Otherwise the employees may think that performance appraisal system used in the organization is inaccurate and will not be interested in it and will give misleading knowledge about the system (Miller and Thornton, 2006)

In today's business world, it seems like a chore for organizations to allocate significant intangible resources for planning, developing and engaging in employee performance appraisals. Despite this allocation, most organizations could not afford to give formal performance evaluations. There are three key reasons why every organization should give performance reviews: (Performance Appraisal, 2006)

- Performance appraisals help employees become better employees by providing feed back about their weaknesses and strengths.
- Performance appraisal measurement can help determine employee compensation, training needs and promotion.
- Performance appraisals can protect organizations against legal cases when countering to claims of illegal activities.

As reported in a recent study (Pulakos, 2004), only ten percent of the employees believe that their firm's performance appraisal help them to improve performance. Appraisal errors, lack of objectivity and nonperformance variables such as age, sex and race cause difficulties in the appraisal process (Miner, 1968; Dornbush and Scott 1975, Winstanley, 1980). Errors in performance appraisal continue to be of interest because of the impact that perceptions of unfairness have on employees, work groups, and organizations. Bias is inherently difficult to study, though, because appraisal differences across groups may reflect unfair preconceptions held by those making evaluations, errors that are inherent in the evaluation process (e.g., halo effects), or real differences in performance.

It has been known that appraisal errors can harm perceptions of pay system fairness by confusing the relationship between true performance differences (Miceli and et al.,1991). Huber, (1983), Kane and et al. (1995) also found that appraisal errors can undermine the potential beneficial of merit appraisals on employee motivation. Wiersma and Latham (1986), Tziner and Kopelman (2002) revealed that if ratees are dissatisfied with appraisal system, the effectiveness of the overall appraisal and feedback process is diminished. Hence, for effectiveness and accuracy of an appraisal system, being free of biases and errors is essential.

B. Comparison of Public and Private Organizations

In addition to the factors mentioned above, there are three environmental factors which make public and private organizations managed differently (Harel and

Tzafirir, 2002): In the private sector, organizations must worry about their profit margins and bottom-line performance. In contrast, public sector managers and organizations are judged by how well they provide service or correct problems, therefore they have high disincentives to control labor cost. The second factor is the legal and constitutional framework which makes the public employment relationship fundamentally different from that in the private sector. Finally, another major difference is the diffusion of authority in the public sector as opposed to a single, formal line or chain of command. In the private sector, employees perceive that they have only one boss while, in the public sector, employees must respond to executive and legislative political superiors.

Many elected officials believe pay-for-performance, based on appraisal ratings, will give employees incentives to improve productivity (Mani, 2002).

There are studies about performance measurements and performance appraisals in public and private organizations. For example, Cederblom, (2002) studied the organizations, incorporating the developments of Total Quality Management (TQM) system and core competencies to the performance appraisals and moving towards performance management from performance appraisal. Jansen, (2004) studied measuring the performance in governmental organizations and controlling the management by performance appraisals. Developing and improving performance appraisal determinants for public organizations were also studied by Jas and Skelcher (2005). Redman and et al., (2000) revealed the effectiveness of performance appraisal systems and their contribution to HRM activities in public organizations. The impact of performance criteria on the performance appraisal was studied by Hennessey and Bernardin (2003). Ndambakuwa and Mufunda (2006) revealed the impact of performance appraisal system on the productivity and job satisfaction.

In Turkey, researches concerning the performance appraisal systems and performance appraisal biases have focused both on the public and private organizations too. Yücel (1999) analyzed performance appraisal as a new concept by emphasizing the importance of performance appraisal for the employees and identified the opinions of public and private sector employees for the performance appraisal. Erdoğan and Beyaz, (2002) revealed the appraisal errors deriving from appraisers in private organizations. Akın (2002) studied the “Coaching” approach as a performance appraising method theoretically. Tak, (2003) analyzed performance appraisal system based on TQM principles and offered several proposals to integrate performance appraisal and TQM. Örucü and Mortaş (2005), studied the resistance of the employees to the new job analysis and performance management system at Muğla University. Örucü and et al (2003), studied the support and confidence of the employees to the new performance appraisal system at Muğla – Yatağan Thermal Station. Bakan and

Kelleroğlu (2003), studied the efficiency of the performance appraisal system from the perspective of the employees at both public and private organizations. Yalçın, (2002) studied the variables influencing the employees to prefer 360 Degree or Group Based performance appraisal system in private organizations.

C. The Concepts of Performance Appraisal

Performance means “a basic instructional method in which the trainee is required to perform, under controlled conditions, the operation, skills, or movement being taught” (Tracey, 1998: 391). **Performance appraisal** is defined as evaluating employees how well do their jobs according to performance standards (Dessler, 2000; 321). After an employee has been selected for a job, has been trained for it, and has worked for a period of time, her or his performance should be reviewed. So, performance appraisal is one of the most significant topics in Human Resource Management.

Traditional performance appraisal relies on economic reward and the threat of punishment to motivate employees to reach desired performance. But this concept does not hold true anymore. Today, performance appraisal is used for developmental and motivational purposes in the organizations.

Performance appraisal is not a still evaluation activity, but a dynamic process, which should be viewed as follows; planning the employees’ performance, evaluation, and improving the performance of the employees’. This process brings the new concept: performance management (Kaynak et al., 2000: 205). This concept is essential for today’s organization to integrate the management goals and employee performance. **Performance management** is a system for integrating the management of organization and employee performance in order to support and improve company’s or organization’s over all business goals (Williams, 1988).

D. Process and Methods of Performance Appraisal

In order to obtain benefits from the appraisal system, managers should give special attention to the design of the appraisal process. In organizations performance appraisal is conducted in a series process. This process can be categorized as follows:

1. Establishing Job Criteria and Appraisal Standard. (A decision regarding what to measure)
2. Timing of Appraisal. (Determining how often Performance Appraisal will be conducted)
3. Selection of Appraisers. (Deciding who will be the appraiser/s)
4. Providing Feedback. (Helping employees to see their strengths and weaknesses)

Performance appraisal could be done by anyone who is familiar with a person's performance according to chosen appraisal technique including the following (Schuler, 1995: 312-316; Mathis and Jackson, 1994: 330-334; Casio, 1995: 290-291): supervisors, subordinates, peers, customers, and self-appraisal and multi-source feedback. Yet, that it is not necessary to choose only one type of performance evaluation method and some organizations successfully use all or some of them.

Formal performance appraisals have become a standard for administrative and developmental purposes. Despite their widespread use, there are dissatisfactions and problems with the feedback systems associated with single source performance appraisal. Therefore, in response to these concerns, considerable emphasis has been placed on developing multi-source feedback systems. Unlike traditional evaluation, which explained above, 360-degree performance appraisal has multiple sources to appraise the performance of the employee. These sources can be supervisors, subordinates, peers or coworkers, self and customers.

Lepsinger defines the 360-degree as follows: "The 360-degree feedback process involves collecting ideas about a person's behavior from the person's boss or bosses, direct reports, colleagues, fellow members of project teams, internal and external customers, and suppliers" (Lepsinger, 1998: 49). The 360-degree appraisal is also known as "multi rater feedback, multi source feedback, full circle appraiser, multi-dimensional evaluation, and up-wards feedback appraisals". The trend for this system is rapidly increasing. Despite some disadvantages of the system, it is still gaining acceptance in the business world. 360-degree feedback has many distinguished features from other traditional types of performance appraisals. First, information and feedbacks about the employee's performance are collected from many sources, as explained above. As a result, gaining a better understanding of others' perspectives allows the individual to see her/his own strengths and weaknesses more clearly. Employee development efforts are more focused and effective.

Second, the raters in the 360-degree feedback know and interact with ratees frequently. Thus they are the right people to appraise the performance of the ratees. So, the person gets a broader range of performance information. She/he has the opportunity to view their performance through the eyes of those with whom he or she works most closely. These ensure that more compressive information is gained in 360-degree Performance Appraisal than the traditional ones (Hurley, 1998: 202-204).

In HRM literature, many appraisal methods can be used to evaluate employee's performance. Most commonly and popular appraisal methods could be categorized in three groups, which are comparative appraisals, behavioral appraisals and output-based appraisals (Fisher, Schoenfeldt and Show, 1996: 469).

E. Biases in Performance Appraisal

In practice, there are many possible errors or biases in the performance appraisal process. The raters cause most of these errors. These errors affect the objectivity of the appraisal very much. The most common errors in appraisal include:

1. The Halo Effect and Horn Effect.
2. Leniency Error. (Loose rater).
3. The Error of Strictness. (Tight rater).
4. The Central Tendency Error.
5. The Recency Effect.
6. The Contrast Error.
7. The Similarity (similar-to- me) Effect.

The halo and horn effect

The halo effect has been defined as “the influence of a rater’s general impression on ratings of specific rate qualities” (Solomonson and Lance, 1997). In other words, rater gives the subordinates good grades although their performances are not worthy. For example, if an employee has few absences, her / his rater may give her / him a high rating because of her / his dependability. This is because; raters sometimes cannot evaluate the employee’s other characteristics separately. Raters often overlook poor performances, if they like the subordinate. Lefkowitz concluded from his review of 24 studies that positive regard for subordinates is often found related to greater halo effect and better interpersonal relationship. (Lefkowitz, 2000).

Horn effect is the opposite of halo effect. It means that rater might give poor grade although ratee’s performance is worthy. In other words, some appraisers have tendencies to view negatively all behaviors or actions of a subordinate because the superior dislikes a particular behavior or action of the subordinate. In order to reduce the halo and horn effects during rating process some solutions have been proposed (Uyargil, 1994: 74). One of these solutions is rater training. Another one is to ask raters to give proof of the events or problems, which cause to them to evaluate the ratees poorly.

The Leniency error

This is the second most common appraisal error. Some raters or managers are concerned about damaging a good working relationship by giving poor or negative rating. For that reason, they have tendencies to give a high rating to ratees. It can be said that the leniency errors may result from the purpose of giving high rating. Per-

sonnel psychologists have said that performance appraisal ratings obtained for administrative purposes (such as pay raises or promotions) would be more lenient than ratings aimed for feedback or employee development purposes (Jawahar and Williams, 1997) Some raters are called “**loose rater**” open to make this error. They have the tendencies to give top rating to all subordinates unless they have a clear deficiency, or they may not want to “adversely impact” the future of subordinates. Also, (by giving the poor grade) they do not want to be perceived as “unwanted person” on the eye of the ratees

The error of strictness

This error is the opposite of the leniency in which raters give unfavorable or poor appraisal regardless of actual performance level of the ratees. In other words, some raters called “**tight raters**” have very high evaluation standards. For example they might say, “The highest I would give is ninety (out of a hundred)”. According to them nobody is perfect so anybody cannot pass the 90 points. In order to reduce this error, rater training (including supervisor subordinate role and being objective), should be reviewed. Some raters make these errors because of the following reasons; (Kaynak et al., 2005: 223)

1. They are afraid that successful employees will replace them in the future.
2. Some raters or managers want to describe themselves as a hard and perfectly based manager. As a result, they avoid giving high rating even if ratee’s performance is worthy.

The central tendency error

Rather than giving extreme poor or good grades, there is a tendency on the part of some raters to evaluate all raters as average score even if performance actually varies. In other words, some raters want to rate the employees in the middle of the scale than too high or too low. For instance, if the appraisal scale ranges 1 to 8, appraisers are willing to avoid giving high score (7, 8) and low score (1 and 2). So, they do evaluate most of the employees between 4 to 6. This error is mostly done for the following reasons (Dessler, 2000: 321):

1. The rater doesn’t know ratee very well, so they prefer this way, which has very low risk to avoid wrong judgment.
2. Also some raters believe appraisal is a waste of time; as a result they give average rating regardless of employee’s actual performance value.

The Recency effect

As we have explained, appraisal is generally conducted once or twice a year. This period might be very long for the rater to remember all performance-relevant information of the employees. As the appraisal time approaches, the rater tries to find information that reflects value of performance. Unfortunately, recent events or behaviors are more noticeable. As a result, recent events are weighted more heavily than should be. So, some raters only see the ratee's latest behavior regardless of employee's actual performance. But performance appraisal should range all appraisal period (Uyargil, 1994: 79). Employees and managers can minimize this error by keeping ongoing behavioral or critical incident files in which good and poor behaviors and outputs are recorded. Although time consuming, they ensure that information for the entire period is incorporated into the appraisal. Yet, there is important point to remember, some ratees work very hard and demonstrate good performance when appraisal time is approaching.

The contrast error

As explained before rating should be done on the basis of standards that are established before the rating process. The contrast error is the tendency to rate people relative to other people rather than to performance standards. If the raters appraise many employees in a short time, it is unavoidable that they do the appraisal by comparing the ratees. In other words, appraisal grade of a ratee might be affected by the grade of the ratee who is appraised just before her or him. For example, if everyone else in a group is standing poor performance, a person performing somewhat better may be appraised as a perfect due to contrast error. The vice-versa is also possible. This error could be minimized but not grouping employees as "successful" or "unsuccessful". All employees must be appraised randomly (Uyargil, 1994: 80).

The similarity effect

Some appraisers overlook the actual performance of the ratee; on the other hand, these kinds of raters have the tendencies to give better rating to those subordinates similar to themselves in behavior, personality or background (Pulakos and Wexley, 1983:131). Also the ratee might cause this error. Since, some ratees have made efforts to demonstrate that their behaviors, tastes and tendencies match those of the superior or hide those not matching with the superior's, with the intention that such would please the superior and consequently receive better ratings.

The best solution to reduce rater errors is appraiser rating training and follow-up training on how to rate performance appraisals. Although training may be the answer in some cases, in others it's not as useful due to other factors that distort ratings (Gilbert, 2006). For example, Roberts, 1998, found four of out ten supervisors agreed that employees receive much of the blame for poor performance when in reality its poor management practices.

There is no way to eliminate these errors completely. In order to minimize these errors, recent studies suggested that raters must be aware of the system very well and organizations should provide rater training for their managers (Keown and Janine, 2001; Noonan and Sulsky, 2001; Roch and O'Sullivan, 2003).

III. Research

A. Description of the Current Performance Appraisal System of the Surveyed Organization

The surveyed organization's current performance appraisal system (PAS) is based on appraisal of an individual's performance by his/her 1st and 2nd superiors. Thus, the system is a typical "superiors rate only" system. The overall performance rating assigned to the annual appraisal. In the organization, a weighted checklist is being used to evaluate the personnel. The weights of the appraisal factors are different, but rater and ratees do not know these weights. The appraisal form or weighted checklist has 35 appraisal criteria, which are the core of the appraisal system for personnel. Each criterion or qualification has up to 3 or 4 or 5 statements describing specific behavior to be marked by the two superiors. These qualifications include public personnel's traits, values and personal competencies (e.g.: Self-confidence and self – control, honesty and integrity with righteous, attaching importance to appearance, level of cultural knowledge, planning ability, mission knowledge). Appraisals are done in May of each year. Rater appraises the ratee by considering ratee's performance considering his or her performance; rater also could write his/her comments on the form. The appraisal forms are filled out manually by the raters who are 1st and 2nd superiors of the ratees. They choose the appropriate choices for the related qualifications. These qualifications are graded to determine ratee's performance appraisal score as a numerical value (% 100 scales). Through a computer program, appraisal score of the rate is calculated by using this form. Rater provides appraisal feedback related to the some qualifications but not for all. Ratees see their strengths and weaknesses about these items which were provided. In other words, 1st rater provides feedback for only the some qualifications to ratees.

B. Hypotheses Tested

Main and sub hypotheses are to be tested in the study are as follows:

H₁: Personnel who work in the organization think that six performance appraisal errors or biases are present in the public PAS. Hypothesis 1 has 6 sub hypotheses which are;

H_{1A}: The personnel think that the raters do the halo error during the appraisal.

H_{1B}: The personnel think that the raters do the horn error during the appraisal

H_{1C}: The personnel think that the raters do the recency error during the appraisal

H_{1D}: The personnel think that the raters do the error of strictness during the appraisal

H_{1E}: The personnel think that the raters do the leniency error during the appraisal

H_{1F}: The personnel think that the raters do the similarity error during the appraisal

H₂: Attitudes of the public persons concerning the appraisal errors are significantly differentiated according to their status. Hypothesis 2 has 6 sub hypotheses which are:

H_{2A}: Attitudes of the public persons concerning the halo effect are significantly differentiated according to their status.

H_{2B}: Attitudes of the public persons concerning the horn effect are significantly differentiated according to their status.

H_{2C}: Attitudes of the public persons concerning the recency effect are significantly differentiated according to their status.

H_{2D}: Attitudes of the public persons concerning the error of strictness are significantly differentiated according to their status.

H_{2E}: Attitudes of the public persons concerning the leniency error are significantly differentiated according to their status.

H_{2F}: Attitudes of the public persons concerning the similarity effect are significantly differentiated according to their status.

H₃: Attitudes of the public persons concerning the appraisal errors are significantly varied according to their ages. Hypothesis 3 also has 6 sub hypotheses which are:

H_{3A}: Attitudes of the public persons concerning the halo effect are significantly varied according to their ages.

H_{3B}: Attitudes of the public persons concerning the horn effect are significantly varied according to their ages.

H_{3C}: Attitudes of the public persons concerning the recency effect are significantly varied according to their ages.

H_{3D}: Attitudes of the public persons concerning the error of strictness are significantly varied according to their ages.

H_{3E}: Attitudes of the public persons concerning the leniency error are significantly varied according to their ages.

H_{3F}: Attitudes of the public persons concerning the similarity effect are significantly varied according to their ages.

C. Method

Sample and Procedures

This study conducted in some branches of a public organization in Istanbul. However, name of the organization was not stated openly, since results of the study might affect image and identity of the organization negatively. Thus, authors and top managements of the organization assessed that title of the organization should be anonymous for reputation and identity of the organization.

Questioners were used to collect data from employees and managers who are members of the organization. The respondents were asked about their perception of appraisal error or biases in the current performance evaluation system. No personal data was collected except some demographics such as, age, status, education and tenure.

Firstly, each question translated from English into Turkish then back to English to make sure of no meaning changes. Then, pilot survey was conducted in 55 randomly selected public personnel. During the pilot survey testing, the questions proved to be easily understood by the participants.

Stratified Random sampling method is used in the study. This sampling method is used to separate the population into subgroups in accordance with some features such as age, sex, marital status, and profession (Seyidođlu, 2000: 40). In this study, respondents are selected in accordance with their status, manager and employees. Considering universe which stands for numbers of personnel who work as manager or employees in surveyed organization's branches in Istanbul, stratified ratio for managers should be 60 % and 40 % for the employees.

154 questionnaires were distributed to the participants and 140 of them returned. For a response rate of 90 percent. 42 percent of the respondents were managers; the remaining 58 percent were employees. Tenure ranged from less than five years (17 percent) to 5-10 years (33 percent), to more than 11 years (50 percent). The participants were predominantly male (99 %), and the median age was between 32 and 41 years old. The Statistical Package for Social Sciences Program (SPSS-14) was used to analyze the gathered data.

Measures

In order to measure perception of appraisal error or biases in the current performance evaluation system of the public personnel, a 9-item appraisal errors scale, which was adapted from Bıçaksız et al (2002), was used in the study. Bıçaksız et al (2002) developed the scale according to literature description of appraisal errors and used the scale for another public organization in Turkey. We made some minor modification on the scale due to purpose of our research and characteristics of the public organization. The performance appraisal errors scale consists of 9 items. To detect presence of appraisal errors in the system, following items were asked to the respondents:

- Questions 6 and 7 measure the halo effects perceptions
- Questions 8 measure the horn effect perceptions
- Questions 9 and 10 measure the recency effect perceptions
- Question 11 measures the error of strictness perceptions
- Question 12 measures the leniency error perceptions, and
- Questions 13 and 14 measure the ratee's similarity effect perceptions.

We asked these questions to understand the presence of the performance appraisal errors in the system with some demographics items. The questionnaire was presented at Appendix-1. The scale consists of the five possible survey responses for second part of the questionnaire, which are "don't know, none, some, most, all". Responses are scored on a scale of 0 to 5, where 0 represents the complete lack of subjective error in question and 5 represents the presence and maximum extent of the subjective error. "Do not know" is assigned a zero value to differentiate it from the "none" response. The Cronbach alpha reliability for the overall scale was 0.86. The coefficient alpha estimate is greater than the level of 0.70 which is the recommended level for researches of the social sciences.

D. Results

Demographics result

Table-1 presents the descriptive statistic of demographics of personnel, working in the public organization. As seen in the table, most of the respondents work as employees rather than managers. 58.5 % of the personnel are 27-36 age years old. High school and below graduates and four year college graduates holds nearly 83 % of all. Gender is predominantly male, and 40 % of personnel have between 7 and 15 years tenure.

Table 1: Descriptive statistic of demographics of respondents

Variable	Group	Frequency	%
Job status	Managerial	58	41,4
	Official	82	58,5
Age	26 and below	18	12,8
	27-36	82	58,5
	37 and above	40	28,5
Education	High school and below	60	42,8
	Two-year-college	12	8,5
	Four-year college	56	40,0
	MA-MBA	10	7,1
Gender	Ph.D.	2	1,4
	Male	121	86,4
	Female	19	13,5
Tenure	6 and below	21	15,0
	7 and 15	57	40,7
	16 and above	42	30,0

Table 2 presents the means and standard deviations of the study variables. According to the Table-2, it can be said that all appraisal biases, investigated in this study are present in the current performance appraisal system. As seen in the table, halo and similarity effect are the highest in average, with 2,52 and 2,66 respectively (out of 5 scale). The leniency and strictness errors are the lowest in average, 1,91 and 2,15 points respectively. It is clear that the system is under the influence of appraisal errors according to attitude of the respondents.

Table 2: Means, Standard Deviations of the variables

Variables	Mean	Standard deviation
1. Halo Effect	2,52	0,64
2. Horn Effect	2,34	0,81
3. Leniency Error	1,91	0,88
4. Error of Strictness	2,15	0,48
5. Recency Effect	2,47	0,75
6. Similarity Effect	2,66	0,62
Overall	2,34	0,70

In order to understand the ideas and attitudes of the persons, we have evaluated descriptive statistic tables of the six appraisal errors below.

Halo effect ideas

In Figure-3, it can be easily seen that approximately 47 % of the public personnel believe that “some” of the raters have the tendencies to view positively all behaviors or actions of subordinates because the superior or rater likes a particular behavior or action of the subordinate. In other words, rater gives the subordinates good scores although their performances are not worthy. Nearly 34 % of the public personnel believe that “most” of the raters do this error. And almost 6 % of them think “all” of the raters do this error.

Horn effect

Horn effect is the opposite of halo effect. It means that rater might give a poor grade although ratee’s performance is worthy. We have asked to find out “how many of your supervisors have tendencies to view negatively all behaviors or actions of a subordinate because the superior dislikes a particular behavior or action of the subordinates”. The result of this variable is presented above. As it is illustrated in Figure-3, approximately 31 % of the public personnel believe that “most” of the raters perform this behaviors, during appraisal, 44 of them think “some” of the raters show this error.

Recency error

In the examination of recency error in Figure-3, 50 % of the public personnel think that “some” of the raters have been influenced by frequent display of behaviors they like and the lack of behaviors they dislike in the evaluation time. Almost, 32 % of them believe that “most” of the raters act this way to give higher appraisal score.

The strictness

The error of strictness or tight rater is the opposite of the leniency or loose rater in which raters give unfavorable or poor appraisal regardless of actual performance level of the ratees. The figure shows that, 48 % of the public personnel think that “some” of the raters have very high rating standards and 14 % than think that “most” of the raters are the tight raters.

Table 3: Descriptive statistic of the appraisal errors

Appraisal Errors Response	Valid Percent (%)					
	The halo	The Horn	The recency	The strictness	The Leniency	The similarity
Don't know	8,4	11,4	10,7	6,6	9,9	11,6
None	5,0	7,4	3,3	28,9	32,2	5,0
Some	47,1	43,8	49,6	47,9	46,3	46,3
Most	33,9	31,4	32,2	14,0	10,7	31,4
All	5,7	5,9	4,1	2,5	, 8	5,8
Total	100,0	100,0	100,0	100,0	100,0	100,0

The Leniency

Some raters or managers are concerned about damaging a good working relationship by giving poor or negative rating. For that reason, they have tendencies to give a high rating to ratees. In Figure-3, we can see the ratio and ideas of the public personnel about this effect. According to the results, 46 % of the public personnel think that “some” of the raters have the tendencies to give to rating to all subordinates unless they have a clear deficiency, or they may not want to “adversely impact” the future of subordinates. When comparing the “loose rater” with the “tight rater”, we can say that “tight rater” ratio is higher than the “loose rater”.

Similarity effect

Some appraisers overlook the actual performance of the ratees; on the other hand, these kinds of raters have tendencies to give better rating to those subordinates similar to themselves in behaviors, tastes and tendencies. The most important danger of the similarity effect is that this error might limit diversity in the organization. In Figure-3, the rater’s similarity effects are presented. 46 % of the public personnel believe that “some” of the raters have tendencies to give better rating to these subordinates similar to themselves in behaviors, tastes and tendencies. 31 % of them think that “most” of the raters show the similarity effect.

To summarize all this data, it can be said that public personnel think that performance appraisal system of the public organization is affected by subjectivity, and under the effects of the six appraisal errors. To be completely objective is very difficult; on the other hand, subjectivity should be minimized as much as possible for effective appraisal system. With these findings, our first hypothesis and its six sub hypotheses (**H1: Personnel who work in the organization think that six performance appraisal errors or biases are present in the public PAS**) were supported. As seen in Table-2, the most common appraisal errors in the system are the similarity (similar-to-me) effect (M: 2, 66) and the halo effect (M: 2, 52).

Effect of Demographics Factors on Appraisal Biases.

In order to determine whether the attitudes of the public persons concerning the appraisal errors are significantly differentiated according to their status or not, we have performed an Independent t test. The means of the overall appraisal biases according to the status and Independent t test table are reported below.

Table 4: Means of the overall appraisal biases according to the status and Independent t test table

Appraisal errors	Means		Independent t test table		
	Manager (n ₁ :58)	Employee (n ₂ : 82)	t	Sig (2-tailed)	Mean difference
1. Halo Effect	1,92	3,12	8,758	0,000	-1,20
2. Horn Effect	1,81	2,87	8,235	0,000	-1,06
3. Leniency Error	1,74	2,08	3,241	0,001	-0,34
4. Error of Strictness	1,68	2,62	7,126	0,000	-0,94
5. Recency Effect	2,05	2,89	6,945	0,000	-0,84
6. Similarity Effect	2,01	3,29	8,823	0,000	-1,28
Total	1,87	2,81	7,245	0,000	-0,94

p <, 01

As seen in the table, means of the managers, working in the organization, are lower than those of the employees. We have executed t test so as to determine this

variation is significant or not in terms of status demographic. Independent t test result demonstrates that mean difference is significant according to the status of (manager or employee) the persons (all $ps < .01$). This result revealed that the attitudes of employees (M: 2,81) was significantly higher in overall and six appraisal biases than those of the managers (M: 1.87). In the survey, all managers are both raters and ratees. On the other hand, the employees are only ratees in the appraisal system. Hence, it can be said that ratees and raters do not share the same perceptions and attitudes concerning the appraisal biases. With these findings, our second hypothesis and its six sub hypotheses (**H₂: Attitudes of the public persons concerning the appraisal errors are significantly differentiated according to their status**) were also supported.

To test our final hypothesis we executed ANOVA. The ANOVA table and means concerning the age demographic are reported below. Before evaluating the ANOVA result, it will be beneficial to examine the means of the attitudes scores of the public personnel in terms of their ages.

Table 5: Means of the Overall appraisal biases according to the age and ANOVA table

Appraisal errors	Means			ANOVA	
	26 and below (n1:18)	27-36 (n2: 82)	37 and above (n3: 40)	F	Sig (2-tailed)
1. Halo Effect	2,45	2,59	2,52	,618	650
2. Horn Effect	2,24	2,42	2,32	1,146	335
3. Leniency Error	1,79	2,02	1,82	1,461	214
4. Error of Strictness	1,98	2,14	2,23	1,654	161
5. Recency Effect	2,32	2,45	2,59	1,924	106
6. Similarity Effect	2,35	2,59	2,45	1,596	175
Overall	2,19	2,37	2,32	,977	420

As seen in the table, attitudes of the respondents, whose ages between 27 and 36, concerning the appraisal biases are generally higher than those of the others. However, the ANOVA table, at the right column of the table, failed to show a significant differentiation. Hence, attitude concerning the overall appraisal biases did not vary

significantly by the age. Hence, our final hypothesis and its sub six hypotheses (**H₃: Attitudes of the public persons concerning the appraisal errors are significantly varied according to their ages**) were not supported.

E. Discussions and Implications

Performance appraisal system has gained considerable importance recently. Just because, all other HR functions such as; recruitment and placement, training and development, managing careers and fair treatment, pay for performance and financial incentives directly or indirectly depend on the performance appraisal system. In public organizations, in order to administer HR activities such as promotion, appointment, and merit system effectively, it is important to have a valid and effective performance appraisal system. The ultimate goal of most performance appraisal systems is to increase employee motivation and productivity. The related literature has begun to identify the characteristics of effective performance appraisal systems (Burke, 1978; Kane and Lawler, 1979; Bernardin and Beatty, 1984; Silverman and Wexley, 1984; Mohrman and Resnick and Lawler, 1989;. Roberts, 1990; Murphy and Cleveland, 1991; Daley, 1992; Cardy and Dobbins, 1994). In order for an appraisal system to be perceived as fair and effective, it must be free of bias. The purpose of this research was to determine attitudes of public personnel about the possible appraisal biases in their current performance evaluation system.

The public personnel think that the current PAS of the surveyed public organization, subject to wide variety of errors or biases, stemmed from appraisal system or raters (overall M: 2,34). According to the results, the most common appraisal errors in the system are the similarity (similar-to-me) effect and the halo effect (H₁). As, regarding H₂, attitudes of the public personnel concerning the appraisal errors are significantly differentiated according to their status (manager or employee). Personnel who work in official jobs (employees) (M: 2,81) was significantly higher in overall appraisal biases than the managers (M:1,87) according to the independent t test result. However, attitudes of the public personnel regarding the appraisal errors are not significantly varied according to their ages (H₃).

As for performance appraisals, Landy and Farr (1980); Latham and Wetley (1981) have suggested that affect is a source of bias in appraisals. Like, dislike, preference, evaluation, displeasure are the major affects in which social interaction is transacted (Zajonc, 1980: 153). These affective reactions are inescapable and usually irrevocable in interpersonal relations. Arvey and Campion (1982); Tsui and Barry (1986) found that feeling or affective reactions towards ratees can potentially influence raters, particularly, similarity effect, halo effect, leniency effect and central

tendency. As Zajonc (1980) said, affect is an irrevocable, perhaps it is impossible for manager to set aside their feelings when participating performance appraisal. However, public organizations should strive not for the suppression of feeling by any one rater but rather for a balance of affect among multiple ratees in performance appraisal.

Kluger and Denisi (1996) stated that it would be difficult to imagine ratees seeing everyone getting similar appraisals and perceiving such a system as fair or even accurate referring to leniency and central tendency effects. Miceli and et al. (1991); Kane and et al (1995) also found that appraisal errors can undermine the potential beneficial of merit appraisals on employee motivation. Considering the results of the study, public personnel think that the current performance appraisal system does not meet their expectations and the demands in terms of appraisal accuracy. For the better and effective PAS, the much importance should be given to performance appraisal which perceived as fair or even accurate process by the ratees.

Raters and ratees must have a shared perception of the purposes and functions of the process and the belief that the appraisal process is useful to them on an individual basis (Maroney and Buckley, 1992). Magnitude of the appraisal errors might cause rater to be unable to appraise the ratee's actual performance (Bowman, 1999). Also, in a survey, over 70 percent of manager reported that they deliberately inflated or deflated evaluations for horn effects (Longenecker and Ludwig, 1990). Finding of our study demonstrated that the attitudes of employees (M: 2, 81) was significantly higher in overall and six appraisal biases than those of the managers (M: 1.87). Hence, it can be said that ratees and raters do not share the same perceptions and attitudes concerning the appraisal biases. The raters can not evaluate the actual performance of the public persons due to the magnitude of the appraisal errors in the system.

The most important reason for wide variety of errors or biases in the system might be due to the method of appraisal. In the public organizations in Turkey, generally, "superior rate only" systems are utilized. These kinds of systems have some disadvantages. First, "the supervisor may have an ethical bias against playing god" (Schuler, 1995: 312). Secondly, the manager or superior sometimes cannot have enough skills to appraise and give good and timely feed-back to the ratees. Thirdly, if the rater exploits the power of appraisal, or there is no accountability of the rater concerning the appraisal process, the ratees or employees may feel threatened. Palmer and Feldman (2005) found that accountability increase the appraisal accuracy by reducing magnitude of appraisal errors. Because, when the rater explicitly told, prior to appraisal task, that they will be held accountable for their appraisal, they will think as accountability direct them and will provide appraisal consistent with contingencies of accountability requirements (Tetlock and Kim, 1987; Kunda, 1990). Therefore, in response to these concerns, considerable emphasis might be

placed on accountability of the rater and developing multi-source feedback systems instead of “superior rate only” system.

Wiersma and Latham (1986), Tziner and Kopelman (2002) revealed that if ratees are dissatisfied with system, the effectiveness of the overall appraisal and feedback process is diminished. Once appraisals have been made, it is important for the ratees to understand and see their weaknesses and strengths. Thus, discussing the appraisal result with the ratee is essential for organizational development. The current appraisal system provides limited feedback for the ratees. This prevents ratees from learning their weaknesses and strengths. So, the PA feed-back should be provided for the ratees including all the criteria.

Bernardin and Pence (1980); Fay and Latham (1982), Smith (1986), Woehr and Huffcutt (1994) suggested that training program will increase rater accuracy during the appraisal process. Also, more recent studies suggested that trained rater provide more accurate appraisal than untrained ones (Keown and Janine, 2001; Noonan and Sulsky, 2001; Roch and O’Sullivan, 2003). Hence, training program for the raters to reduce the appraisal errors might actually have the effect of increasing appraisal accuracy in the surveyed public organization.

The research has some limitations as well. First, all factors (six appraisal errors) were assessed through self reports which might create the potential for common method bias. This bias might have inflated or deflated the appraisal biases. Another limitation of our research is that because of time and financial constraints, this study was conducted in only one public organization. The findings from this study should be viewed cautiously due to the limited sample size. Also, in the study, only six appraisal errors (the halo effect, the horn effect, the recency effect, the error of strictness, the leniency error and similarity effect) were tried to be measured. But, as it is known there might be other appraisal errors such as, central tendency, contrast error, and attribution error. So, these appraisal errors should be studied as further study. Despite the limitations, the results are potentially important since, to date, no previous studies have examined the magnitude of the appraisal errors in the public organization in Turkey. Hence, the research may prove useful for guiding future researches. Additional research is necessary to further analyze the relationship between rater attitudes/beliefs and the design and implementation of the public organization’s appraisal system. It would be valuable for further studies to determine specifically, the reasons, causes of these errors, and ways for minimizing the errors as well.

Bir Kamu Örgütünde Performans Değerleme Hataları: Ampirik Bir Araştırma

Özet: Bu araştırmanın amacı, Türkiye'de bir Kamu Kuruluşunda kullanılan Performans Değerleme Sisteminde olası performans değerlendirme hatalarının varlığını tespit etmektir. Araştırmada, hale etkisi, horn etkisi, benzerlik etkisi, sıkı değerlendirici etkisi, gevşek değerlendirici etkisi ve benzerlik etkisi olmak üzere altı adet değerlendirme hatasının mevcudiyeti sorgulanmıştır. Konuyla ilgili teorik çerçevenin ortaya konulmasından sonra, 150 kamu personelinden oluşan bir örneklem kullanılarak, performans değerlendirme sisteminde değerlendirici ve değerlendirilen olarak yer alan personelin altı adet değerlendirme hatasına yönelik tutumları, araştırmanın hipotezleri çerçevesinde test edilmiştir. Araştırmanın sonuçlarına göre, kamu personeli performans değerlendirme sisteminde altı adet değerlendirme hatasının var olduğuna inanmaktadırlar. Araştırmaya dahil edilen personelin değerlendirme hatalarına ilişkin tutumları statülerine göre anlamlı düzeyde farklılaşmaktadır. Yaş demografik değişkenine göre ise anlamlı bir farklılaşma tespit edilememiştir.

Anahtar Kelimeler: Performans Değerlendirme, Başarı Değerleme, Performans Değerlendirme Hataları, Kamu Örgütleri

References

- Akın, Adnan, "İşletmelerde İnsan Kaynakları Performansını Değerleme Sürecinde Coaching (Özel Rehberlik)", C.Ü.İktisadi ve İdari Bilimler Dergisi, 3 (1), 2002, 97 - 113.
- Anderson, D., **Performance Appraisal**, FBI Law Enforcement Bulletin, 63(4), 1994.
- Aplaki Soner, "A Methodology Proposal for Personal Appraisal System in Hierarchic Organizations", (Unpublished Master Thesis), Yeditepe University Graduate Institute of Science and Engineering, İstanbul, 1999.
- Arvey R.D and Campion J.E. "The Employment Interview: A Summary And Review Of Earlier Research", Personnel Psychology, 35, 1982, 281-322.
- Bakan İsmail and Kelleroğlu Hakan, "Performans Değerlendirmenin Etkinleştirilmesi Yönünde Çalışanların Düşünceleri: Tutumlar, Beklentiler ve Sonuçlar Konusunda Bir Alan Çalışması", Yönetim ve Ekonomi - Celal Bayar Üniversitesi İ.İ.B.F. Dergisi, 2003, Cilt:10, Sayı:2, 63 - 74
- "Benefits of Appraisal". Available On Site:
<http://www.performanceappraisal.com/benefits.htm> (28.02.2006).
- Bernardin H.J. and Beatty R. W., **Performance Appraisal: Assessing Human Behavior at Work**, Boston: Kent Publishing Company, 1984.
- Bernardin, H.J. and Pence, E.C. "Effects of Rater Training: Creating New Response Sets and Decreasing Accuracy", Journal of Applied Psychology, 65, (1980), 60-66.

- Bıçaksızı Adnan et al, "**Personel Değerlendirme Sistemi Projesi**", Ankara, 2002.
- Boland, Tony, and Fowler, Alan, "**A Systems Perspective Of Performance Management In Public Sector Organizations**", The International Journal Of Public Sector Management, 2000, Vol:13, Iss:5, 417 - 430
- Bowman JS, "**Performance Appraisal: Verisimilitude Trumps Veracity**", Public Personnel Management, , 28, 1999, 557-577
- Boyett Joseph H. and Conn, Henry P., **Maximum Performance Management: How To Manage And Compensate People To Meet World Competition**, London: Gland bridge Publishing Ltd., 1988.
- Burke R. J., "**Characteristics of Effective Employee Performance Reviews and Development Interviews: Replication and Extension**", Personnel Psychology, 31, 1978, 903-919.
- Byham, W.C. and Thornton, G.C Assessment Centers. In R.A.Berk (Ed.), **Performance Assessment: Methods And Applications**, London: The Jones Hopkins University Press, 1986.
- Cardy R.L. and Dobbins, G.H. **Performance Appraisal: Alternative Perspectives**, Cincinnati: South-Western Publishing Co., 1994
- Cascio, W.F., **Managing Human Resources**, N.Y.: McGraw-Hill, op.cit., 1995.
- Cederblom D., "**From Performance Appraisal To Performance Management; One Agency's experience**", Public Personnel Management, Summer 2002, Vol:31, Iss:2, 131 - 141
- Cleveland, J.N., Murphy, K.R. and Williams, R. E., "**Multiple Uses Of Performance Appraisal: Prevalence And Correlates**", Journal of Applied Psychology, 74, 1989, 130-135.
- Corbett, Barry and Kenny, Brian, "**Appraisal And Learning In A Government Agency**", The Learning Organization, 2001, Vol:8, Iss:1, 21 - 35
- Cynthia D. Fisher, Lyle F. Schoenfeldt and James B. Show, **Human Resource Management**, Boston: Houghton Mifflin Company Proud, 1996.
- Daley D., **Performance Appraisal in the Public Sector**, Westport, Ct.:Quorum Books, 1992
- David A Waldman; Leanne E Atwater, and David Antonio , 05/01/1998, Available On Site: http://www.zigonperf.com/resources/pmnews/360fb_amok.html (27.05.2006).
- DC Freedman, Hj Arnold,' **Managing Individual and Group Behavior in Organizations,** McGraw Hill Book Co., Auckland, 1983.
- Dessler, Garry., **Human Resource Management**, New Jersey: Prentice-Hall, 2000.
- Dicle, Ülkü.,**Yönetmelik Başarının Değerlendirilmesi Türkiye Uygulaması**, Ankara : ODÜ,İdari Bilimler Fakültesi,Yayın no.43, Ankara,1982.
- Dornbush , S. M. and W. R Scott, **Evaluation and the Exercise of Authority**, Jossey-Bass, 1975
- Erdoğan, Nihat and Beyaz Medihan, "**Başarı Değerlemede Atfetme Hatası ve Bir Araştırma**", Kocaeli Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, (3), 2002, 65-83
- F. J. Landy, And Farr, J. L. "**Performance Rating**", Psychological Bulletin, 87, 1980, 72-107

- Forest Jourden and Chip Heath, "The evaluation Gap in performance Perceptions: Illusory Ideas or groups and Individuals", *Journal of Applied Psychology*, 81 (4), 1996, 369-379
- Fox, S., & Dinur, Y., "Validity Of Self Assessment: A Field Of Evaluation", *Personnel Psychology*, 41, 1988, 141-56
- Gilbert, Charles A., 13 December 1994, "Performance Appraisals: A Process Analysis of Existing Methods and Trends", Available on Site:
<http://home.fuse.net/cgilbert/appraise.htm> (19.04.2006)
- Grote, D., *The Complete Guide to Performance Appraisal*, New York: American Management Association, 1996
- Harel, H. Gedaliahu, Tzafirir, S. Shay, "HRM Practices In The Public And Private Sectors: Differences and Similarities", *Public Administration Quarterly*, Fall 2001 / Winter 2002, Vol:25, Iss: 3 / 4, 316 - 356
- Hennessey H W Jr, Bernardin H John, "The Relationship Between Performance Appraisal Criterion Specificity And Statistical Evidence Of Discrimination", *Human Resource Management*, Summer 2003, Vol:42, Iss:2, 143 - 158
- Huber, V.L., "An Analysis of Performance Appraisal Practices in the Public Sector," *Public Personnel Management Journal*, November, 1983, 258-267
- Hurley, Simon, "Application of team-based 360° feedback systems", *Team Performance Management*, 4 (5), 1998, 202-210
- Jai, Ghorpade, 02.01.2000, "Five Paradoxes of 360 Degree Feedback", Available on Site:
<http://pages.stern.nyu.edu/~wstarbuc/mob/ghorpade.htm> (06.06.2006)
- Jansen E Pieter, "Performance Measurement In Governmental Organizations: A Contingent Approach To Measurement And Management Control", *Managerial Finance*, 2004, Vol:30, Iss:8, 54 - 68
- Jas Pauline and Skelcher Chris, "Performance Decline And Turnaround In Public Organizations: A Theoretical And Empirical Analysis", *British Journal Of Management*, September 2005, Vol:16, Iss:3, 195 - 211
- Jawahar I.M. And Williams C.s, "Where All the Children Are Above Average: The Performance Appraisal Purpose Affect", *Personnel Psychology*, 50, 1997, 905-925.
- Kane J.S. and Lawler, C.E., "Performance Appraisal Effectiveness. Its Assessment and Determinants", *Research in Organizational Behavior*, ed. B.M. Straw (Greenwich, Ct.: JAI Press, 1979).
- Kane JS, Bernardin JH, Hagan C, and Ross S, "The Effects Of A 360-Degree Appraisal System On Managerial Performance". In W.W. Tornow (Chair), *Upward feedback: The ups and downs of it*. Symposium, 10th Annual Conference, Society for Industrial and Organizational Psychology, Orlando, Florida, 1995
- Kasten, R. And Weintrau, Z., "Rating Errors and Rating Accuracy: A Field Experiment", *Human Performance*, 12 (2), 1999, 137-153

- Kaynak, Tuğray et al., **İnsan Kaynakları Yönetimi**, İstanbul : İ.Ü.İşletme Fakültesi, Yayın No:7, 2000.
- Keaney T. and McGann A., "A comparison of behavioral Expectation Scales and Graphic Rating Scales", *Journal of Applied Psychology*, 60, 1975, 695-703.
- Kluger, A. N., And Denisi, A., "The Effects Of Feedback Interventions On Performance: A Historical Review, A
- Kunda, Z.. "The Case For Motivated Reasoning", *Psychological Bulletin*, 108, 1990, 480-498.
- Landy, F.J. and Farr, J.L., **The Measurement of Work Performance: Methods, Theory and Applications**, N.Y.: Academic Press. 1983.
- Lefkowitz, J., "The Role of Interpersonal Affective Regard in Supervisory Performance Ratings: A Literature Review and Proposed Causal Model", *Journal of Occupational & Organizational Psychology*, 73, 2000, 67-85.
- Lepsinger, R. and Lucia D., "Creating Champions For 360 Degree Feedback", *Training and Development*, Feb, 52, 1998, 49-52
- Longenecker, C.O. and Ludwig, D.C. "Ethical Dilemmas in Performance Appraisal Revisited", *The Journal of Business Ethics*, Number 9, 1990, pp. 53-61
- Mani, G. Bonnie, "Performance Appraisal Systems, Productivity, And Motivation: A Case Study", *Public Personnel Management*, Summer 2002, Vol:31, Iss:2, 141 - 160
- Maroney, B., and Bucakley R., "Does Research In Performance Appraisal Influence The Practice Of Performance Appraisal? Regretfully Not". *Public Personnel Management*, 21, 1992, 185-196.
- Mathis, R.L. and Jackson, J.H., **Human Resource Management**, Min/St. Paul: West Publishing, 1994
- McEvoy, G.M., "Evaluating the Boss", *Personnel Administration*, September, 33 (9), 1988, 115-119
- Meta-Analysis, And A Preliminary Feedback Intervention Theory**" *Psychological Bulletin*, 119, 1996, 254-284.
- Meyer, H. H. , Kay, E., & French, R. P., "Split Roles in performance Appraisal" , *Harvard Business Review*, 65 (1), 1965, 123-9
- Miceli, M.P., I. Jung, J.P. Near, and D.B. Greenberger. "Predictors and Outcomes of Reactions to Pay-for-Performance Plans", *Journal of Applied Psychology*, 76, 1991, 508-521.
- Milkovich, G.M., & Boudreau, J.W, **Human Resource Management**, (8th ed.). Chicago: Irwin, 1997
- Miller, E. Corey, Thornton, L. Carl, "How Accurate Are Your Performance Appraisals", *Public Personnel Management*, Summer 2006, Vol:35, Iss:2, 153 - 163
- Miner, R. "Management Appraisal: A Capsule Review And Current References", *Business Horizon*, October, 1968, 83-99

- Mohrman, A.M., Resnick-West S.M. and Lawler, E.E. **Designing Performance Appraisal Systems**, San Francisco: Jossey-Bass, 1989
- Mohrman, A., Lawler, E and Resnick, S., "Performance Appraisal Revisited", *Organizational Dynamics*, 13 (2), 1984, 20-35,
- Murphy K.R. and Cleveland, J.N. **Performance Appraisal: An Organizational Perspective**, Boston: Allyn and Bacon, 1991
- Murphy, K. R., & Cleveland, J. N., **Understanding Performance Appraisal: Social, Organizational, and goal-based perspectives**, California: Sage, Thousand Oaks, 1995.
- Ndambakuwa Y. and Mufunda J., "Performance Appraisal System Impact on University Academic Staff Job Satisfaction And Productivity", *Performance Improvement Quarterly*, Spring 2006, Vol:19, Iss:1, 117 - 127
- Noonan, L., and Sulsky, L.M. "Examination Of Frame-Of-Reference And Behavioural Observation Training On Alternative Training Effectiveness Criteria In A Canadian Military Sample", *Human Performance*, 14, 2001. 3-26.
- Örücü and Mortaş, "Kamu Kurumlarında İş Analizi Çalışmaları - Muğla Üniversitesi Örneği", *Yönetim ve Ekonomi - Celal Bayar Üniversitesi İ.İ.B.F. Dergisi*, 2005, Cilt:12, Sayı:1, 25 -38
- Örücü and et al, "KİT'lerde Performans Değerleme Çalışmalarına Astların Destek ve Güveni - Muğla / Yatağan Termik Santrali Örneği", *Yönetim ve Ekonomi - Celal Bayar Üniversitesi İ.İ.B.F. Dergisi*, 2003, Cilt:10, Sayı:2, 51 - 62
- Palmer, J.K. and Feldman, Jack M, "Accountability and Need for Cognition Effects on Contrast, Halo, and Accuracy in Performance Ratings", *Journal of Psychology*, 139 (2), 2005, 119-137
- "Performance Appraisal", Available on Site: <http://www.performance-appraisal.com/basic.htm> (10.09.2006)
- Pulakos, Elaine D. and Wexley, Kenneth N. "The Relationship among Perceptual Similarity, Sex, and Performance Ratings in Manager-Subordinate Dyads", *Academy of Management Journal*, 26 (1), 1983, 129-139
- Redman and et al, "Performance Appraisal In An NHS Hospital", *Human Resource Management Journal*, 2000, 10 (1), 48 - 63
- Reilly, R.R. & McGourty, **Performance Appraisal In Team Setting**, San Francisco: Jossey-Bass, 1998.
- Roberts G.E., "The Influence of Participation, Goal Setting, Feedback and Acceptance on Measures of Performance Appraisal Effectiveness.", *Dissertation Abstracts International* (Doctoral dissertation, University of Pittsburgh, 1990
- Roch, S.G and O'Sullivan, B. J. "Frame Of Reference Rater Training Issues: Recall, Time And Behavior Observation Training". *International Journal of Training & Development*, 7(2), 2003, 93-107
- Schuler, R. S., **Managing Human Resource**, Min Ist. Paul: West Publishing, 1995

- Seyidođlu, Halil **Bilimsel Arařtırma ve Yazma el Kitabı**, Geliřtirilmiř 8 nci baskı, İstanbul, 2000
- Silverman S.B. and Wexley K.N, "**Employees Reaction to Performance Appraisal Inter-views as a Function of Their Participation in Rating Scale Development**," *Personnel Psychology*, 37, 1984, 703-710.
- Solomonson A. and Lance C., "**Examination of the Relationship between True Halo and Halo Effort in Performance Ratings**", *Journal of Academic Applied Psychology*, 82 (5), 1997, 665-674.
- Steers R.M. and Lee T.W., "**Facilitating Effective Performance Appraisals: The Role of Employee Commitment and Organizational Climate**," in *Performance Management and Theory*, eds. E.J. Landy and S. Zedeck, (1983), 75-85.
- Tak, Bilçin, "**Performans Deđerlendirme Sisteminin Toplam Kalite İlkeleri Açıřından Analizi ve Uyumlařtırma Önerileri**", *Marmara Üniversitesi Sosyal Bilimler Enstitüsü Hakemli Dergisi ÖNERİ*, 19 (5), 2003, 55 - 65.
- Tetlock, P. E., and Kim, J. "**Accountability And Judgment İn A Personality Prediction Task**". *Journal of Personality and Social Psychology: Attitudes and Social Cognition*, 52, 1987, 700-709.
- Thorton, G., III, "**Psychometric Properties of Self-Appraisal of Job Performance**," *Personal psychology*, 33 (2), 1980, 263-271.
- Tracey, W. R., **The Human Resources Glossary**, Washington DC.:St. Lucie Press, 1998.
- Tsui, A.S. and Barry, B. "**Interpersonal Affect And Rating Errors**", *Academy of Management Journal*, 29, 1986, 586-599.
- Tziner A. and Kopelman, R. E., "**Is there a Preferred Performance Rating Format? A Non-psychometric Perspective**", *Applied Psychology*, 51 (3), 2002, 479-503
- Uyargil Cavide, **İřletmelerde Performans Yönetimi Sistemi**, İstanbul: İ.Ü. İřletme Fakültesi Yayınları, 1994.
- Welbourne, T. M. & Johnson, D.E, & Amir, E., "**The Role-Based Performance Scale: Validity Analysis of A Theory-Based Measure**", *Academy of Management Journal*, 41, 1998, 540 - 555.
- Wexley, K. N., and G. P. Latham, **Developing And Training Human Resources İn Organizations**. Dallas TX:Scott, 1981
- Wiersma V.C. and Latham G.B. "**The Practicability Of Behavior Observation Scales, Behavioral Expectation Scales, And Trait Scales**", *Group And Organizational Management*, 20, 1986, 297-309
- Williams, R., **Performance Management**, Thomson Learning Press: London, 1988.
- Woods, H.Robert, Sciarini Michael and Breiter Deborah, "**Performance Appraisal In Hotels**", *Cornell Hotel and Restaurant Administration Quarterly*, 39 (2) 1998, 25-29.

136 Sait Gürbüz, Onur Dikmenli

Yalçın, Azmi, "Farklı Performans Değerleme Sistemlerine İlişkin İşgören Tercihlerinin Belirlenmesi Üzerine Bir Araştırma", Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 2002, Cilt:9, Sayı:9, 3 - 17

Yücel, Recep, "İnsan Kaynakları Yönetiminde Başarı Değerlendirme", Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 3 (1), 1999, 110 - 128

Zayonc, R.B. "Feeling And Thinking. Preferences Need No Inferences", American Psychologist, 35, 1980, 151-175.

Appendix-1

Demographics and Performance Appraisal Errors Scale Items Used As Measure

Part 1: Demographics

1. Your Status

a Managers b Employee

2. Number of years completed in the organization (tenure)

a 6 and below b Between 7 and 15 c 16 and above

3. Your age

a 26 and below b Between 27- 36 c 37-and above

4. Your Education level

a High school or below b Two-year college c Graduate (four year college) d MA-MBA (Master's) e Ph.D.

5. Your gender

a Male b Female

Part 2: Your Perceptions Concerning Performance Appraisal Errors

In this part, answer the following questions considering the behavior of your rater-superiors towards you, or other people in similar position to you, and your predictions and perceptions on how such behavior may be reflected on your appraisal report.

Questions 6, 7 measure the halo perceptions.

6. How many of your rater-superiors to date have the tendency to view affirmatively all behaviors or actions of a subordinate because the superior likes a particular behavior or action of the subordinate?

- a Don't know b None c Some d Most e All

7. How many of your rater-superiors to date have made clear by his words or acts that he places special emphasis on one or several behaviors or qualities specified in the evaluation form from the start of your working together?

- a Don't know b None c Some d Most e All

Questions 8 measure the horn effect

8. How many of your rater-superiors to date have the tendency to view negatively all behaviors or actions of a subordinate because the superior dislikes a particular behavior or action of the subordinate?

- a Don't know b None c Some d Most e All

Questions 9-10 measure the recency effect perceptions

9. How many of your rater-superiors to date do you think have been influenced by frequent display of behaviors they like and the lack of behaviors they dislike in the evaluation time?

- a Don't know b None c Some d Most e All

10. How many of your work mates have spared more effort in the evaluation time to display behaviors liked by the superiors and avoid errors disliked by the superiors?

- a Don't know b None c Some d Most e All

Question 11 measures the error of strictness perceptions

11. Some raters called "tight raters" have very high evaluation standards. For example they may say "the highest I would give is ninety (out of a hundred). How many of your rater-superiors would you qualify as "tight raters" to date?

- a Don't know b None c Some d Most e All

Question 12 measures the leniency error perceptions

12. Some raters called “loose raters” have the tendency to give top rating to all subordinates unless they have a clear deficiency, or they may not want to “adversely impact” the future of subordinates. How many of your rater-superiors would you qualify as “loose raters” to date?

- a Don't know b None c Some d Most e All

Questions 13-14 measure the similarity effect perceptions

13. How many of your rater-superiors have the tendency to give better rating to those subordinates similar to themselves in behavior, tastes and tendencies?

- a Don't know b None c Some d Most e All

14. How many of your work mates have made efforts to demonstrate that their behaviors, tastes and tendencies match those of the superior or hide those not matching with the superior's, with the intention that such would please the superior and consequently receive better ratings?

- a Don't know b None c Some d Most e All