Eskişehir Osmangazi Üniversitesi İİBF Dergisi
Aralık 2017, C. 12, S. 3, 49 - 64

Başvuru : 20.07.2017
Kabul : 20.09.2017

An Investigation on the Ratings from Four Sources for Different Positions in a 360 Degree Feedback System

Emin Kahya[1]
Fatih Çemrek[2]

360 Derece Performans Değerleme Sisteminde Farklı Pozisyonlar İçin Dört Kaynaktan Performanslar Üzerine Bir Araştırma

Öz

Günümüzde en popular değerlendirme sistemi olarak kabul edilen 360 derece performans değerleme sistemi, farklı kaynaklardan elde edilen değerlemelerin daha objektif ve kapsamlı olduğu görüşüne dayanmaktadır. Bu system, çalışana kendisini ve diğerlerini değerlendirme şansı tanıdığından, yaygın tercih edilmektedir. Çok kaynaklı değerleme sürecinde, dört kaynaktan (kendisi, yönetici, arkadaş ve ast) değerlemelerinde değerleyicinin aynı hoşgörüye sahip olup olmadığı araştırılmaktadır. Bu çalışmada, orta ölçekli bir mobilya işletmesinde, hem beyaz hem de mavi yakalı personeli kapsayan tüm görev pozisyonları için 360 derece performans değerleme sistemin kurulması ve uygulaması ele alınmıştır. Tüm fonksiyonel alanlardan ve pozisyonlardan toplam 200 çalışan (39 beyaz yakalı, 17 ustabaşı, 144 işçi) çalışmaya katılmıştır. Bulgular, değerlemelerde önyargı etkisinin, kendi, ast ve arkadaş değerlemelerinde yönetici değerlemesinden daha fazla olduğunu göstermiştir.

Anahtar Kelimeler: Çok Kaynaklı Değerleme Süreci, Performans Kriterleri, Iş Performansı, Deneyim, Eğitim.

An Investigation on the Ratings from Four Sources for Different Positions in a 360 Degree Feedback System

Abstract

The 360 degrees performance appraisal system, which is considered as the most popular evaluation system today, is based on the idea that the evaluations obtained from different sources are more objective and comprehensive. Since the system gives the employee a chance to evaluate him/herself and others, it is commonly preferred. It is investigated whether rater has a similar effect on the leniency of ratings from four sources (self, supervisor, peer, and subordinate) in multi-source assessment process. In this study, the establishment and implementation of a 360-degree performance appraisal system for all task positions including both white and blue-collar personnel in a medium-sized furniture business is discussed. A total of 200 employees (39 white collar employees, 17 foremen and 144 workers) from all functional areas and reputational roles participated in the study. The findings indicate that the influence of bias on ratings was significantly greater in self, subordinate and peer feedback than in supervisor feedback.

Keywords: Multi-Source Assessment Process, Performance Criteria, Job Performance, Experience, Education.

1. Introduction

Multi-source assessment (MSA) or 360 degree feedback process is used for assessing employee behaviors based on evaluations by two or more sources. For many organizations, MSA or the use of multiple raters to assess employee performance has become the cornerstone of the performance management process (Brutus et al., 2005). In 360 degree feedback process, employees receive ratings from four resources; they assess themselves, and they receive assessments from their supervisors, from their peers, and from their subordinates, if they are managers (Antonioni and Park, 2001). According to some estimates, MSA is used in more than 90% of Fortune 1000 companies and is continuing to spread among smaller firms as well (Coates, 1998).

[1] Prof. Dr., Eskişehir Osmangazi Üniversitesi Müh.ve Mim. Fak., Endüstri Müh. Bölümü, ekahya*@ogu.edu.tr*
[2] Doç. Dr., Eskişehir Osmangazi Üniversitesi Fen ve Edeb. Fak., İstatistik Bölümü, fcemrek@ogu.edu.tr

It is argued that these ratings should be used for development, rather than for evaluative purpose, although general evaluations are indirectly embedded in developmental feedback (e.g. Murphy and Cleveland, 1995; Waldman et al., 1998). In fact, their overwhelming use has been for employee development (Fletcher and Baldry, 1999). For example, employees are often required to present personal development plans which have to be met before the next administration of the 360 degree feedback (Beehr et al., 2001). The developmental use of 360 degree feedback is judged more positively by user than is administrative use (Bettenhausen and Fedor, 1997). Organizations primarily use 360 degree feedback for developmental purposes, to provide information to ratees about how raters perceive their leadership and work behaviors (Antonioni and Park, 2001).

In spite of the popularity of 360 degree feedback process and recent research on it, much is still unknown about ratings used and their relationships to other important work-related variables (Fletcher and Baldry, 1999). Researchers have suggested that the advantages of using multiple raters include the ability to observe and rate various job facets of each ratee's performance (Borman, 1974), greater reliability, enhanced fairness, and increased rate acceptance (Antonioni and Park, 2001). Previous empirical research has addressed the benefits of 360 degree feedback (London and Beatty, 1993; Tornow, 1993), the benefits of peer and upward appraisals (Bettenhausen and Fedor, 1997), and the extent of self-other agreement in ratings (Atwater et al., 1998). However, several researchers have argued that research on 360 degree feedback has not kept pace with the practice and that there are insufficient research models and data available to guide organizations through the use of this type of feedback (Antonioni and Park, 2001). There is a risk that specific rater characteristics may influence ratings. This oversight is cause for some concern because 360 degree feedback programs depend on the quality of ratings from multiple sources (Antonioni and Woehr, 2000).

As long as employers continue to rely on rating instruments to evaluate the performance of employees, the quality of ratings will be of continuing interest to both managers and researchers (Tsui and Barry, 1986). Previous studies have demonstrated that performance ratings are influenced by various factors such as rater and ratee demographic characteristics (age, gender, race, education level, job experience) (e.g., Schmidt et al., 1986; Schmitt and Robertson, 1990; Sundvik and Lindeman, 1998; Ferris et al., 2001), cognitive process (e.g., DeNisi and Williams, 1988; Robbins and DeNisi, 1994; Feldman, 1981; DeNisi et al., 1984; Robbins and DeNisi, 1998; Murphy and Cleveland, 1991), and interpersonal affect (e.g., friendship, liking) (Varma et al., 2005). There exist an extensive literature on the relation between rater's bias and ratings. The evidence suggests that interpersonal affect and opportunity to observe cerate bias in ratings assigned in a performance evaluation (e.g., Tsui and Barry, 1986; Robbins and DeNisi, 1998; Antonioni and Park, 2001; Varma et al., 2005). Interpersonal affect is defined as a "like-dislike relationship" between a supervisor and his/her subordinate, and has been shown to occur very early in stimulus observation and performance evaluation (Cardy and Dobbins, 1986). Liking is an emotional reaction (positive, neutral, or negative) to a specific person (Zajonc, 1980). In other words, if a supervisor likes his/her subordinate, s/he is deemed to have a high bias towards that subordinate. In this connection, research has consistently indicated that rater's interpersonal affect towards a ratee is difficult to separate from performance information when assigning ratings (Cardy and Dobbins, 1986; Robbins and DeNisi, 1994). DeNisi et al. (1984) suggest that a rater's consideration of (i) the purposes and consequences of an appraisal, and (ii) the ratee's awareness of ratings may influence the assignment of ratings at the last stage in the

evaluation process. Interpersonal affect may be a basis for a rater's attempt to preserve friendship in situations where appraisals will be used for promotions and rewards. Although many studies have examined the role of undifferentiated affect in performance evaluations, others have looked at the role of differentiated affect, or liking for another individual, in the appraisal process (e.g. Cardy and Dobbins, 1986; Varma et al., 1996). These reviews have concluded that interpersonal affect has an influence on performance ratings, but the mechanism for this influence is not clear (Varma et al., 2005).

The present study investigates the influence of raters' bias towards ratees on ratings from 360 degree feedback. The literature on the role of this topic has focused on performance appraisals from a single resource, primarily traditional downward or peer appraisals. The primary purpose of the present study was to support and extend previous research by focusing on the extent to which the joint role of performance evaluations was tested in a wider range of demographic characteristics, rater resources, performance measures used for assessing employees, and reputational roles in organization.

It has been suggested that individual characteristics may influence ratings. Therefore, the experience and education level of the ratee and rater will be used as controls to better ascertain the incremental relationship between rater's position and ratings in this study. The age and gender of the ratee and rater were not considered because of weak on performance ratings of these variables. Most studies on this topic have focused on ratings from a single source, and only one employee position. It remains to be seen valuable whether performance ratings from multi-sources are influenced by the same factors. We investigated the effect of positon on ratings from four rater resources (shelf, peer, supervisor and subordinate).

In this study, the establishment and implementation of a 360 degree feedback system for all task positions involving both white and blue-collar personnel in a medium-sized furniture business was discussed. The reputational roles in manufacturing organizations vary from worker (blue collar employee) to director (department manager). The present study was designed to reveal how employee position might be related to bias in ratings. Three employee groups were (i) blue collar employees who are workers, (ii) foremen who are the first-line managers of the workers, (iii) white collar employees including the officers, sub-department managers-chiefs, and department managers. There is evidence that bias is associated with several dimensions of contextual performance. The present study is an attempt to reveal whether performance categories (interpersonal citizenship, organizational citizenship and job dedication) used to measure employee performance are associated with differential effects on ratings.

2. 360 Degree Feedback Process

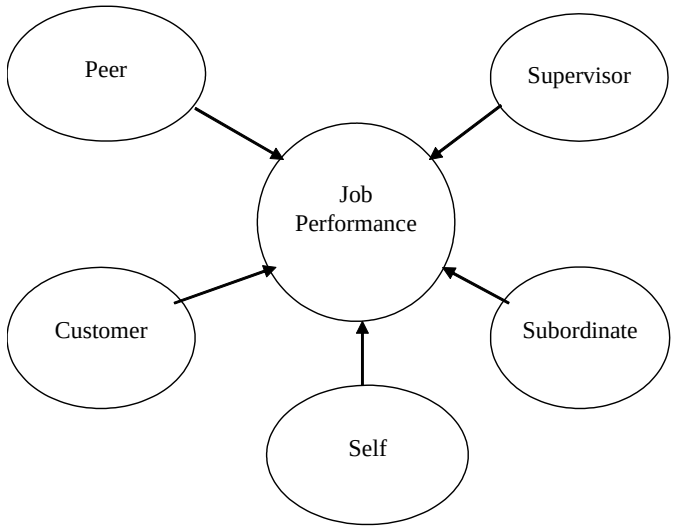
Performance is a term indicating how far away is a person, group or organization from target point in a certain period or in unit time. In other words, it states what they can provide as qualitative and quantitative (Akal, 2005). The future expectation, sense of duty, work discipline, ability and skill level of each employee are different from each other. These differences, starting from human nature, also make different in one's success on the job. While some employees fulfill the task expected from them completely, some can not show the expected success. The degree of success in enterprises can be determined by performance evaluation.

Performance evaluation is objective analysis and synthesis to determine how well the skills of the staff fit in with the qualities and requirements of the job or how well they perform their expected tasks (Sabuncuoğlu, 2000). Performance evaluation is one of the important functions

of human resources management and it is used by individuals in the direction of organizational goals and in the analysis of the results they produce in a certain period and in various fields (such as wage, promotion, etc.) (Akdemir, 2009). One of the most recent and popular approaches to performance evaluation is the use of multi-source performance evaluation and feedback. A 360 Degree Performance Appraisal System has been required due to a large number of employee in organizations and providing a more comprehensive and accurate feedback in line with different perspectives on employees (Uygur and Sümerli Sarıgül, 2015). The 360 Degree Performance Appraisal System aims to interrogate a multi-dimensional and continuous understanding within the performance evaluation methodology and it is a system that is assessed by the supervisors as well as the evaluate him/herself (self-evaluation), his/her colleagues (peers), subordinates and customers in the business line, and provides feedback on performance (Barutçugil, 2002).

In the appraisal of performance, according to traditional approaches, it is argued that only the supervisors can appraise subordinates. In practice, however, supervisors are the least qualified persons for appraising the key points of the individual's performance. The 360 degree appraisal system is a mixed evaluation approach in which, unlike traditional performance appraisal methods, a large number of people and measures are used instead of managers only for an employee's behaviors. It is the participation of managers, peers, internal and external customers (especially in service systems), lower level staff (subordinate) and self in evaluating (Figure 1).

Figure 1: 360 Degrees Appraisal System



360 Degree Appraisal is named in various forms such as; "Multi-Source Evaluation", "360 Degree Performance Appraisal and Feedback", "360 Degree Feedback". According to a survey on the prevalence of 360 degree appraisal, 40% of enterprises in practice use this method (Antonioni and Park, 2001).

360 degree performance appraisal based on different rater resources has some certain advantages. These advantages can include:

a) The ability of each rater resource to present information about ratee from his own perspective,

b) Multiple ratings have a higher validity than the information from a single source,

c) Providing the requisite behavioral change by increasing the self-awareness of ratee with the feedback obtained from multiple sources,

d) The ability to see how employees are perceived not only by their supervisors but also by their peers and subordinates (Baltacı and Burgazoğlu, 2014).

3. Method

3.1. Participants

This study was conducted at a medium sized furniture company with a total of 200 employees (39 white collar employees, 17 foremen, each of them was the first manager of a manufacturing line, and 144 blue collar employees-workers) from all functional. All the employees were participated. All participants received a two-hours workplace training program designed to improve appraisal skills in large groups. Such a program includes principles of assessment such as reliability, validity, and fairness, definitions of the criteria used to evaluate employee performance, and halo errors in ratings.

3.2. Performance Criteria

The performance assessed in this study was forms of contextual performance. Contextual performance is defined as individual efforts that are not directly related to their main task function but are important because they shape the organizational, social, and psychological context that serves as the critical catalyst for task activities and processes (Werner, 2000). Although task performance traditionally has received more attention than contextual performance, researchers have begun to empirically demonstrate that contextual performance yields a competitive advantage for organizations (Witt, et al., 2002). It has been suggested that contextual performance benefits organizations in several ways (Van Scotter, 2000). Contextual performance behaviors involving persistence, effort, compliance, and self-discipline are expected to increase the effectiveness of individual workers and managers (Motowidlo et al., 1997). Helpful, considerate, and cooperative behaviors are expected to increase work group effectiveness and improve organizational coordination and control by reducing friction among organizational members and promoting a social and psychological context that facilitates task performance (Borman and Motowidlo, 1993). Innovative and voluntary behaviors enhance an organization's ability to solve unanticipated problems and adapt to change. In the aggregate, these behaviors are expected to improve organizational efficiency by freeing up resources that would otherwise be needed to handle disciplinary problems, solve communication difficulties, resolve conflicting demands, or provide closer monitoring of employee performance (Motowidlo et al., 1997).

Twenty-five contextual performance criteria were generated from previously cited literature (Borman and Motowidlo, 1993; Goodman and Svyantek, 1999; Coloman and Borman, 2000; Van Scotter, 2000; Moorman and Wells, 2003) and performance evaluation systems implemented by the companies. Four performance criteria also were suggested by the evaluation committee, consisted of the department managers. They selected 22 of 29 criteria to measure job performance, which were important and linked to the company's vision and values. They were categorized in three groups as settled by Coleman and Borman (2000) (see Appendix 1).

3.3. Control Variables

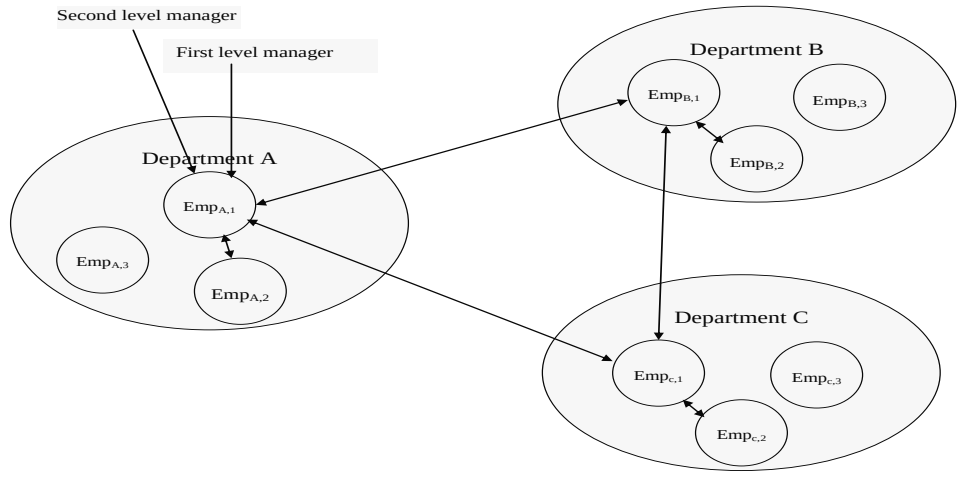
The studies cited so far have examined the effects of various ratee characteristics on ratings. Raters tend to favor younger employees over older ones. Moreover, men have been shown to give more positive ratings than women. Major findings were that ratee age, gender, and race have little effect on ratings. Borman et al. (1991) stated that ratee race and gender have been shown to account for from less than 1% to a maximum of about 5% of the variance in ratings. In addition, ratee age does not appear to be significant factor in ratings. Education level and job experience as employee's demographic information might explain some of the variance in performance ratings. Job experience (number of years that employee is working in the company) and education level of each ratee and rater were obtained from the company's archival sources.

- Education level was a grouped measure that was coded as 1: High school under, 2: High school, 3: Professional school, 4: Associate degree program, 5: University, 6: Master degree.
- Job position level was a grouped measure that was coded as 1: Director, 2: Chief, 3: Other white collar employee, 4: Foreman, each of them was the first manager of a manufacturing line, and 5: Blue collar employees-workers
- Rater resource was a grouped measure that was coded as 1: Self, 2: Peer, 3: First level manager, 4: Second level manager, 5: Subordinate.

3.4. Procedure

This study was part of a larger HRM project. In the present 360 degree feedback process, employees received ratings from four different sources (shelf, peer, manager and subordinate); one employee received assessments from two peers (from workers) or three peers, from the first and also second level managers and from subordinates. Peers for each employee were randomly selected by the author (Figure 2). To illustrate at this point, one employee assessed one peer from the same department and two peers from the nearest departments.

Figure 2: Evaluating Model



Although some studies have showed that differences in rating format have only minimal effects on the quality of ratings (Landy and Farr, 1980), different formats involve different psychological process that may influence the relationship between individual differences or contextual variables and ratings (Yun et al., 2005). The graphic rating scale format is widely used by organizations for performance rating purposes. It consists of a definition of performance with a bipolar adjective scale. Raters use a five-point scale ranging from 1="fails to meet expectations" to 5="clearly and consistently exceeds expectations" to rate ratee's performance. However, these concise scale definitions are open to interpretation and not to permit accurate assessments. A defining desirable behaviors (specific behavioral anchors) in depth for each scale on each criterion reduces idiosyncratic variations, increases the likelihood of understanding the scale, and then such a format yields more consistent ratings. An evaluation form including employee (name, id. no), and job (position, name, department name etc.) characteristics, and scales of the criteria to check the appropriate one was designed.

Assessments were administrated during small group meetings. Forms per employee and also one guide explaining desirable behaviors for each scale of the criteria were given to each rater. Participants were guaranteed that all responses would remain strictly confidential, and also informed that participation in this study was voluntary. The raters from different organizational positions and sources used different criteria in making performance judgements. To illustrate at this point, workers assessed their own co-workers with 11 criteria though they received assessments with 16 criteria from their first and also second level managers. The final performance score was calculated by averaging the ratings to each of the criteria used. Complete data existed for 1176 ratings such as

- 199 self ratings
- 446 peer ratings
- 363 manager ratings
- 168 subordinate ratings

4. Results

4.1. Ratings

Table 1 reveals descriptive statistics. All the directors and department chiefs have university degree. All white collar employee have high school and upper degree. However, foremen who are the first supervisor in manufacturing lines have under high school and professional school degree. Job experiences are ranging from 0.50 to 15 years.

Table 1: Descriptive Statistics

Position	Education Level						Sex		
	High school under	High school	Professional school	Associate degree program	University	Master degree	Male	Female	Experience
Director (n=4)					4		3	1	11.80
Chief (n=13)			3		10		9	4	7,68
Employee (n=21)		5	5	5	5	1	8	13	5,01
Foreman (n=17)	13			4			13	-	8,45
Worker (n=144)	118	4	21	1			144	-	6,00

The performance scores obtained from each source in the rated (position) category are given in Table 2.

Table 2: Performance Ratings

Position (Ratee)	Resource				
	Self	Peer	Fist Manager	Second Manager	Subordinate
Director (n=4)	3.93	3.86	3.39	-	3.52
Chief (n=13)	3.48	3.03	3.65	3.03	3.52
Employee (n=21)	3.47	3.23	3.13	3.31	-
Foreman (n=17)	3.36	3.29	3.10	3.37	3.66
Worker (n=144)	3.98	3.47	3.00	2.96	-

When examined the director ratings, self-assessment and peer evaluations are around 3.9, while general director rating is 3.39. While the directors find themselves very successful, the general director is not of the same opinion. The ratings of chiefs are not too high, 3.52. The chiefs are not in a position to protect their managers.

While in chief ratings, self-evaluation, manager evaluation and office staff (sub) evaluations are close to each other and around 3.5, peer evaluations are very low (3.03). The chefs are either not aware of each other's performances or are in a harsh (negative) attitude (prejudice) towards each other. Another remarkable result is about the general director ratings on chiefs. The rates as second supervisor are very low. This shows that second-level supervisors do not have much knowledge on the performance of employees. If the supervisor does not know the employee's performance very well, it is expected that he will tend to appreciate his performance as "moderate".

Employees rated themselves at 3.47, while the other raters rated them at 3.13-3.31. These staff tend to score high for themselves. In other words, employees do not know themselves enough.

Foremen's ratings, when compared to others, are more objective. While self-evaluation, peer rating and manager rating are very close to each other (3.30-3.40), chief ratings are slightly lower, but ratings of workers are higher (3.66). Workers are in the instinct of protecting their foremen.

In worker ratings, self-evaluation scores are very high. Workers are not objective towards themselves. When performance evaluation forms are examined, it has been found that many workers give themselves "perfect" (5) for each criterion. Due to this non-objective attitude, self-evaluation has been removed for workers in subsequent periods. The workers are rated as "moderate" (3) by both the foremen and the chiefs.

It is determined as the result of SPSS 22 package software, "One-way ANOVA: Post Hoc Multiple Comparison" (Analyze/Compare Means) Tukey test analyzes, with 95% confidence level;

a) There are no significant differences among raters in manager ratings (the significance level is between 0.25 and 0.90),
b) There is no significant difference among raters, self evaluation, manager (firs supervisor) and clerk (sub) ratings alongside peer and general director (second supervisor) ratings (Significant difference occured between these two),
c) There are no significant differences among raters in white collar personnel ratings,
d) There are significant differences between peer ratings and chief (first supervisor) and worker (sub) ratings in foreman ratings,
e) In workers evaluations, there is no difference between first (foreman) and second (chef) supervisor ratings, and there are significant differences between other raters (Table 3.a).

In addition, the same analysis was done for all positions and also for self-assessment (Table 3.b). It has been determined that there is a significant difference between the self-evaluations of the workers and the personnel in all other categories. This means that workers make the most optimistic assessment of themselves, in comparison with all other employees.

4.2. Correlations Among All The Variables

Pearson correlation coefficients of the variables are depicted in Table 4. Since the possibility of prejudgement, seperated correlations for white-collar staff (Table 4.a), foremen (Table 4.b) and workers (Table 4.c) were generated only through supervisor ratings. Education level, experience, three performance categories and job scores were used as variables.

While there was a positive relationship between experience and education level for foremen (r=0.317), there was a negative relationship between white-collar staff (r=-0.404) and workers (r=-0.201). This shows that the vast majority of experienced personnel have a lower level of education.

In the ratings of white-collar personnel and foremen, in general, the level of education increases the performance values (r=0.274 ve r=0.176). However, the same result has not been achieved for blue-collar workers (r=-0.025), and it can not be said that the education has improved performance. On the other hand, educated white-collar workers have been observed to have lower performance (r=0.169) in terms of organizational citizenship criteria. However, organizational citizenship performance criteria are lower for foremen (r=0.034). There is no consistency in the performance rates of workers.

Concerning experience, the results are thought-provoking. Similar results were obtained for the staff in all three categories. As the experience of the staff increases, the performance results in all three categories are negative both for the white-collar (r=-0.041) and foremen (r=0.278) and for workers (-0.062). Namely, the performance decreases as the experience increases

Table 3: One-way ANOVA (Tukey) Testi

a. Workers

(I)	(J)	Mean Difference (I-J)	Std. Error	Sig.
Self	Peer	,515904*	,053079	,000
	Foreman	,972870*	,061761	,000
	Chief	1,019358*	,062336	,000
Peer	Self	-,515904*	,053079	,000
	Foreman	,456965*	,053702	,000
	Chief	,503453*	,054362	,000
Foreman	Self	-,972870*	,061761	,000
	Peer	-,456965*	,053702	,000
	Chief	,046488	,062866	,881
Chief	Self	-1,019358*	,062336	,000
	Peer	-,503453*	,054362	,000
	Foreman	-,046488	,062866	,881

* The mean difference is significant at the 0.05 level

b. Self Evaluation

(I)	(J)	Mean Difference (I-J)	Std. Error	Sig.
Director	Chief	,51026	,29981	,435
	Employee	,46580	,28890	,491
	Foreman	,57222	,29604	,303
	Worker	-,04840	,27301	1,000
Chief	Director	-,51026	,29981	,435
	Employee	-,04446	,16519	,999
	Foreman	,06197	,17737	,997
	Worker	-,55866*	,13548	,001
Employee	Director	-,46580	,28890	,491
	Chief	,04446	,16519	,999
	Foreman	,10642	,15824	,962
	Worker	-,51420*	,10924	,000
Foreman	Director	-,57222	,29604	,303
	Chief	-,06197	,17737	,997
	Employee	-,10642	,15824	,962
	Worker	-,62062*	,12691	,000
Worker	Director	,04840	,27301	1,000
	Chief	,55866*	,13548	,001
	Employee	,51420*	,10924	,000
	Foreman	,62062*	,12691	,000

*The mean difference is significant at the 0.05 level.

Table 4: Pearson Correlation Coefficients

a. White collar employees (n=57)

	1	2	3	4	5	6	7
1. Rater's experience	-						
2. Rater's education level	0.109						
3. Ratee's experience	0.182	0.244					
4. Ratee's education level	0.258	-0.125	-0.404**				
5. Inter. citizenship	0.114	-0.099	-0.076	0.253			
6. Org. citizenship	-0.108	-0.200	-0.132	0.169	0.541**		
7. Job dedication	-0.088	0.011	0.031	0.259	0.627	0.689**	
8. Job performance	-0.040	-0.079	-0.041	0.274*	0.808**	0.835**	0.934**

Notes : ** p<0.01 (2-tailed), * p<0.05 (2-tailed)

b. Foremen (n=29)

	1	2	3	4	5	6	7
1. Rater's experience	-						
2. Rater's education level	-0.466*						
3. Ratee's experience	0.064	-0.130					
4. Ratee's education level	0.051	-0.144	0.317				
5. Inter. citizenship	-0.115	0.371*	-0.369	0.034			
6. Org. citizenship	0.100	0.066	-0.145	0.161	0.358		
7. Job dedication	0.147	0.160	-0.271	0.254	0.725**	0.679**	
8. Job performance	0.114	0.150	-0.278	0.176	0.707**	0.883**	0.908**

Notes : ** p<0.01 (2-tailed), * p<0.05 (2-tailed)

c. Blue collar employees (Workers) (n=277)

	1	2	3	4	5	6	7
1. Rater's experience	-						
2. Rater's education level	-0.509**						
3. Ratee's experience	0.017	0.000					
4. Ratee's education level	0.225**	0.009	-0.201**				
5. Interpersonal citizenship	0.088	-0.167*	-0.026	0.044			
6. Organizational citizenship	-0.208**	-0.147*	-0.061	-0.030	0.405**		
7. Job dedication	-0.147*	-0.058	-0.065	-0.077	0.672**	0.674**	
8. Job performance	-0.129*	-0.159*	-0.062	-0.025	0.759**	0.880**	0.892**

*Notes : ** p<0.01 (2-tailed), * p<0.05 (2-tailed)*

5. Conclusion

The 360 degree performance appraisal method is based on the principle that the rater person is appraised not only by the managers as in the classical methods, but also by his peers, subordinates and him/herself. In general, the basic idea underlying this method is that the results will be more objective and comprehensive with the ratings to be taken from different sources instead of the single source (manager). It is aimed to reduce the errors in the rating

system and achieve more accurate results by increasing the acceptability of the evaluation system by ensuring employee participation from all levels.

However, this system has some disadvantages. Most importantly, there are different outcomes among ratings from different sources. Especially, the most common situation is that there are differences between self-evaluations and other raters. This situation has shown itself in white-collar ratings and even in manager ratings. While the general director's score was 3.39, the directors gave more credit themselves with a score of 3.93. This is more grave in the workers' results. While the foremen rated the workers at 3.00, the workers rated themselves at 3.98 score.

In performance appraisal, supervisors are considered to have a) continuous communication and physical proximity with employees, b) knowledge and experience and superior qualifications than employees. If these two features do not take place, the errors in the ratings will appear. There is high communication and physical proximity between department chiefs and directors and employees and foremen. Therefore, except for the protection attitude, it is believed that the performance appraisal results of first level managers of employees are more reliable. As a matter of fact, it is clear that the second-level managers appraised the white-collar staff and foremen with lower or higher ratings.

In performance appraisal literature, individual characteristics (such as age, gender, experience), observation time, interpersonal affect, rating format, workplace deviant behaviors have been considered in many studies. The argument for distinguishing between task and contextual performance gains force if they are correlated with different demographic characteristics. Borman and Motowidlo (1993) suggested that the major source of variation in job performance is the proficiency with which a person can carry out task activities. This means that individual differences in knowledge, skills, and abilities should covary more with job performance. Experience should be more strongly correlated with job performance. In Van Scotter and colleagues' studies, experience has significantly correlated with job performance (r=0.30-0.40). Although our results (r=-0.041 for white collar employee) differed from their findings, they not only support the above sight but also overlap too much with Moser et al. (1999) and Baltacı and Burgazoğlu (2014)'s results. However, experienced employees may, generally, get difficulty adjusting to social or new situations or engaging in self-development to improve own effectiveness.

To our knowledge, this study is the first investigation to indicate the influence of four different rater categories. The participants were one general director as only a rater, 4 directors, 13 chiefs, 21 employees and 17 foremen who are the first level managers of the manufacturing lines and 144 blue collar workers. The general director rated the directors and their chiefs. The directors, chiefs, employees, foremen and workers rated to each other (self, downward, peer and upward ratings). In self-evaluation, while the foremen, employees and chiefs rated with 3.4 - 3.5 score, workers appreciated 3.98. Both level managers, however, believe that workers have a "moderate" (around 3) performance. This clearly shows that workers are protecting themselves and even their friends and making prejudiced rating. A similar situation is seen also in the workers' rating on their foremen. While rating foremen, it is clear that workers rate taking into consideration their personal relationships with them rather than the aims of workplace or in such a way that they do not conflict with the manager and not draw his reaction. For this reason, it is not advisable to operate a 360-degrees performance appraisal system for workers, unless they are adequately trained in objective rating.

In order to generalize the findings to other task environments, further research should seek to define the conditions that reduce or enhance the impact of job performance across a range of different occupations. In this study, it is not intended to drive the job performance appraisal criteria to be used by the company to assess employees. They were designed to measure the effectiveness of the employee. As a next research, the other appropriate job performance criteria should be extended.

References

Akal, Z. (2005), *İşletmelerde Performans Ölçüm ve Denetimi: Çok Yönlü Performans Göstergeleri,* 6.Baskı, Milli Prodüktivite Merkezi Yayınları No: 473. Ankara.

Akdemir, A. (2009), *İşletmeciliğin Temel Bilgileri*, Ekin Yayınları, Bursa.

Antonioni, D. ; H. Park (2001), "The relationship between rater affect and three sources of 360-degree feedback ratings", *Journal of Management*, 27: 479-495.

Antonioni, D. ; D.J. Woehr (2000), "Improving the quality of multi-source rater performance", In D.W. Bracken, C.W. Timmreck, and A.H. Church (Eds.). *Handbook of Multisource Feedback* (pp. 114-129). San Francisco: Jossey-Bass.

Atwater, L.E.; C. Ostroff; F.J. Yammarino; J.W. Fleenor (1998), "Self-other agreement: Does it really matter?", *Personnel Psychology*, 51 : 577-598.

Baltacı, A. İ.; H.Burgazoğlu (2014), "Değerlendiriciler Arası Güvenilirlik ve Tatmin Bağlamında 360 Derece Performans Değerlendirme", *Marmara Üniversitesi Öneri Dergisi*, 11 (41): 57-76.

Barutçugil, İ. (2002), *Performans Yönetimi,* 2. Basım, Kariyer Yayıncılık, İstanbul.

Beehr, T.A.; L. Ivanitskaya; C.P. Hansen; D. Erofeev; D. Gudanowski (2001), "Evaluation of 360 degree feedback ratings: Relationships with each other and with performance and selection predictors", *Journal of Organizational Behavior*, 22: 775-788.

Bettenhausen, K.L.; D.B. Fedor (1997), "Peer and upward appraisals: A comparison of their benefits and problems", *Group and Organizational Management*, 22: 236-263.

Borman, W.C. (1974), "The rating of individuals in organizations: An alternative approach", *Organizational Behavior and Human Performance*, 12: 105-124.

Borman, W.C.; S.J. Motowidlo (1993), "Expanding the criterion domain to include elements of contextual performance", In N.Schmitt and W.C. Borman (Eds), *Personnel Selection in Organizations* (pp. 71-98), New York, Jossey-Bass.

Borman, W.C.; L.A. White; E.D. Pulakos; S.H. Oppler (1991), "Models of supervisory job performance ratings", *Journal of Applied Psychology*, 76 (6): 863-872.

Brutus, S.; S. Petosa; E. Aucoin (2005), "Who will evaluate me? Rater selection in multi-source assessment contexts", *International Journal of Selection and Assessment*, 13 (2): 129-138.

Cardy, R.L. and G.H. Dobbins (1986), "Affect and appraisal accuracy: Liking as an integral dimension in evaluating performance", *Journal of Applied Psychology*, 71: 672-678.

Coates, D.E. (1998), "Don't tie 360 feedback to pay", *Training*, 35: 68-78.

Coleman, V.I. ; W.C. Borman (2000), "Investigating the underlying structure of the citizenship performance domain", *Human Resource Management Review*, 10 (1) : 25-44.

Decotiis, T. ; A.Petit (1978), "The performance appraisal process: A model and some testable propositions", *Academy of Management Review*, 3: 635-646.

DeNisi, A.S.; T.P. Cafferty; B.M. Meglino (1984), "A cognitive view of the performance appraisal process: a model and research propositions", *Organizational Behavior and Human Performance*, 33 (3): 360-396.

DeNisi, A.S. ; K.J. Williams (1988), "Cognitive approaches to performance appraisal", *Research in Personnel and Human Resources Management*, 6: 109-155.

Feldman, J.M. (1981), "Beyond attribution theory: cognitive process in performance appraisal", *Journal of Applied Psychology*, 66: 863-872.

Ferris, G.R.; T.A. Judge; K.M. Rowland; D.E. Fitzgibbons (1994), "Subordinate influence and the performance evaluation process: Test of a model", *Organizational Behavior and Human Decision Process*, 58: 101-135.

Fletcher, C.; C.Baldry (1999), "Multi-source feedback systems: A research perspective", *International Review of Industrial and Organizational Psychology*, 14: 149-193.

Goodman, S.A. ; D.J. Svyantek (1999), "Person-organization fit and contextual performance: Do shared values matter", *Journal of Vocational Behavior*, 55: 254-275.

Landy, F.J.; J.L. Farr (1980), "Performance ratings", *Psychological Bulletin*, 87: 72-107.

London, M.; R.W. Beatty (1993), "360-degree feedback as a competitive advantage", *Human Resource Management*, 32: 353-372.

Moorman, R.H.; D.L. Wells (2003), "Can electronic performance monitoring be fair? Exploring relationship among monitoring characteristics, perceived fairness, and job performance", *Journal of Leadership and Organizational Studies*, 20 (2): 2-16.

Moser, K.; H. Schuler; U.Funke (1999), "The moderating effect of raters' opportunities to observe ratees' job performance on the validity of an assessment centre", *International Journal of Selection and Assessment*, 7 (3): 133-141.

Motowidlo, S.J.; W.C. Borman; M.J. Schmit (1997), "A theory of individual differences in task and contextual performance", *Human Performance*, 10 (2): 71-83.

Murphy, K.R.; J.N. Cleveland (1991), *Performance appraisal: An organizational perspective.* Englewood Cliffs, NJ: Prentice Hall.

Murphy, K.R. ; J.N. Cleveland (1995), *Understanding performance appraisal: Social organizational, and goal-based perspectives.* Sage: Thousand Oaks, CA.

Robbins, T.L.; A.S. DeNisi (1994), "A closer look at interpersonal affect as a distinct influence on cognitive processing in performance evaluations", *Journal of Applied Psychology*, 79: 341-353.

Robbins, T.L.; A.S. DeNisi (1998), "Mood vs. interpersonal affect: Identifying process and rating distortions in performance appraisal", *Journal of Business and Psychology*, 12 (3): 313-325.

Sabuncuoğlu, Z. (2000), *İnsan Kaynakları Yönetimi***,** Ezgi Kitabevi, Bursa.

Schmitt, N. ; I. Robertson (1990), "Personnel selection", *Annual Review of Psychology*, 41: 289-319.

Sundvik, L.; M. Lindeman (1998); "Performance rating accuracy: Convergence between supervisor assessment and sales productivity", *International Journal of Selection and Assessment*, 6 (1): 9-15.

Tornow, W.W. (1993), "Editor's note: Introduction to special issue on 360-degree feedback", *Human Resource Management*, 32: 211-220.

Tsui, A.S.; B.Barry (1986), "Interpersonal affect and rating errors", *Academy of Management Journal*, 29 (3): 586-599.

Uygur, A.; S. Sümerli Sarıgül (2015), "360 Derece Performans Değerleme ve Geri Bildirim Sistemi", *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 33: 189-201.

Van Scotter, J.R. (2000), "Relationships of task performance and contextual performance with turnover, job satisfaction, and affective commitment", *Human Resource Management Review*, 10 (1): 79-95.

Waldman, D.A.; L.E. Atwater; D. Antonioni (1998), "Has 360-degree feedback gone amok?" *Academy of management Executive*, 12: 86-94.

Varma, A.; S.Pichler; E.S. Srinivas (2005), "The role of interpersonal affect in performance appraisal : Evidence from two samples – the US and India", *International Journal of Human Resource Management*, 16(11): 2029-2044.

Werner, J.M. (2000), "Implications of OCB and contextual performance for human resource management", *Human Resource Management Review*, 10 (1): 3-24.

Witt, L.A.; K.M.Kacmar; D.S. Carlson; S. Zivnuska (2002), "Interactive effects of personality and organizational politics on contextual performance", *Journal of Organizational Behavior*, 23: 911-926.

Yun, G.J., L.M.Donahue; N.M. Dudley; L.A. McFarland (2005), "Rater personality, rating format, and social context: Implications for performance appraisal ratings", *International Journal of Selection and Assessment*, 13 (2): 97-107.

Zajonc, R.B. (1980), "Feeling and thinking: Preferences need to inferences", *American Psychologist*, 35: 151-175.

Appendix 1: Performance Criteria

Criterion	Definitions
A. Interpersonal Citizenship	
A.1. Altruism	
Communication	Communication co-workers with personal matters
Supporting	Supporting and encouraging a coworker with a problem
A.2. Conscientiousness	
Cooperating	Cooperating with others to solve problems
Leadership	Leadership
*Group activity**	*Engaging responsibly in meetings and group activities*
B. Organizational Citizenship	
B.1. Allegiance/Loyalty	
Respect	Treatment the supervisor with respect
Absenteeism	Absenteeism
Arriving on time	Exhibiting punctuality arriving at work on time in the morning and after lunch breaks
Working systematically	Working systematically
Not complaining	Not complaining about organizational conditions
*Meeting**	*Participating in training meeting*
*Engaging**	*Keeping others engaged in individual problems*
*Health**	*Health condition for the job*
B.2. Compliance	
Bearing	Display proper appearance or bearing
Reliability	Reliability
Safety	Opeying occupational health and safety rules
Rules	Following organization rules and proper procedures
C. Job Dedication	
Analyzing	Effectively handling new situations
Multi-tasking	Ability to carry out tasks not part of own job
Innovation	Generating new ideas to make things (tasks) better (innovation)
Planning	Planning and organizing work
Responsibility	Participating responsibility in the organization
Productivity	Working harder than necessary
Effectiveness	Spending the resources with effectively
Concentration	Concentrating to the duties
Decision making	Decision making
*Quality**	*Not making errors*
*Creativity**	*Creativity to solve a work problem*
*Self-development**	*Engaging in self-development to improve own effectiveness*

(*) Not used