Süleyman Demirel Üniversitesi
İktisadi ve İdari Bilimler
Fakültesi Dergisi
Y.2018, C.23, S.4, s.1505-1519.

Suleyman Demirel University
The Journal of Faculty of Economics
and Administrative Sciences
Y.2018, Vol.23, No.4, pp.1505-1519.

THE EFFECTS OF INTERPERSONAL AFFECT AND OPPORTUNITY TO OBSERVE ON RATINGS IN MULTI-SOURCE ASSESSMENT PROCESS

ÇOK KAYNAKLI DEĞERLEME SÜRECİNDE KİŞİLERARASI ETKİ VE GÖZLEMLEME İMKÂNININ DEĞERLEMELER ÜZERİNE ETKİSİ

Emin KAHYA[*]

[*] Prof. Dr., Eskişehir Osmangazi Üniversitesi, Mühendislik ve Mimarlık Fakültesi, Endüstri Mühendisliği Bölümü, ekahya@ogu.edu.tr, http://orcid.org/0000-0001-9763-2714

ABSTRACT

In related literature, although various rater demographic characteristics, opportunity to observe and interpersonal affect have been considered in many studies, a small number of investigations has indicated the influence of interpersonal affect and opportunity to observe on ratings in multi-source assessment process. In this study, we investigated whether rater affect has a similar effect on the ratings from three sources in multi-source assessment process and whether there is an interaction between rater's affect and the opportunity to observe the rate. All the white collar employees (39 persons) within a medium-sized manufacturing company participated in the study. The findings indicate that the influence of interpersonal affect on ratings was significantly greater in subordinate and peer feedback than in supervisor feedback.

Keywords: *Multi-source assessment process, Performance appraisal, Interpersonal affect, Opportunity to observe*

Jel Codes: *J24, J53.*

ÖZ

İlgili literatürde, çeşitli değerleyici karekteristikleri, gözlemleme imkânı ve kişilerarası etki çoğu araştırmada dikkate alınmasına rağmen, çok az araştırma, çok kaynaklı değerleme sürecinde kişilerarası etki ve gözlemleme imkânının değerlemeler üzerine etkisini araştırmıştır. Bu çalışmada, değerleyici etkisinin, çok kaynaklı değerleme sürecinde, üç kaynaktan (yönetici, arkadaş ve ast) değerlemelerin aynı etkiye sahip olup olmadığı ve değerleyici etkisi ile değerleyicinin gözlemleme imkânı arasında bir ilişki olup olmadığı araştırılmıştır. Orta ölçekli bir üretim işletmesinde tüm beyaz yakalı personel (39 kişi) çalışmaya katılmıştır. Bulgular, kişiler arası etkinin değerlemeler üzerine etkisinin, yönetici değerleme ile karşılaştırıldığında, ast ve arkadaş değerlemesinde daha fazla olduğunu göstermektedir.

Anahtar Kelimeler: *Çok kaynaklı değerleme süreci, Performans değerlemesi, Kişisel etki, Gözlemleme imkânı*

Jel Kodları: *J24, J53.*

1. INTRODUCTION

Companies are able to improve workforce quality through quality assurance system, job evaluation and payment system, proposal-incentive systems and initiatives that enhance quality of work life. Change in workforce can be gauged through systematic and regular performance appraisal system. One of the most popular approaches to performance assessment is

the use of multi-source performance evaluation.

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 MSA, 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).

Performance appraisal systems are employed for various purposes: administrative (e.g. for deciding on employee compensation and promotion), developmental (training or learning needs assessment), role-definition (defining and communicating roles) and strategic (goal orientation and self-monitoring) (Razzaq et al., 2016). 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. 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 MSA (Beehr et al., 2001). 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, greater reliability, enhanced fairness, and increased rate acceptance (Antonioni and Park, 2001). Previous empirical research has addressed the benefits of 360 degree feedback, the benefits of peer and upward appraisals, and the extent of self-other agreement in ratings (Kahya and Çemrek, 2017). 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 revealed that performance ratings are influenced by various factors such as rater and ratee demographic characteristics (age, gender, education level, job experience) (e.g., Sundvik and Lindeman, 1998; DeNisi and Murphy, 2017), cognitive process (e.g., Spence and Keeping, 2011; Roch et al., 2012) and interpersonal affect (e.g., friendship, liking) (e.g., DeNisi and Sonesh, 2011). There exist an extensive literature on the relation between interpersonal affect and ratings. The results showed that interpersonal affect creates bias in performance evaluation ratings (e.g., Antonioni and Park, 2001; Varma et al., 2005; Ng et al., 2011; Sutton et al., 2013; Razzaq et al., 2016). 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 (Zajonc, 1980), 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 interpersonal affect toward 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 (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 common denominator in these relationships. For example, 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. As such, if interpersonal affect for a ratee develops before the rater processes performance-related information, and is difficult to disconnect from actual performance, it is logical to argue that interpersonal affect is a source of bias in performance appraisal, diminishing rater accuracy (Varma et al., 2005).

Performance evaluation literature has recognized that interpersonal affect may play a significant role in performance appraisal. It is expected that rater affect would influence ratings regardless of whether the rater was a supervisor or a peer. Raters who like ratees may give higher ratings than raters who dislike ratees which may contribute to lenient or severe ratings. Such rating errors can lead ratees to believe that their performance are better or worse than their actual performance. There is reason to believe that rater affect would have a stronger influence on peer and upward ratings than on downward ratings because peers and upwards have less experience in evaluating others.

A number of studies have investigated interpersonal affect in performance ratings. Although many of these 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. Varma et al., 1996). These reviews have concluded that although interpersonal affect has an influence on performance ratings, the mechanism for this influence is not clear (Varma et al., 2005).

The present study investigates the influence of raters' interpersonal affect 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. Few have explained performance ratings from multiple sources. The primary purpose of the present study is to support and extend previous research by focusing on the extent to which the joint role of performance evaluations is tested in a wider range of demographic characteristics, rater resources, type of criteria used for assessing employees, and reputational roles in an organization.

It has been suggested that individual characteristics may influence on ratings. Therefore, the experience and education level of the ratee and rater will be used as control variables to better ascertain the incremental relationship between interpersonal affect 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 rater affect on ratings from three rater resources (peer, supervisor, and subordinate). The current study makes three unique contributions to the previous research.

First, some authors have received the observation opportunity from the participants as a statement such as "I frequently have the opportunity to observe the work behaviors of the person [name]". In this study we suggest two measures (observation time, and physical distance) for the opportunity to observe by considering the relationship among the departments or manufacturing lines.

Second, reputational roles in manufacturing organizations vary from worker to department manager. The present study was designed to reveal how employee position might be related to affect in ratings. The employee positions were ;

a) Workers (blue collar employees)
b) supervisors (first-line managers of the workers)
c) white collar employees including officers, sub-department managers (chiefs) and department managers.

Thirdly, the recent studies have examined the relation between affect and new dimensions of performance. Researchers have expanded the domain of performance to include contextual performance. There is an evidence that affect is associated with several dimensions of contextual performance. The present study is an attempt to reveal whether performance criteria used to measure employee performance is associated with differential effects of interpersonal affect on ratings.

2. MULTI-SOURCE PERFORMANCE EVALUATION

Performance is a term indicating how far away is a person, group or organization from target point in a certain period or a 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 task expected from them completely, some can not show expected success. The degree of success can be determined by performance evaluation (Kahya and Çemrek, 2017).

Performance evaluation is objective analysis and synthesis to determine how well the skills of the staff fit in with the qualities and requirements of job or how well they perform their expected tasks (Sabuncuoğlu, 2000). Performance evaluation is one of 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 are produced in a certain period and in various fields (such as wage, promotion, etc.) (Akdemir, 2009).

One of the most recent and popular approaches for performance evaluation is the use of multi-source performance evaluation. It 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).

The 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 Sarıgül, 2015). It 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 performance appraisal, according to traditional approaches, it is argued that only supervisors can assess 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 to evaluate employees' behaviors. The participants are managers, peers, internal and external customers (especially in service systems), lower level staff (subordinate) and self in evaluating (Kahya and Çemrek, 2017).

Several recent studies have focused on the implementation and effectiveness of 360 degree performance appraisal based on different rater resources (Baltacı and Burgazoğlu, 2014; Uygur and Sarıgül, 2015; Karkoulian et al., 2016; Kanaslan and Iyem, 2016; Kahya and Çemrek, 2017).

3. INTERPERSONAL AFFECT AND OPPORTUNITIES TO OBSERVE IN RATINGS

Most of the studies examining interpersonal affect have been conducted in laboratory, where interpersonal affect and performance levels can be easily manipulated. The laboratory is not the best setting for studying interpersonal affect, and the external validity of laboratory results is often suspect (Lefkowitz, 2000). Lefkowitz (2000) and DeNisi and Sonesh (2011) specifically pointed to the importance of understanding the role of rater affect in performance ratings. Of the various rater effects proposed to influence rating quality, rater interpersonal affect, or liking, has frequently been implicated as a pervasive source of bias as in performance evaluations (Suttan et al., 2013). Indeed, interpersonal affect develops over time between a supervisor and a subordinate, and systematically influences the performance rating process (Robbins and DeNisi, 1998). Further, the relationship between a supervisor and subordinate is also a developmental process that is a function of the length of relationship. As such, the effects observed in the laboratory may differ in important ways to the effects of interpersonal affects in the field (Varma et al., 2005).

Based on the discussion in the literature, we affirmed that a rater affect towards a ratee may influence ratings regardless of the source of downward, upward, or peer. One of the related studies of the relationship between rater affect and ratings errors, Tsui and Barry (1986) reported that interpersonal affect was positively related to halo. The results of their study indicated that raters with positive affect tend to exhibit to higher ratings (high leniency) and negative affect is related to lower ratings (high severity) in all three types of performance evaluation. Positive affect results in higher levels of halo in ratings. In this study, we tested the relationship between rater affect and performance ratings.

It seems plausible that ratings should be based on an acceptable amount of opportunities to observe ratee performance. Although some evidence exists that opportunity to observe influences the reliability of ratings, less is known about the consequences of varying opportunity to observe on the results of validation studies (Moser et al., 1999). The foregoing findings show that rater affect influences raters' attention and interpretation in rating process. Through repeated observations of ratees, raters may store up increasing amounts of affect-consistent information and interpretation (Antonioni and Park, 2001). According to this approach, less opportunity to observe results in less reliable ratings and may lead to a reduction in the validity coefficient. We investigated whether a rater's observation time is moderated with the influence of interpersonal affect on ratings.

Most of the studies on this topic have focused on performance appraisal from a single source, primarily traditional downward appraisals. Few have compared assessments from multiple sources. Antonioni and Park (2001) investigated whether rater affect had a similar effect on the leniency of ratings from sources of 360 degree feedback. In their study, an attractive relationship scale developed by Tsui and Barry (1986) measured interpersonal affect. They indicated that the influence of rater affect on the leniency of ratings was significantly greater in upward and peer feedback than in downward feedback and the influence grows with increasing raters' observation time. Because peers may face more ambiguity when giving performance evaluations and may feel less accountable for their ratings than managers do, peer and upward ratings may tend to be more susceptible to rater affect

than downward ratings. The influence of rater affect on the ratings may be stronger in peer and upward ratings than in downward ratings.

Interpersonal affect may be a basis for a rater's attempt to preserve ratee in situations where appraisals will be used for promotions and rewards. It is expected that worker raters, who are blue collar employees, tend to avoid conflict with co-workers and reflect a friendship bias much more than white collar employees do. Therefore, we will investigate the role of employee position on interpersonal affect.

4. METHOD

4.1. Participants

This study was conducted on a medium sized furniture company with about 200 employees. All the white collar employees (39 persons) in the company were participated in the study. All participants received a two-hours workplace training program including principles of assessment such as reliability, validity, fairness, definitions of the criteria used to evaluate employee performance, and halo errors in ratings.

4.2. Performance assessment

In this study, employees received ratings from three different sources; one employee received assessments from three peers, from the first and also second level managers and from subordinates. The peers for each employee were randomly selected. One employee assessed one peer from the same department and two peers from the nearest departments.

The performance assessment in this study was a form of contextual performance. Twenty-two contextual performance criteria to evaluate employee performance were taken from Kahya and Çemrek (2017). 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. 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. The raters from different sources used different criteria to assess ratees. To illustrate at this point, workers assessed their own co-workers with 11 criteria though they received assessments with 15 criteria from their first managers and also second managers. The final performance score was calculated from the ratings to each of the criteria used. All data was 213 ratings as follows:

- 57 supervisor (32 first and 25 second line managers) ratings
- 45 subordinate ratings
- 111 peer ratings.

4.3. Opportunity to observe

It seems plausible that ratings should be based on an acceptable amount of opportunities to observe ratees' performances. Less opportunity to observe results in less reliable performance ratings and may lead to a reduction of validity coefficients. Such raters are more likely to give inflated ratings and less likely to distinguish among ratees. By giving uniformly high ratings, they can avoid the potentially unpleasant consequences of assigning high ratings to some ratees and low ratings to others. Several authors have argued that new employees should be evaluated only after a certain time has passed because supervisors need adequate opportunities to observe ratees' behaviors (Rothstein, 1990; Moser et al., 1999).

There exist various options for measuring opportunity to observe. In Antonioni and Park (2001)'s study, three criteria measured the amount of rater observation time. Rater responded to statements such as "I frequently have the opportunity to observe the work behaviors of the person [name] I am evaluating" using a seven-point response scale ranging from strongly

disagree (-3) to strongly agree (+3). This kind of subjective measure can be confounded with such variables as raters' willingness to rate the ratee, attractiveness of the ratee and even the performance appraisal itself. Some authors such as Moser et al. (1999) used the number of months or years as a measure of "opportunity to observe". Even though this is only a rough estimate of opportunity to observe, it is easily measurable and manageable.

In this study, "opportunity to observe" was measured by two separate items: a) observation time, b) physical distance. Observation time was the number of years that the rater was working with the ratee together. We equated the length of time to "rater-ratee acquaintance", which is the minimum length (restricted with five years) of experience years of ratee and rater. To determine the physical distance level of between a rater and a ratee, the closeness levels between departments were investigated by conducting a questionnaire to the supervisors. They were asked to rate each department in terms of the relationship in their particular work environment on a five-point scale from 1 (unconnected) to 5 (highly connected). Because accurate ratings are impossible without some kind of acquaintanceship, it is plausible to assume that less physical distance produces more accurate ratings. "Opportunity to observe" was produced with multiplying by two items and then the scores are ranging from 1 to 25.

4.4. Interpersonal affect

In organizational psychology, interpersonal affect has taken many different forms and been measured in many different ways. Some forms include friendship, personal acquaintance, physical attractiveness, and the extent of familiarity between raters and ratees (Tsui and Barry, 1986). An attractive relationship scale developed by Tsui and Barry (1986) measures interpersonal affect. The measure in their study included the interpersonal feelings of admiration, respect, and liking based on rater-ratee interactions. Raters respond to statements such as "I like this person" using a seven-point scale ranging from strongly agree (1) to strongly disagree (7). This approach has an important drawback because they hesitate to rate the best. In the current study, therefore, interpersonal affect of rater toward ratee was evaluated by rater's supervisor. Supervisors were asked to rate each of their staff on how much the rater likes the ratee on a five-point scale ranging from strongly dislike (1) to strongly like (5). In order to reduce some problems associated with common influence, data were collected at separate times.

5. RESULTS

5.1. Correlations among all the variables

Table 1 reveals descriptive statistics (minimum, maximum, means and standard deviations). All the directors and department chiefs had university degree. All white collar employee were high school and upper degree. However foremen who are the first supervisor in manufacturing lines had under high school and professional school degree. Job experiences were ranging from 0.50 to 15 years. The raters also tended to rate favorably (Mean 3.29). In generally, the raters had positive affect toward the ratees (Mean 3.57) and opportunity to observe the ratees (Mean 14.49).

The ratings for all the rater categories are given in Table 2. In generally, department directors tended to be more lenient toward the peers and their chiefs (Mean 3.86 and 3.67), however the average rating for employees was under the expected score (Mean 3.31). It can be the result of not to observe nearly subordinates. Similarly, sub-department chiefs tended to rate neutral to the peers (Mean 3.03). That is, the more rates observed ratees, the influence of rater affect on ratings increased.

Pearson correlation coefficients of the variables are depicted in Table 3. There are several interesting findings that should be noted.

Tablo 1: Descriptive Statistics

	Min	Max	Mean	Std.dev.
Rater's experience	0.50	15.00	7.97	4.56
Rater's education level	Under high school	Master		
Ratee's experience	0.50	15.00	6.99	4.36
Ratee's education level	Under high school	Master		
Opportunity to observe	0.75	25.00	14.49	7.51
Interpersonal affect	2.00	5.00	3.57	0.70
Job performance	2.33	4.44	3.29	0.45

Table 2: Performance Ratings

	Ratee		
Rater	Director	Chief	Employee
General Director (n=1) (only rater)	3.39	3.03	-
Director (n=4)	3.86	3.65	3.31
Chief (n=13)	3.52	3.03	3.13
Employee (n=21)	-	3.52	3.23
Foreman (n=17) (only subordinate rater)	-	3.66	-

Table 3: Pearson Correlation Coefficients

a.All the ratings

	1	2	3	4	5	6
1. Rater's experience	-					
2. Rater's education level	-0.071					
3. Ratee's experience	0.227**	0.157*				
4. Ratee's education level	0.136	-0.087	-0.264*			
5. Opportunity to observe	0.458**	-0.026	0.478**	-0.075		
6. Interpersonal affect	0.000	0.093	-0.022	0.044	0.131	
7. Job performance	-0.114	-0.047	0.006	0.180**	0.009	0.220**

b.Superwiser ratings

	1	2	3	4	5	6
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. Opportunity to observe	0.124	0.156	0.751**	-0.183		
6. Interpersonal affect	0.181	0.331*	0.180	0.086	0.211	
7. Job performance	-0.040	-0.079	-0.041	0.274*	0.115	0.281

c.Peer ratings

	1	2	3	4	5	6
1. Rater's experience	-					
2. Rater's education level	-0.272**					
3. Ratee's experience	0.338**	0.134				
4. Ratee's education level	0.166	0.044	-0.255**			
5. Opportunity to observe	0.384**	0.084	0.419**	-0.116		
6. Interpersonal affect	-0.148	0.005	-0.150	0.000	0.083	
7. Job performance	-0.102	-0.004	0.005	0.004	-0.046	0.187*

d.Subordinate ratings

	1	2	3	4	5	6
1. Rater's experience	-					
2. Rater's education level	-0.375*					
3. Ratee's experience	0.060	0.515**				
4. Ratee's education level	0.078	-0.107	-0.498**			
5. Opportunity to observe	0.761**	-0.323*	0.153	-0.034		
6. Interpersonal affect	0.007	0.080	0.103	0.193	-0.034	
7. Job performance	-0.294	0.175	-0.169	0.313*	-0.294	0.279

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

There was a positive relationship between the leniency of performance ratings and rater affect toward rate ($r=0.220$, $p<0.01$) (Table 3.a). Surprisingly, job performance was neutral with opportunity to observe ($r=0.009$) for all the raters. It is expected that increased observation time gives observers more information for accurate ratings, but it may also amplify the influence of rater affect. Rater education levels was negatively correlated with job performance ($r=-0.047$) and ratee education levels was positively correlated with job performance ($r=0.180$); thus, job performance may diminish somewhat with increasing education, but no significant relationship was found among them.

As depicted at some investigations, rater' experience were weakly but negatively correlated with job performance ($r=-0.114$) which means that experienced employers prefer to give lower ratings (severity).

However, they influenced from ratees' experiences (r=0.227).

As expected, opportunity to observe was positively correlated with rater's experience (0.458, p<0.01) ratee's experience (r=0.478, p<0.01).

Job performance was positively correlated with opportunity to observe (r=0.115) and interpersonal affect (r=0.281) for the superwiser raters (Table 3.b). However, peer ratings were negatively correlated with opportunity to observe (r=-0.046) and positively correlated with interpersonal affect (r=0.187, p<0.05) (Table 3.c). The results show that the influence of interpersonal affect was stronger in superwiser than in peer. Surprisingly, the influence of rater affect on the leniency of rating was weak in peer and upward ratings than in downward ratings. It means that the raters are not to tend a friendship bias toward the ratees.

5.2. Multiple regression analyses

Multiple regression analysis were conducted to shed further lights, the relative contribution of each variable to the prediction of performance. The specific linear multiple regression model is drawn as below

$$P_j = \beta_0 + \beta_1 x_{1,j} + \beta_2 x_{2,j} + \beta_3 x_{3,j} + \ldots + \beta_k x_{k,j} \quad (1)$$

where P_j is performance rating of j^{th} ratee, $x_{i,j}$ is the level scale for i^{th} variable in the j^{th} ratee, $\beta_0, \beta_1, \beta_2, \beta_3 \ldots \beta_k$ are parameters, and n is the number of ratees.

To control the incremental effects of the most significant variables (opportunity to observe and interpersonal affect) in the regression analysis, four hierarchical regression models were established; in all the models, job performance rating was dependent variable, and four variables (Rater's experience, Rater's education level, Ratee's experience and Ratee's education level) were fixed as independent variables. The base model (Model 1) for job performance ratings is given as

$$P_j = \beta_0 + \beta_1(\text{Rater's experince})_j + \beta_2(\text{Rater's education level})_j + \beta_3(\text{Ratee's experince})_j + \beta_4(\text{Ratee's education level})_j \quad (2)$$

For the next models, opportunity to observe and interpersonal affect were extended to the previous model.

Model 2 = Model 1 + opportunity to observe

Model 3 = Model 1 + interpersonal affect

Model 4 = Model 1 + interpersonal affect + opportunity to observe

Such an approach shows the effect of the extended variable on performance ratings. Table 4 reports the standardized coefficient β, R^2 (adjusted), ΔR^2 (the change in R^2).

Table 4: Hierarchical Multiple Regression Analyses

Independent variables	Step 1	Step 2	Step 3	Step 4
Rater's experience	-0.017	-0.020	-0.017	-0.018
Rater's education	-0.021	-0.019	-0.028	-0.028
Ratee's experience	0.012	0.009	0.012	0.011
Ratee's education	0.091	0.092	0.088	0.088
Opportunity to observe		0.005		0.002
Interpersonal affect			0.137	0.134
R^2 (Adjusted)	0.044	0.043	0.085	0.081
ΔR^2		-0.001	+0.041	+0.037
ΔR2		-0.001	+0.041	+0.037

The first column in Table 4 (Model 1) shows that only ratee's education level (β=0.091, $p<0.01$), significantly, explained the variance in the ratings of job performance. Interestingly, raters tended to be more severe toward educated employees, and then expect from them to be much more efficient than the other. Other variables were not significant predictors. Nevertheless, all the variables accounted for 4.4% of the total variance of ratings (R^2=0.044) which is quite small. In all the analyses (Model 1-4), other three fixed variables were not found to be significant (all p's>0.05).

When "Opportunity to observe" was entered into regression equation (2) (Model 2), the change in R^2 was -0.1%. As predicted, "Opportunity to observe" did not account for additional, significant variance in predicting job performance after controlling the main effects of the variables in the base model. In Model 3 (including interpersonal affect), ΔR^2 was 4.10%. Interpersonal affect along with the control variables accounted for %4.1 of the total variance in the leniency of ratings. In support of our predictions, interpersonal affect was a statistically significant predictor of job performance, This variable has an important effect on ratings (β=0.137). Raters tend to give higher ratings to the some ratees, and this causes to more lenient ratings. The results of the Model 4 show that the most significant predictors on ratings were interpersonal affect (β=0.134), ratee's education level (β=0.088), and ratee's experience (β=0.011), respectively.

6. DISCUSSION

In performance evaluation literature, individual characteristics (such as age, gender, experience), observation time, interpersonal affect, rating format, workplace deviant behaviors, have been considered in many studies. In this study, we investigated whether rater affect has a similar effect on the ratings from three sources in multi-source assessment process and whether there is an interaction between rater's affect and the opportunity to observe the rate. All the white collar employees within a medium-sized manufacturing company participated in the study.

The results suggest that the nature of the rating content may be an important determinant of the relationship between affect and ratings. Varma et al. (1996) concluded that the influence of rater affect was stronger in trait-like ratings than in task-related ratings. These findings suggest that affect may serve to help raters interpret or "make sense" of ambiguous performance cues. If so, we should expect to be greater affect for trait ratings.

We found that different resources in the multi-source assessment process were influenced by interpersonal affect unequally. A major contribution of this study is the finding that the influence of interpersonal affect was stronger in downward and peer ratings than it was in upward ratings. This findings rise some questions. It is believed that each resource may be influenced differently by the some factors.

In addition to demonstrating a relationship between job performance and interpersonal affect, this study provides further support for the assertion that there was an interaction with job performance. In many studies correlation coefficient among them varies. In Antonioni and Park (2001)'s study, a total of 433 employees of a midsize insurance company was volunteer to be part of the study. Complete data existed for 163 downward ratings, 103 upward ratings and 1027 peer ratings. Their findings indicated that influence of rater affect on the leniency of ratings was significantly greater in upward and peer than in downward and that the influence increased as raters' observation time increased. Interpersonal affect (r=0.45) and observation (r=0.10) were positively associated with ratings. Varma and Pichler (2007), using data from

190 supervisors, further delineate the relationship between affect and job performance. They found that interpersonal affect was strongly and positively correlated with job performance (r=0.66). This results are not overlapping with ours (r=0.22). In accordance with past laboratory research, results indicate that performance level has a significant effect on performance ratings (Varma and Pichler, 2007). Laboratory research suggests that affect should have a significant influence on performance ratings. Our results buttress laboratory findings; they indicate that affect has a significant and positive influence on performance ratings.

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' study, experience had significantly correlated with job (task) performance (r=0.30-0.40). Although our results (r=0.006) differed from their findings, they not only support the above sight but also overlap too much with Moser et al. (1999)'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 five different rater resources. The participants were one general director as only a rater, 4 directors, 13 chiefs, 21 employees and 17 foremen (only subordinate raters). The general director rated the directors and their chiefs. The directors, chiefs and employees rated to each other (downward, peer and upward ratings). The foremen are blue-collars employees and rated only the chiefs as a upward raters. The other one of the major contributions of the present study is that "opportunity to observe" was measured by two separate items: a) observation time, b) physical distance. Observation time was the number years the rater was working with the ratee together, which is the minimum length of experience years of ratee and rater. To determine the physical distance level of between a rater and a ratee, the closeness levels between departments were investigated by conducting a questionnaire to the supervisors. "Opportunity to observe" was produced with multiplying by two items.

We believe that this study makes a significant contribution to the literature on interpersonal affect, but it has some limitations. This study conducted on a middle sized furniture company and all the data were collected from 39 white collar employees. Although the study may not be generalized to other manufacturing companies, the findings can partly depend on the criteria and attitude of the manager working in the company. This is not a dilemma. Organizational culture and performance evaluation training program designed to improve appraisal skills of raters can lessen some disadvantages on ratings.

As the future research, 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, we don't intend to drive the job performance evaluation 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

1. AKAL, Z. (2005). İşletmelerde Performans Ölçüm ve Denetimi: Çok Yönlü Performans Göstergeleri, 6.Baskı, Milli Prodüktivite Merkezi Yayınları No: 473, Ankara.
2. AKDEMİR, A. (2009). İşletmeciliğin Temel Bilgileri, Ekin Yayınları, Bursa.
3. ANTONIONI, D. and PARK, H. (2001). "The relationship between rater affect and three sources of 360-degree feedback ratings". Journal of Management, 27: 479-495.
4. ANTONIONI, D. and WOEHR, D.J. (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.
5. BALTACI, A. İ. and BURGAZOĞLU, H. (2014). "Değerlendiriciler Arası Güvenilirlik ve Tatmin Bağlamında 360 Derece Performans Değerlendirme". Marmara Üniversitesi Öneri Dergisi, 11(41): 57-76.
6. BEEHR, T.A., IVANITSKAYA, L., HANSEN, C.P., EROFEEV, D. and GUDANOWSKI, D. (2001). "Evaluation of 360 degree feedback ratings: Relationships with each other and with performance and selection predictors". Journal of Organizational Behavior, 22: 775-788.
7. Barutçugil, İ. (2002). Performans Yönetimi, 2. Basım, Kariyer Yayıncılık, İstanbul.
8. BORMAN, W.C. and MOTOWİDLO, S.J. (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.
9. BRUTUS, S., PETOSA, S. and AUCOIN, E. (2005). "Who will evaluate me? Rater selection in multi-source assessment contexts". International Journal of Selection and Assessment, 13(2): 129-138.
10. CARDY, R.L. and DOBBINS, G.H. (1986). "Affect and appraisal accuracy: Liking as an integral dimension in evaluating performance". Journal of Applied Psychology, 71: 672-678.
11. DENISI, A. S. and MURPHY, K. R. (2017). "Performance appraisal and performance management: 100 years of progress?". Journal of Applied Psychology, 102(3): 421-433.
12. DENISI, A. S. and SONESH, S. (2011). "The appraisal and management of performance at work". In S. Zedeck (Ed.), Handbook of industrial and organizational psychology (pp. 255–280). Washington, DC: APA Press.
13. DENISI, A.S., CAFFERTY, T.P. and MEGLINO, B.M. (1984). "A cognitive view of the performance appraisal process: a model and research propositions". Organizational Behavior and Human Performance, 33(3): 360-396.
14. FLETCHER, C. and BALDRY, C. (1999). "Multi-source feedback systems : A research perspective". In International Review of Industrial and Organizational Psychology, 14: 149-193.
15. KAHYA, E. and ÇEMREK, F. (2017). "An Investigation on the Ratings from Four Sources for Different Positions in a 360 Degree Feedback System". Eskişehir Osmangazi Üniversitesi İİBF Dergisi, 12(3): 49-64.
16. KANASLAN, E. K. and IYEM, C. (2016). "Is 360-degree feedback appraisal an effective way of performance evaluation?". International Journal of Academic Research in Business and Social Sciences, 6(5): 172-182.

17.KARKOULIAN, S., ASSAKER, G. and HALLAK, R. (2016). "An Empirical Study of 360-degree Feedback, Organizational Justice, and Firm Sustainability". Journal of Business Research, 69: 1862-1867.

18.LEFKOWITZ, J. (2000). "The role of interpersonal affective regard in supervisory performance ratings: A literature review and proposed causal model". Journal of Occupational and Organizational Psychology, 73: 67-85.

19.MOSER, K., SCHULER, H. and FUNKE, U. (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.

20.NG, K.Y., KOH, C., ANG, S., KENNEDY, J.C. and CHAN, K.Y. (2011). "Rating leniency and halo in multisource feedback ratings: testing cultural assumptions of power distance and individualism-collectivism". Journal of Applied Psychology, 96(5): 1033-1044.

21.RAZZAQ, S., IQBAL, M.Z., IKRAMULLAH, M. and PROOIJEN, J.W.V. (2016). "Occurrence of rating distortions and ratees' fairness perceptions per raters' mood and affect". Career Development International, 21(7): 726-743.

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

23.ROBBİNS, T.L. and DENISI, A.S. (1998). Mood vs. interpersonal affect: Identifying process and rating distortions in performance appraisal. Journal of Business and Psychology, 12(3), 313-325.

24.ROCH, S.G., WOEHR, D.J., MISHRA, V. and KIESZCZYNSKA, U. (2012). "Rater training revisited: an updated meta-analytic review of frame-of-reference training". Journal of Occupational and Organizational Psychology, 85(2): 370-395.

25.ROTHSTEIN, H.R. (1990). "Interrater reliability of job performance ratings; Growth to asymptote level with increasing opportunity to observe". Journal of Applied Psychology, 75(3): 322-327.

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

27.SPENCE, J.R. and KEEPING, L. (2011). "Conscious rating distortion in performance appraisal: a review, commentary, and proposed framework for research". Human Resource Management Review, 21(2): 85-95.

28.SUNDVIK, L. and LINDEMAN, M. (1998). Performance rating accuracy: Convergence between supervisor assessment and sales productivity. International Journal of Selection and Assessment, 6(1): 9-15.

29.SUTTON, A.W., BALDWIN, S.P., WOOD, L. and HOFFMAN, B.J. (2013). "A meta-analysis of the relationship between rater liking and performance ratings". Human Performance, 26(5): 409-429.

30.TSUI, A.S. and BARRY, B. (1986). "Interpersonal affect and rating errors". Academy of Management Journal, 29(3): 586-599.

31.UYGUR, A. and SARIGÜL, S.S. (2015). "360 Derece Performans Değerleme ve Geri Bildirim Sistemi". Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 33: 189-201.

32.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.

33. VARMA, A., DENİSİ, A.S. and PETERS, L.H. (1996). "Interpersonal Affect and Performance Appraisal: A Field Study". Personnel Psychology, 49: 341–59.

34. VARMA, A. and PICHLER, S. (2007). "Interpersonal affect : Does it really bias performance appraisals?". Journal of Labor research, 28(2): 397-412.

35. VARMA, A., PICHLER, S. and SRINIVAS, E.S. (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.

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