Journal of Economics, Business, and Accountancy Ventura Vol. 19, No. 3, Decmber 2016 – March 2017, pages 393 – 404

Halo effect in subjective performance evaluation bias

Andi Ina Yustina[1], Gudono[2]

[1] University of President, Ki Hajar Dewantara Street RT 2/RW 4, Mekarmukti, Cikarang Utara, Bekasi, 17550, West Java, Indonesia
[2] University of Gadjah Mada, Bulaksumur, Caturtunggal, Depok, Sleman, 55281, DIY, Indonesia

A R T I C L E I N F O

Article history:
Received 25 June 2016
Revised 23 November 2016
Accepted 24 January 2017

JEL Classification:
G18

Key words:
Subjective Performance Evaluation, Objective Performance Evaluation, Controllability, and Halo Effect.

DOI:
10.14414/jebav.v19i3.621

A B S T R A C T

This study aimed not only to examine the effect of the objective measure and controllability on subjective performance evaluation but also to prove empirically the halo effect phenomenon, which is present in the evaluation process when evaluators are faced with two or more different measurement dimensions. This study used a 2 × 2 factorial web-based experiment involving 62 undergraduate students and 77 sales managers in the Telecommunications industry. The results reveal the subjective performance evaluation manager is directly influenced by objective measurement based on sales performance. Subjective evaluation of performance evaluator will be high when the objective performance information managers showed a high score and vice versa. The level of controllability affects undergraduate students in conducting subjective performance rating. This evidence suggests that the two subjects of this research using their discretion in conducting the performance appraisal rating. Halo effect is proven to have high correlation with two different dimensions of performance measurement.

A B S T R A K

Penelitian ini bertujuan tidak hanya menguji pengaruh ukuran obyektif dan pengendalian evaluasi kinerja subjektif tetapi juga untuk membuktikan secara empiris fenomena halo effect yang ada dalam proses evaluasi ketika evaluator menghadapi dua atau lebih dimensi pengukuran. Penelitian ini menggunakan eksperimen berbasis web faktorial 2 × 2 melibatkan 62 mahasiswa sarjana dan 77 manajer penjualan di industri Telekomunikasi. Hasil menunjukkan manajer evaluasi kinerja subjektif secara langsung dipengaruhi oleh pengukuran yang objektif berdasarkan kinerja penjualan. Evaluasi subjektif kinerja evaluator akan tinggi ketika kinerja manajer informasi yang objektif menunjukkan skor tinggi dan sebaliknya. Tingkat pengendalian mempengaruhi mahasiswa dalam melakukan subjective performance rating. Bukti ini menunjukkan bahwa kedua subjek penelitian ini menggunakan kebijaksanaan mereka dalam melakukan rating penilaian kinerja. Hallo effect terbukti memiliki korelasi yang tinggi dengan dua dimensi yang berbeda dari pengukuran kinerja.

1. INTRODUCTION

Performance evaluation system models that organization used are objective performance evaluation and subjective performance evaluation. Objective performance evaluation is based on quantitative data, using real measurement and the final result will reflect on the organizational targets such as productivity, profitability, and sales growth (Bella Vance et al. 2013). However, objective performance evaluation is defined as judgment based on personal impressions that is unquantified such as innovation, creativity, loyalty, the ability to work together, sharing knowledge, leadership, or the ability to communicate (Baker et al. 1994; Bella Vance et al. 2013, Bol 2008). Both models are interplay one another (Bol and Smith 2011), can be adopted by the company simultaneously (Prendergast 1999), and actually not mutually exclusive but rather complementary (Breuer et al. 2013; Bommer et al. 1995; Prendergast and Topel 1996).

Previous studies discussed the benefits of

* Corresponding author, email address: [1] a.inayustina@gmail.com.

subjective performance evaluation. Subjective performance evaluation is more reliable in formulating and designing incentive systems contracts (Ittner et al. 2003; Gibbs et al. 2004; Breuer et al. 2013), developing training programs (Goldman and Bhatia 2012) or to promote employee positions (Gibbs et al. 2004; Kren and Tyson 2009). However, previous studies also revealed the weakness of the subjective performance evaluation model. Subjective performance evaluation, which is based on the human judgment, is likely cause a number of problems. Employers who used a subjective evaluation tend to behave favoritism, and will cause the manager gives a performance evaluation, which does not reflect the actual performance evaluation and even different evaluation from one subordinate to another subordinate (Prendergast and Topel 1996). Another impact of the subjective performance evaluation is the manager will use discretion to make score adjustments that induce to inaccuracies in the provision of rating (Moers 2005; Bol and Smith 2011).

Bol and Smith (2011) used the subjective and objective performance evaluation models to detect the possibility of bias of rating. Their results prove that the evaluators tend to use their discretion to raise the rating of subjective performance evaluation when the objective performance information indicates low scores and low controllability. Controllability is the employee's ability to control the conditions of the uncertainty toward the outcomes obtained (Tan and Lipe 1997). The managers will increase their subjective evaluation when controllability lead to low outcomes in order to avoid the unfairness from the employee that will impact on the motivation and job satisfaction (Colquitt et al. 2001). Otherwise, when controllability leads to high outcomes, managers will not use their discretion to correct the judgment. Further findings of Bol and Smith (2011) also revealed that the evaluators with a high rating on one dimension of measurement will also give a high value on other dimensions. Indirectly this behavior indicates a problem of a halo effect in providing appropriate performance evaluation.

The psychology literature confirmed that the halo effect is common in performance evaluation. Murphy et al. (1993) said that the more dimensions of measurement used in assessing the performance, the greater potential of halo effect will happen, thus allowing bias. This argument is also confirmed by studies from Bechger et al. (2010) which said that the halo effect will arise and affect the rating when the evaluator is faced two different measurements. Evaluators who have information or a general impression that precedes a more detailed assessment of performance will lead to an attitude of consistency to make the initial information as a benchmark to provide an evaluation of the subordinate (Utami et al. 2014). Thus, the evaluator will provide an interpersonal assessment to subordinates based on the general impression that they received (Nisbett and Wilson 1977).

This study examined the research of Bol and Smith (2011) about the bias from the evaluators. Two things that need to be criticized from research of Bol and Smith (2011) and will be used for improvement in this study by considering the role of the halo effect in rating evaluation, which are:

First, it is related to subjects of research. Subjects were given prior objective information with score then followed with subjective information. Indirectly, the subjects will have cognitive distortions caused by the weighting of information. The manager will give a high rating without need to see the subjective information by itself. The author assumed that the rating bias found by Bol and Smith (2011) is because they do not consider the possibility of a halo effect. Bol and Smith (2011) acknowledged this in the limitations of their research:

> "All participants receive the qualitative information for their subjective evaluation after observing the manager's sales score. Yet the generalizability of the distortion and adjustment result we observe may not be limited to this order information presentation" (Bol and Smith 2011: 1227).

Therefore, the authors reverse the information by providing subjective information and then objective information to prove whether the findings of Bol and Smith (2011) remained consistent when treatment is modified and to prove whether there is a halo effect that interfere evaluators in performance evaluation.

Second, the subjects are the supervisors in University. They were asked to assume themselves as the managers of business organizations to evaluate their subordinates. The use of this subject may have affected the results. This study used sales manager as a subject because they have more experience and better knowledge of the actual conditions in the field, so that they will understand the case provided better. This is in line with that proposed by Nahartyo (2012) that the most appropriate subjects for research that focus on the business is the business itself (Nahartyo 2012: 172). This study also employed the undergraduate

students as a comparison with the consideration that the students have not sufficiently known the reflect real-world conditions to evaluate the performance, so it is likely to affect the results of performance assessment (Tan and Lipe 1997).

2. THEORETICAL FRAMEWORK AND HYPOTHESES

Halo Effect

Halo effect was first identified by Wells (1907) and Webb (1915) as a phenomenon of "errors" that occur in the process of judgment (Jacobs and Kozlowski 1985). Halo effect is a strong tendency of a person to give a high rating (good) or a low rating (inferior) based on the general impression (Thorndike 1920). Halo effect is also considered as one of the response biases of evaluators to generalize positive impression and a negative impression of subordinates, which are assessed, based on certain characteristics. Evaluators who have limitations to observe the performance of subordinates will rely on their minimal information to provide an assessment (Friedson and Rhea 1965). The impression that emerged will give a great influence on the subsequent assessment (Tetlock 1983).

Several researchers described the character and the impact of this phenomenon of halo effect. Jacobs and Kozlowski (1985) said that the halo effect is considered common and can be found everywhere, especially in performance evaluation. Halo effect is also considered as an assessor's inability to discriminate specific attributes possessed by individuals assessed (Murphy and Jako 1989). The same thing is also expressed by Bernardin and Beatty (1984) that the majority cause of the halo effect is the overall impression of an evaluator (rater) to individuals assessed (ratée) so they ignore the specific attribute assessment. Bowditch and Buono (2001) said that the halo effect needs to be alert because it can interfere with the performance evaluation process caused by certain characteristics inherent to the individual.

They said that the employee evaluation can not be based only one performance measurement. Someone who has shown the best performance to the company will not necessarily be given a high rate because there are other factors considered by evaluators. For example, an employee shows a maximum sales performance during his work, yet he is often not timely in the works. If the initial information (high sales performance) is received by the evaluator, then the evaluator is likely to give a lower assessment with the presence of the following information (not on time) and will affect the overall rating of the evaluators.

The possibility of a halo effect emerged in performance evaluation when the company uses more than one dimensions of performance measurement has been a long discussion in the Psychology literature (Bechger et al. 2010; Belzer and Sulsky 1992; Kozlowski and Kirsch 1987; Murphy et al. 1993; Nisbett and Wilson 1977). The evaluator will give an independent evaluation based on the different measurements for weighting on one measurement (Nisbett and Wilson 1977). Initial information by the manager of his subordinates is considered to be the most important before evaluating the other criteria and make this knowledge as a benchmark (anchor) in making decisions (Murphy et al. 1993 and Grcic 2008).

Bol and Smith (2011) said that the cognitive limitations of evaluators will preclude him in making decisions related to performance evaluation. The implication is that the evaluator will make the wrong and biased decision because it is not based on a real evaluation. Consequences of the characteristics of this halo effect can reduce the quality of an evaluator evaluation and lead to inaccuracies in rating. Several approaches can be taken to lower or reduce the halo effect, which are through the provision of training, increasing familiarity with the individuals assessed, or using statistics as a control in assessing (Cooper 1981).

Subjective Performance Evaluation

Performance evaluation is an important instrument for the company. Basically, performance evaluation uses two models, which are objective performance evaluation and subjective performance evaluation. The Company uses objective performance assessment based on measurement or data that can be quantified, real, and refers to the organizational targets such as productivity, profitability, and sales growth (Bellavance et al. 2013). However, based on several studies, the company not only uses quantitative data, but also uses measurements based on human judgment such as creativity, sharing knowledge, loyalty, leadership, communication skills and so on (Baker et al. 1994; Bol 2008; Bellavance et al. 2013). Other researchers have also proven that both models are feasible to be adopted simultaneously because both models did not eliminate each other, yet they act as a complementary (Bommer et al. 1995; Prendergast 1999; Bol and Smith 2011).

Subjective performance evaluation is used as a complement to objective measurement because not all aspects of performance can be measured by

using quantified objective measurements (Breuer et al. 2013). Otherwise, the researchers acknowledged that the subjective measurement is more reliable for contract design, arranging training of employees, and can be used as a reference in the recruitment of employees (Goldman and Bhatia 2012). In addition to provide a number of benefits for the company, the subjective research model is also susceptible to biases that encourage supervisors to use discretion, thereby reducing the rate and accuracy in planning compensation (Moers 2005; Bol and Smith 2011; Bol 2011).

Controllability

Controllability is defined as an employee's ability to control the conditions of uncertainty that affect outcomes (Tan and Lipe 1997). According to Tan and Lipe (1997), the measurement of that can be controlled will lead to better decisions and drive up the successful outcome, whereas measurements can not be controlled will lead to failure and drive worse decisions. Furthermore, Holmstrom (1979) said that a controlled performance evaluation is more favorable because it is closely related to the contract design so the employees that are evaluated on these conditions will feel fair and appreciated while the low controllability reduces in formativeness of the performance evaluation.

In the real conditions, measurements that can actually be controlled are almost rare, for example, the compensation system in the company. The compensation system is an objective measurement which is always in controlled by the company, but not according to empirical research done by Baker et al (1994). They said that the compensation system as a controlled measurement used as a reference in objective measurement, then sometimes results in ambiguous information because it is not based on a real evaluation. Other studies said that unfair compensation has a negative effect on motivation and job satisfaction (Greenberg 1987; Baker et al. 1994; Colquitt et al. 2011). When this condition occurs (low control) evaluators will always strive to make discretion by correcting the evaluation into higher rate (adjusted upward).

Hypothesis Development

Performance measurements that are generally used by the company are formal measurements, rather than unique measurements (informal). An objective performance measurement is more real (verifiable), quantified, and the results will lead to organizational targets such as productivity, profitability, and sales growth (Bellavance et al. 2013). The resulting data is more easily processed and provide the time efficiency for managers to make decisions quickly and accurately in evaluating their subordinates.

Several researchers revealed that using a formal measurement is not sufficient to perform a more comprehensive performance evaluation. Evaluators need other elements that can not be captured by a quantifiable measurement (Bol 2008) but can only be obtained by using informal performance evaluation. This informal measurement refers to subjective evaluation that is based on human judgment such as innovation, creativity, and loyalty (Baker et al. 1994), communication skills, attitude, leadership, and ability to work in a team (Bellavance et al. 2013).

The use of two-dimensional measurement in performance evaluation is likely to occur as mentioned previously, so does in real condition. If one of the information is already known by the evaluators and dominates their judgment, there is a tendency to give an inaccurate rating. This influence is caused by manager's cognitive distortion (Bond et al. 2007 and Bol and Smith 2011). This cognitive distortion occurs when one of the inherent information (anchor) followed by other information that is likely to influence the judgment from superiors to subordinates. In this condition, it is potential that the halo effect presents and affects the evaluator's decision in assessing the performance (Murphy et al. 1993). Managers who have had good preliminary information regarding the performance of subordinates will use the additional information obtained to provide a good weight (high) or vice versa a bad (low) assessment on other measurements. Therefore, the halo effect is thought to have a correlation with the use of different measurements so the evaluators will be high or low in weighting the rating evaluation. Thus, the first hypothesis is:

H1: Supervisor's subjective performance evaluation will be higher (lower) when the employee's objective performance measure is high (low).

Ideally, objective measures that can be controlled are the best for employees who are being assessed. But, the measurements that can be completely controlled are extremely rare (Bol and Smith 2011). For example, evaluators who rely on objective measurements were not able to capture all aspects that are not quantified (Prendergast and Topel 1996), therefore evaluators will consider other aspects that tend to be more subjective (Baker et al. 1994; Boomer et al. 1995).

Controllability influences the evaluator's evaluation since it is closely related to outcome (Tan and Lipe 1997). If the employee's ability to control the low uncertainty factors leads to low-performance evaluation, then employee will feel unfairness. Employees feel that performance measure does not capture the whole effort and contribution that have been given to the company. Perception of unfairness from an employee would induce demotivation or dissatisfaction (Colquitt et al. 2001). Evaluators would try to use their subjective judgment to adjusting upward. If controllability leads to higher performance evaluation, the negative effect in the work will not appear. The evaluator will give score in accordance with the actual performance. This is consistent with empirical research done by Cropanzano and Konovsky (1995) which said that the evaluators will not consider fairness when the outcome is high but will consider the fairness factor when the outcome is low. Thus, the second hypothesis is:

H2: Supervisor's subjective performance evaluation will be higher (lower) when the controllability level of objective performance measurement is lower (higher).

3. RESEARCH METHOD

Design and Instrument Development

The researcher employed a 2 × 2 between subjects and used experiments design that has been developed and tested previously by Bol and Smith (2011). Participants assumed themselves as the regional director for an industrial pipe and fitting company. The regional director had supervisory authority on ten district managers, which are responsible for two divisions, sales division, and administration division. Regional director evaluates the performance to one of district managers based on objective information and will use personal notes and some other information collected from the interview of office staff. The subjects were grouped into 4 groups and got different treatments, which are objective performance information (high and low) and controllability (high and low).

Participant

Participants in this research are 62 undergraduate students majoring in accounting in sixth semester and 77 managers of sales division in the telecommunications industry in Indonesia.

Experiment Design and Procedure

The access of experiment links case made by www.esurv.org consists of 4 links in accordance with the treatment in the experimental design of both undergraduate students and manager. Before manager completed the case, two screening questions are given to avoid error in subject selection. In order to ensure that participants understand the experiment case given, manipulation checks are conducted by answering questions and providing scores presented on a 1 to 7 scale.

The Dependent and Independent Variable

The dependent variable is based on the subjective performance of the administrative office district managers. Participants will be given information about the personal notes of district managers and the interviews of some of the subordinate staff of managers. Then based on the information, the participants will provide score with the score range of 0-10.

The independent variables manipulated are objective performance level and controllability. The level of objective performance is based on individual sales during six months period. High sales will be given a score of 9 and for low sales will be given a score of 2, while the controllability is manipulated by the presence of external factors that could affect the level of the district manager.

In order to detect the involvement of the halo effect in performance evaluation, participants will be asked to answer the questions that have been tested by Utami et al. (2014) adapted to the study: (1) participants asked to give agreement about the influence of granting a score of objective performance toward performance appraisals of evaluated administration office managers, (2) participants asked to assess the influence of granting score to the evaluated district manager on subjective evaluation.

Pilot Test

The experimental material used in this study was obtained from the research of Bol and Smith (2011). However, since the research was made on different time, place, and using different subjects, then the validity test of the instrument was done by conducting translation of experiment materials from English to Indonesian and Indonesian to English. The first pilot test was conducted on 10 students of Master of Science UGM and the second pilot test was conducted on the 16 employees who have worked for over 3 years to test the 4 links that have been created through www.esurv.org.

Table 1
Participation Information

Panel A: Demographics (n=62)

		Number	Percent
Age:	19 – 23 years	62	100.00%
Gender:	Male	27	43.50%
	Female	35	56.60%

Panel B: Background Information (n=62)

		Number:	Percent
Experience:	< 1 year	3	4.80%
	> 1 year	1	1.60%
	Inexperienced	58	93.50%

4. DATA ANALYSIS AND DISCUSSION

This section outlines in detail about data analysis and hypothesis testing on two subjects that are used in this research.

Study I: Undergraduate Student

Demographic

The first subject used 98 undergraduate students in Indonesia. There were 13 participants who must be removed due to not answering completely (8 participants) and did not pass the manipulation check (5 participants). In total, participants that can be processed are 62 people (63%) consist of 19 people in the cell 1, 14 people in the cell 2, 16 people in the cell 3, and 13 people in the cell 4. Characteristics of undergraduate students are set out in Table 1.

Based on a statistical test, there was no significant difference in the performance evaluation scores given by participants. F results of gender testing = 0.070 and significance value 0.792. This is consistent with Maas and Torez-Gonzalez (2011) research which proved that gender does not affect the subjective performance evaluation. Work experienced tests show F values = 0.106 and sig value 0.746, which means there is no difference in determining the subjective performance evaluation scores, whether it is given by the experienced and inexperienced participants. Homogeneity test shows that groups of data samples are coming from populations with same variance. Levene's Test results shows 0.864 with p-value 0.465.

Manipulation Checks

The researcher included two step manipulation check in post experimental questionnaire (on 7 point scale). First step is that respondents were asked whether target manager's individual sales score was higher than others manager (mean = 4.403; s.d. = 2.161) and the extent to which level of controllability affect to manager's individual sales (mean = 3.742; s.d. = 1.717). Both of our manipulation measures for this study were successful. Second step, we asked to our respondents to rate about personal and interview record of target manager to make judgment related to office administration performance. The results of this measure shows an average value was 5.177 (s.d. = 1.337), which indicates a significantly higher value than the midpoint of 3.5. We also ask to our respondent about level of confidence in providing subjective performance evaluation. The result reveal that the average value was 5.113 (s.d. = 1.189) that is higher than the scale midpoint of 3.5. The last, we ask about both performances (administrative and sales) were interrelated. The result show that value of 3,855 (s.d. = 1.658) was higher than midpoint. From all manipulation check questionnaire results suggest that all respondent in study I had understanding to give evaluation.

Test of Hypotheses

Table 2 is a descriptive statistics and the result of our hypothesis on 62 undergraduate students that will explain the average number in each group and their standard deviations. The average value of the subjective performance evaluation scores (mean = 7.31) for the low controllability conditions, is higher when objective performance information than the low objective performance information. In contrast to the high controllability condition, the mean value of the subjective performance evaluation scores (mean = 6.37) is higher when objective performance information lower than the high objective performance information. Overall, the average value of the subjective performance evaluation is higher in the low controllability condition (mean = 6.97). These results are consistent with the predictions of this study that the subjective performance evaluation scores will be higher in the low controllability condition.

The first hypothesis predicts that participants

Table 2
Descriptive Statistics and Result of Testing Hypotheses

Panel A: Descriptive Statistics

Controllability	Sales		
	High	Low	Overall Means
	n=35	n= 27	n= 62
High (n= 33)	6.37 (1.300)	6.57 (1.284)	6.45 (1.277)
Low (n=29)	7.31 (1.138)	6.54 (0.967)	6.97 (1.117)
Overall Means	6.80 (1.302)	6.56 (1.121)	

Panel B: Result of Testing Hypotheses

Source	df	Means Square	F-test	p-value
Corrected Model	3	2.887	2.029	0.120
Sales	1	1.238	0.870	0.355
Controllability	1	3.151	2.214	0.142
Sales * Controllability	1	3.623	2.547	0.116

Table 3
Pearson Correlation of Halo

	Pearson Correlation	
Overall rating to objective performance evaluation	0.948**	P value= 0.000
Overall rating to subjective performance evaluation	0.412**	P value=0.001

** Pearson correlation sig. 0.001

will provide high subjective performance evaluation when the individual sales are high. Instead, participants will give a low subjective performance assessment when the individual sales are low. The hypothesis 1 is a test with the level of sales (high and low) that are at a high level of controllability. The result of hypothesis 1 testing on the subject of students shows the average value of performance evaluation on the condition of high sales is smaller than the average value in the condition of low sales. This result is in contrary to the first hypothesis, which states that when sales are high, the subjective evaluation will also be high and vice versa. Subjective evaluation given by the undergraduate students are not significant (F = 0.199, p = 0.659) and did not reveal any main effect or interaction effect of each independent variable on the dependent variable. Thus, hypothesis 1 is not supported by the data.

Halo is observed by comparing the coefficient correlation between the overall ratings given by participants with specific measurement dimension (Thorndike 1920). Pearson correlation on Table 3 shows (p-value = 0.001). That is, the use of two-dimensional measurements when evaluating has a strong correlation with the detection of a halo.

The second hypothesis predicts that the subjective performance evaluation by evaluators will be higher (lower) when the controllability level of objective performance measurement is lower (higher). The homogeneity test value of 0.362, show the subjective performance evaluation will be higher when the level of controllability is lower, and vice versa when a higher level of controllability, subjective performance evaluation will be lower evaluator. F test value of 5,124 (p = 0.030), therefore the second hypothesis is supported.

Study II: Practitioner
Demographic

The second study used 118 sales managers at telecommunications companies in Indonesia. All participants passed the screening questions that given in the beginning of the experiment. Data that can be used is 77 practitioners (65%) consisted of 18 people in the cell 1, 19 people in the cell 2, 19 people in the cell 3, and 21 people in the cell 4. The demographic characteristics of these participants are explained more in Table 4.

The tests of equality of gender, age, and work experience were done. F value for statistical test of gender is 0.191 with a significance value of 0.664 which means there is no significant difference in the subjective performance evaluation given whether the participants' gender is male or female. This result is also in compromise with the empiri-

Table 4
Participation Information

Panel A: Demographics (n=77)

		Number	Percent
Age:	23 – 35 years	38	49.40%
	36 – 45 years	35	45.50%
	46 – 55 years	4	5.20%
Gender:	Male	65	84.40%
	Female	12	15.60%

Panel B: Background Information (n=77)

		Number	Percent
Experience:	5 years	2	2.60%
	6 – 8 years	10	13.00%
	9 – 12 years	22	28.60%
	>12 years	43	55.80%
Level:	Supervisor	44	57.10%
	Manager	33	42.90%

cal research of Maas and Tores-Gonzalez (2011) which states that gender does not influence the subjective performance evaluation. Statistical analysis for age test shows the value of F = 1.244 with a significance value of 0.294. Result of statistical test for work experience demonstrates the value of F = 0.528 with a sig 0.665. The purpose of homogeneity test is to show that two or more cells of sample data are derived from the population that have the same variance. Levene's Test shows the result of the homogeneity test 2.700 with p-value 0.052.

Manipulation Checks

Like this study, manipulation check was conducted with two steps in post experimental questionnaire (7 point scale) on sales manager. First step, question about whether target manager's individual sales score was higher than other manager and whether level of controllability affect to manager's individual sales. Both of our manipulation questions were higher than midpoint (mean = 4.130; s.d. = 2.308 and mean = 4.208; s.d. = 1.757) which means that all manipulation measurements were success. Second step question about rating respondent to target manager's personal notes and interview in make judgment regarding to office administration performance. We found that value of 5.338 (s.d. = 1.324), significantly higher than midpoint of 3.5. We also asked to respondent about the level of confidence in rating subjective performance. Our result show the average value of 5.753 (std.dev = 1.041) which means significantly higher than the scale of midpoint. The last question about interrelated between objective performance and administration office. Our result value of 3.403 (s.d. = 1.703) indicated that both of performance measures cannot separated. Finally, all there results suggest that respondents felt sufficient in giving evaluation.

Test of Hypotheses

Based on Table 5, the average value of the subjective performance evaluation scores (mean = 7.56) for high controllability conditions, is higher when the objective performance information is higher than the low objective performance information. In contrast to the low controllability condition, the average value of the subjective performance evaluation (mean = 6.89) is higher when objective performance information of sales conditions is higher than low objective performance information. Overall, the average value of the subjective performance evaluation is higher when the objective performance information is high (mean = 7.22). This proves that the participants provide a more logical assessment on objective performance district manager. Experience possessed by the participants in assessing the performance of the participants made use of a common strategy to provide a more real valuation (Tan and Lipe 1997).

The first hypothesis stated that participants would provide high subjective performance assessment when individual sales are high. Instead, participants will give a low subjective performance evaluation when the individual sales show a low result. To test the hypothesis 1, the conditions of sales levels (high and low) used are at a high level of controllability. It is based on the fairness perceptions of employees assessed. In high controllability conditions, the evaluation will be

Table 5
Descriptive Statistics and Result of Testing Hypotheses

Panel A: Descriptive Statistics

Controllability	Sales		
	High	Low	Overall Means
	n=37	n= 40	n= 77
High (n= 37)	7.56 (0.784)	6.47 (1.020)	7.00 (1.054)
Low (n=40)	6.89 (1.524)	6.52 (1.601)	6.70 (1.556)
Overall Means	7.22 (1.250)	6.50 (1.340)	

Panel B: Result of Testing Hypotheses

Source	df	Means Square	F-Statistic	p-value
Corrected Model	3	4.640	2.772	0.048
Sales	1	10.126	6.049	0.016
Controllability	1	1.789	1.069	0.305
Sales * Controllability	1	2.425	1.448	0.233

fair because the achievements will be rewarded and performance appraisal will be high. The result of hypothesis 1 testing on the business practitioner subjects show the average values in conditions of high sales is greater than the average value in conditions of low sales (cell 1> 2 cells). The mean value is 10.819 and subjective evaluation is significant (F = 12,976, p = 0.001). This shows that the subjective performance evaluation is strongly influenced by high or low individual sales achieved by the district manager. Thus, hypothesis 1 is supported.

Overall rating of the participants was compared with the dimensions of a particular measurement (Thorndike 1920). The result of Pearson correlation test shows p-value = 0.000 which indicates that the use of two performance measurement has a significant correlation to the overall performance evaluation.

The second hypothesis states that the subjective assessment evaluators will be higher (lower) when the controllability level objective performance measurement is lower (higher). F value of 2.705 (p-value = 0.109). Thus, the second hypothesis is not supported.

5. CONCLUSION, IMPLICATION, SUGGESTION, AND LIMITATIONS

It can be concluded that, when the objective performance measurement of a manager is low, the subjective performance evaluation scores given to the manager is also low. The mean value of high sales should be greater than low sales. However, result from the undergraduate students show the mean value of high sales lower than low sales. This fact indicates that the undergraduate students have cognitive distortions that lead to biased judgments. Subjective performance information that is presented at the beginning of the experiment stored in memory and the students become the benchmark to anchor for judging. However, due to subsequent information is not profitable (low sales) the undergraduate students ignore this information and use the preliminary information to provide subjective performance evaluation. This argument is convinced by the answers given by the subject when the researchers asked whether personal notes and interviews assist participants in making decisions related to subjective performance. The mean value of high sales and low sales are significantly different (high sales = 4.84, low sales = 5.21). The mean value for practitioner in high sales is higher than low sales. This is consistent with the hypothesis. Personal notes and interviews information affect the manager in making a rating. There is no significant differences of the average values are shown (high sales and low sales = 5.44 = 5.26). Practitioners are giving logical and rational rating in experience in assessing performance since they are supported by their experiences in working and performance evaluation. The empirical evidence confirms the result of Tan and Lipe (1997) research. The use of two-dimensional measurements in assessing the performance also proves the strong correlation of the potential presence of a halo with the overall performance assessment. Both subjects empirically prove this conjecture.

The second hypothesis is supported by the undergraduate students but is not supported by the practitioners. Subjective performance evaluation scores will be high in the low controllability

condition and the subjective performance evaluation scores will be low in the high controllability. The second hypothesis testing uses high sales and high controllability and high sales and low controllability. According to the theory, the average value of high sales and low controllability should be higher than the average value of high sales and controllability.

The data obtained from the undergraduate students are in line with the theoretical logic, but in contrast to the result shown by the practitioners. Subject students tend to ignore the uncontrollable conditions (low controllability) because the objective performance information has already given a positive outcome (high). In addition, from Attribution Theory said that someone will give attributes to the uncontrolled conditions as external factors outside of the individual so that the high valuation would be considered fair (Heider 1958). In contrast to the practitioners, empirical evidence shows the average value of the high sales is higher than the low controllability. Assessment decision is based on the achievement of objective performance. High controllability or low controllability condition does not affect the rate a business practitioners in rating the performance of managers. The decisions made by the practitioners tend to be fair because they use their experience in conducting performance evaluations.

This research gives strong evidence of the influence of objective performance information with subjective performance evaluation. The evaluators will provide high subjective judgments when objective information also indicates a high score. Conversely, when objective information indicates a low score, the subjective performance evaluation will also be low. The managers prove it; however, undergraduate students do not. It is most likely caused by the difference in work experience and experience in conducting performance evaluations. Managers have had a long working experience (over 5 years) and have experience in conducting performance evaluations while the students have not had a work experience. The results of this study confirm the research of Bol and Smith (2011).

Controllability on objective measurements also affects the subjective performance evaluation. This is proven in the second hypothesis testing. Subject of the undergraduate students prove the relationship level with subjective performance evaluation rating. While different result is shown in the subject of managers. Undergraduate students are having cognitive distortions when controllability is low, which influence the subjective evaluation. This study also confirms with the research of Tan and Lipe (1997).

The use of the subject should be considered in conducting research as well as possible. This study proves that the subject of undergraduate students and managers differ in how to make decisions related to performance evaluation.

This study has some limitations. First, experiments were carried out on the subject of business practitioners through web. Therefore, the control on the respondents was very low. The respondents did not obtain an adequate explanation from the researcher. Thus, the perception of potential bias is very high. Second, this study uses only one industry, so the result of this study cannot be generalized. Third, this study detects halo when two different performance measurements are used, but do not test how to reduce the halo.

Future studies need to improve the control of the respondents that the respondent's perception of bias can be minimized. By using a variety of industries, it will obtain the best results that can be generalized and will increase external validity. In order to get a more actual assessment results, it is necessary to use certain measurements to detect and reduce the halo effect in performance evaluation.

REFERENCES

Baker, G, Gibbons, R & Murphy, KJ 1994, 'Subjective Perormance Measures in Optimal Incentive Contracts', *The Quarterly Journal of Economics*, 109 (4), 1125-1156.

Bechger, TM, Maris, G & Hsiao, YP 2010, 'Detecting Halo Effect in Performance-Based Examinations', *Applied Psychological Measurement*, 34 (8), 607-619.

Belzer, WK & Sulsky, LM 1992, 'Halo and Performance Appraisal Research: A Critical Examination', *Journal of Applied Psychology*, 77 (6), 975-985.

Bellavance F, Landry, S & Schiehll, E 2013, 'Procedural Justice in Managerial Performance Evaluation: Effect of Subjectivity, Relationship Quality, and Voice Opportunity', *The British Accounting Review*, 45, 149-166.

Bernardin, HJ, and Beatty, RW 1984, *Performance Appraisal: Assessing Human Behavior at Work*, Boston, Kent.

Bol, JC 2008, 'Subjectivity in Compensation Contracting', *Journal of Accounting Literature*, 27, 1-23.

Bol, JC 2011, 'The Determinant and Performance

Effects of Managers' Performance Evaluation Biases', *The Accounting Review*, 86 (5), 1549-1575.

Bol, JC & Smith, SD 2011, 'Spillover Effect in Subjective Performance Evaluation: Bias and the Asymmetric Influence of Controllability', *The Accounting Review*, 86 (4), 1213-1230.

Bommer, HW, Johnson, JL, Rich, GA, Podsakoff, PM & Mackenzie, SB 1995, 'On the Interchangeability of Objective and Subjective Measure of Employee Performance: A Meta-Analysis', *Personnel Psyhcology*, 48, 587-605.

Bond, SD, Carlson, KA, Meloy, MG, Russo, JE & Tanner, RJ 2007, 'Information Distortion in the Evaluation of a Singgle Option', *Organizational Behaviour and Human Decision Processes*, 102, 240-254.

Bowditch, JL & Buono, AF 2001, *A Primer on Organizational Behavior,* Fifth Edition, John Wiley & Sons, New York.

Breuer, K, Nieken, P & Sliwka, D 2013, 'Social Ties and Subjective Performance Evaluation: An Empirical Investigation', *Review Management Science*, 7, 141-157.

Colquitt, JA, Conlon, DE, Wesson, MJ, Porter, COLH & Yee Ng, K 2001, 'Justice at the Millineum: A Meta-Analytic Review of 25 Years of Organizational Justice Resarch', *Journal of Applied Psychology*, 86 (3), 425-445.

Cooper, WH 1981, 'Ubiquitous Halo', *Psychological Bulletin*, 90, 218-244.

Cropanzano, R & Konovsky, MA 1995, 'Resolving the Justice Dilemma by Improving the Outcome: The Case of Employee Drug Screening', Journal of Business and Psychology, 10 (2), 221-243.

Friedson, E & Rhea, B 1965, 'Knowledge and Judgment in Professional Evaluations', *Administrative Science Quarterly*, 10 (1), 107-124.

Gibbs, M, Merchant, KA, Van der Stede, WA & Vargus, ME 2004, 'Determinants and Effect of Subjectivity in Incentives,' *The Accounting Review*, 79 (2), 409-436.

Goldman, J & Bathia, S 2012, 'Performance Evaluation Inflation and Compression', *Accounting, Organizations and Society*, 37, 534-543.

Grcic, J 2008, 'The Halo Effect Fallacy', *Electronic Journal of Philosophy*, 1-58.

Greenberg, J 1987, 'Reaction to Procedural Justice in Payment Distribution: Do the Means Justify the Ends?', *Journal Of Applied Psychology*, 72 (1), 55-71.

Holmstrom, B 1979, 'Moral Hazard and Observability', *The Bell Journal of Economics*, 10 (1), 74-91.

Ittner, CD, Larcker, DF & Meyer, MW 2003, 'Subjectivity and the Weighting of Performance Measures: Evidence from a Balanced Scorecard', *The Accounting Review*, 78 (3), 725-758.

Jacobs, R & Kozlowski, SW 1985, 'A Closer Look at Halo Error in Performance Ratings', *The Academy of Management Journal*, 28 (1), 201-212.

Kozlowsky, SWJ & Kirsch, MP 1987, 'The Systematic Distortions Hypothesis, Halo, and Accuracy: An Individual-Level Analysis', *Journal of Applied Psychology*, 72 (2), 252-261.

Kren L & Tyson, T 2009, 'Trade-off in Objective and Subjective Performance Evaluation: A Case Study Examining the Validity of Agency Theory Predictions', *Management Accounting Quarterly*, 10 (2), 12-23.

Maas, VS & Torres-Gonzalez, R 2011, 'Subjective Performance Evaluation and Gender Discrimination', *Journal of Business Ethics*, 101, 667-681.

Moers, F 2005, 'Discretion and Bias in Performance Evaluation: The Impact of Diversity and Subjectivity', *Accounting, Organization and Society*, 30, 67-80.

Murphy, KR & Jako, RA 1989, 'Under What Condition are Observed Intercorrelations Greater than or Smaller than True Intercorrelations?', *Journal of Applied Psychology*, 74, 827-830.

Murphy, KR, Jako, RA & Anhalt, RL 1993, 'Nature and Consequences of Halo Error: A Critical Analysis', *Journal of Applied Psychology*, 78 (2), 218-225.

Nahartyo, E 2012, *Desain dan Implementasi Riset Eksperimen*, UPP STIM YKPN, Yogyakarta.

Nisbett, RE & Wilson, TD 1977, 'The Halo Effect: Evidence fo Unconscious Alteration of Judgments', *Journal of Personality and Social Psychology*, 35 (4), 250-256.

Prendergast, C 1999, 'The Provision of Incentives in Firm', *Journal of Economic Literature*, 37 (1), 7-63.

Prendergast, C & Topel, RH 1996, 'Favoritism in Organization', *Journal of Political Economy*, 104 (5), 958-978.

Tan, HT & Lipe, MG 1997, 'Outcome Effect: The Impact of Decision Process and Outcome Controllability', *Journal of Behavioral Decision Making*, 10, 315-325.

Tetlock, PE 1983, 'Accountability and the Perseverence of First Impressions', *Social Psychology Quarterly*, 46 (4), 285-292.

Thorndike, EL 1920, 'A Constant Error in Psychological Rating', *Journal of Applied Psychological*,

82 (5), 665-674.
Utami, I, Kusuma, IW, Gudono & Supriyadi, 2014, 'Halo Effect in Analytical Procedure: The Impact of Client Profile and Information Scope', *Global Journal of Business Research*, 8 (1), 9-26. www.esurv.org.