Performance Improvement of Shif Production through the Appraisal Agreement Level

Bambang Biantoro[1], Aris Trimarjoko[2], Choiri Purwanto[3], Zulfa Fitri Ikatrinasari[4]

[1,2,3,4]Mercu Buana University, Jakarta, Indonesia

Abstract: **Performance appraisal has importance role in the reward system especially to giving fairness to the employee that can effect to working motivation. Appraisals that do by multiple raters - hierarchy level will generate differences result. The company that operates in 24 hours need system working in the shift pattern and will cause different interaction between employee and leader's that give raters. The purpose of this study is to analyze agreement level of raters in performance appraisal hierarchy model using Cohen's kappa method. This study shows that the agreement level inter rater was very good in quantifiable aspect example discipline, but hierarchy model has moderate and poor agreement level for behavior aspect. This different raters will be higher as the intensity of interaction between rater to the employee. To improve the agreement level was discussed in the end of this article.**

Keywords: **Performance Appraisal, Multirater, Agreement Level**

1. Introduction

Performance Assessment (performance appraisal) is a formal system that examines, reviews and evaluates the performance of individuals and groups involving leadership and members, in an effort to determine the level of achievement of members' achievements for rewards or other further requirements. Assessment performance is usually done by a supervisor who directly relates as a source of information or can also be done by a colleague or superiors, also called multi rater or 360 degrees (Aguinis, 2013). The analysis of multi rater performance is an ongoing discussion of the different assessments between the assessors so that the level of agreement between the assessors of the multi rater assessment method becomes a separate research topic (Denisi & Murphy, 2017). According to (Aguinis, 2013) multi-rater or 360-degree ratings are more used to defining career paths and will be able to find problems that can be used to address the weaknesses and strengths of employees. The purpose of this study is to know the level of agreement of performance appraisal on the pattern of process of shif production process that is done in multi-stage (multi rater).

2. Literature survey

Performance evaluation and performance management are the management activities implemented to build and improve employee performance through the development of human resource efficiency and development and perform performance evaluations over time by comparing agreed-upon performance standards and following systematic performance evaluation results to enhance individual and group performance within the organization thoroughly (Abdullah, 2014). Performance evaluation is a formal system used to assess the performance of employees on a regular basis as determined by organizations that are useful for measuring, evaluating and evaluating the performance of individuals or groups within the organization. In the performance evaluation according to Ivan Cevih (1992) as mentioned (Abdullah, 2014) has the following objectives:

- Development, with performance appraisals, can be seen where fewer employees have been able to provide training to cover the shortcomings.
- Giving remuneration, performance appraisal results can be used to determine pay increases, incentives and promotions although some organizations use to set aside workers.
- Motivation, performance evaluation can be used to motivate employees, develop initiatives and work confidence.
- HR planning, with performance appraisal results can be useful in developing the potential, expertise and skills of employees for HR planning.
- Compensation, with performance evaluations can be seen how to balance fairly according to performance levels.
- Communication, performance evaluation is a means of continuous communication between superiors and subordinates to employee performance.

Performance appraisal is a system that is useful for analyzing and evaluating the performance of workers both individually and in groups, of course performance appraisal in order to produce good results requires a good performance appraisal system, good performance assessment needs to fulfill several conditions (Mayasari, Haryanti, Hindiarto, 2012), that is:

- Reliability that is the factor / instrument of assessment must be able to be believed, the measure of achievement must be consistent if there are two or more assessors in assessing the same worker they can make conclusions on the same case against the results of the employee's performance evaluation.
- Relevant, compliance between achievement steps and output of worker performance.

International Journal of Research in Engineering, Science and Management
Volume-1, Issue-12, December-2018
www.ijresm.com | ISSN (Online): 2581-5792

594

- Sensitivity, which is a certain measure must reflect the difference between the high and low stages of appearance. Appearance must be able to distinguish carefully about the difference in achievement.
- Practical, may be used and easily understood in a practical manner and lack of data does not interfere with the scoring system.

With reference to the good performance appraisal conditions as mentioned above, this study will test the Assessment of the company Assessment, especially on the Criteria of Reliability, which is the level of consent of the three assessors (feeds) on existing valuation items using the Cohen Kappa method approach. Cohen Kappa is a measure that states a measurement assessment agreement made by two assessors (rater) or an agreement between two measurement / assessment methods or can also measure the agreement between two measuring tools (http://research-indonesia.blogspot.com/2012) To calculate the coefficient of Cohen's Kappa (K), use the following formula:

$$\kappa = \frac{\Pr(a) - \Pr(e)}{1 - \Pr(e)}, \qquad \text{K = (kappa coeficien value = -1 to 1}$$

Note, K: Cohen coefficient, s Kappa Pr (a): Percentage of consistent number of measurements between rater. Pr(e): Percentage of changes in measurement between rater. The value of the Cohen coefficient, s Kappa can be interpreted (Altman, 1991) as quoted (parametered.wordpress.com) as follows:

Table 1
Agreement Level Index

K value	Value agreement (Strength of Agreement)
<0.20	Low (Poor)
0.21 - 0.40	Fair
0.41 - 0.6	Moderate
0.61 - 0.80	Strong (Good)
0.81 - 1.00	Very strong (Very Good)

In addition to the calculations as in the formula above, the Cohen's Kappa coefficient calculation can be done with the help of SPSS software.
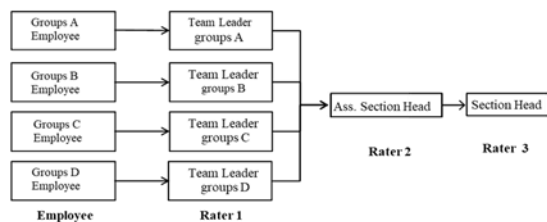
3. Research method

Fig. 1. Tiered Performance Assessment Model

This research is carried out at the final inspection department which forms part of the manufacturing involved in the automotive industry. Performance evaluation is an annual assessment to determine the rewards the company will give to employees. The 24-hours production system used in the company's work system makes employees divided into 4 groups and 3 changes. The performance appraisal model uses a multilevel system involving 3 assessors as in Fig. 1. The first assessor has interaction with employees during working hours. The second assessor interacts with the employee as long as the employee is in shift1 position (1 week in a month) while the third assessor rarely interacts with the employee.

Workers who are the object of the assessment (as many as 93 people), using SPSS software, the acquired data (employee assessment data) carried out the approval test with the Cohen's Kappa approach. This approval exam is carried out on 7 (seven) assessment instruments with 3 (three) rater which aims to determine the strength of the agreement of each party in the existing performance assessment instrument. The performance appraisal form for companies has 7 (seven) achievement assessment instruments, each of which has a different estimated weight, that is:

- Work knowledge that has an assessment weight of 25%
- Quantity and quality of work which has a weight rating of 25%
- Discipline has a weight rating of 15%
- Integrity has an assessment weight of 10%
- Initiatives have an assessment weight of 10%
- Collaboration has an assessment weight of 10%
- Loyalty has a weighting rate of 5%

The 7 (seven) assessment instruments there are 4 (four) consecutive assessment categories rather than the most damned ones to the best which must be filled by 3 (three) players so that the final achievement value is agreed based on 7 (seven) assessment instruments. With the Cohen Kappa method, the assessment decisions of 7 (seven) assessment instruments conducted by 3 (three) rater will be tested for the applicable approval stage to analyze the assessment system during the same time there are good weaknesses or still need to be corrected so that future assessment achievements are expected to be can fulfill the overall estimation of workers. The decision to make a decision regarding the Cohen Kappa exam based on the comparison of the Approximate importance value with the α value used in the study using the value α is 0.05, then the test hypothesis in this study is:

- H0> α, there is not agreement between the ratters.
- H1 <α, there is agreement between the ratters.
- For the strength of the agreement between the raffles, refer to schedule 1. Above.

4. Decision and discussion

A. Decision

From the data collection assessment of 93 workers with 3 (three) rater on 7 (seven) assessment instruments, the consistency test agreement for each instrument using the Cohen SPSS Kappa approach obtained the following kappa coefficient values.

B. Knowledge instruments regarding work: Rater 2 to rater

The work knowledge instrument of 93 workers assessed

International Journal of Research in Engineering, Science and Management
Volume-1, Issue-12, December-2018
www.ijresm.com | ISSN (Online): 2581-5792

595

between rivals 2 and rater 1 has a significant value of 0.000 less than the value of α, that is 0.05 (0,000 <0.05) then accepts H1 and rejects H0, which means an agreement between the two with the level of agreement (value) together with 0.514 simple categories.

Table 2
Kappa Coefficient Rater 2 to rater 1

	Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Symmetric Measures				
Measure of Kappa Agreement	0,514	0,109	4,981	0,000
N of Valid Cases	93			

C. Rater 3 to rater 2

Table 3
Kappa Coefficient Rater 3 to rater 2

	Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Symmetric Measures				
Measure of Kappa Agreement	0,027	0,038	0,645	0,519
N of Valid Cases	93			

The work knowledge structure of 93 workers assessed between raw 3 and rater 2 had significant values lower than 0.519 more than the value of α which was 0.05 (0.519> 0.05) then accepted H0 and refused H1 which means there are not agreement between the two raters.

D. Instrument Quality and quantity of work Rater 2 to rater 1

Table 4
Kappa Coefficient Rater 2 to rater 1

	Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Symmetric Measures				
Measure of Kappa Agreement	0,219	0,120	2,281	0,023
N of Valid Cases	93			

Instrument quality and work quantity than 93 workers assessed between competitors 2 and rater 1 have a value that is almost the value of 0.023 smaller than the value of α 0.05 (0.023 <0.05) then accept H1 and reject H0 which means the agreement between the two with graduation ratings (value) 0.219 fair categories.

E. Rater 3 to rater 2

Table 5
Kappa Coefficient Rater 3 to rater 2

	Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Symmetric Measures				
Measure of Kappa Agreement	-0,045	0,062	-0,744	0,457
N of Valid Cases	93			

In the instrument of quality and quantity of work of 93 employees assessed between rater 3 and rater 2 having an approximate value significance 0.457 is greater than the value of α which is 0.05 (0.457 > 0.05) then accept H0 and reject H1

which means there are not agreement between the two raters.

F. Instrument of Discipline Rater 2 to rater 1

Table 6
Kappa Coefficient Rater 2 to rater 1

	Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Symmetric Measures				
Measure of Kappa Agreement	1,000	0,000	11,489	0,000
N of Valid Cases	93			

In the instrument of discipline 93 workers rated between rivals 2 and rater 1 had a significant value of 0.000 smaller than the value of α which was 0.05 (0.000 <0.05) then received H1 and refused H0 which means there are agreement between two raters was a good agreement categories with value of 1.000.

G. Rater 3 to rater 2

Table 7
Kappa Coefficient Rater 3 to rater 2

	Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Symmetric Measures				
Measure of Kappa Agreement	0,941	0,041	10,777	0,000
N of Valid Cases	93			

in the work discipline instrument 93 workers assessed between 3 and 2 initially had significant values which 0.000 was smaller than the value of α, that is 0.05 (0.000 <0.05) then received H1 and refused H0 which means there are agreement between two rater was agreements (value) 0.941 very good categories.

H. Integrity Instruments Rater 2 to rater 1

Table 8
Kappa Coefficient Rater 2 to rater 1

	Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Symmetric Measures				
Measure of Kappa Agreement	0,377	0,121	4,493	0,000
N of Valid Cases	93			

The integrity instruments of 93 workers rated between rivals 2 and rater 1 had significant values which 0.000 was smaller than the value of α, that is 0.05 (0.000 < 0.05), then accepted H1 and refused H0, which means there are agreement between two raters with a value of 0.377 fair category.

I. Rater 3 to rater 2

Table 9
Kappa Coefficient Rater 3 to rater 2

	Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Symmetric Measures				
Measure of Kappa Agreement	0,245	0,145	2,538	0,011
N of Valid Cases	93			

International Journal of Research in Engineering, Science and Management
Volume-1, Issue-12, December-2018
www.ijresm.com | ISSN (Online): 2581-5792

596

In the integrity instrument of 93 employees assessed between rater 3 and rater 2 having an approximate significance value of 0.011 smaller than the value of α which is 0.05 (0.011 <0.05) then accept H1 and reject H0, which means there are agreement between the two raters with value 0.245 an agreement fair category.

J. Initiative instruments Rater 2 to rater 1

Table 10
Kappa Coefficient Rater 2 to rater 1

Symmetric Measures				
	Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Measure of Kappa Agreement	0,173	0,090	2,551	0,011
N of Valid Cases	93			

In the initiative instrument of 93 employees assessed between rater 2 and rater 1 have approximate significance value of 0.011 was smaller than the value of α which was 0.05 (0.011 <0.05) then accept H1 and reject H0 which means an agreement between the two rater with value agreement 0.173 that is poor categories.

K. Rater 3 to rater 2

Table 11
Kappa Coefficient Rater 3 to rater 2

Symmetric Measures				
	Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Measure of Kappa Agreement	0,164	0,122	1,916	0,055
N of Valid Cases	93			

In the initiative instrument of 93 employees assessed between rater 3 and rater 2 have approximate significance value of 0.055 was greater than the value of α which was 0.05 (0.055> 0.05) then accept H0 and reject H1 which means there was no agreement between the two rater.

L. Cooperation Instrument. Rater 2 to rater 1

Table 12
Kappa Coefficient Rater 2 to rater 1

Symmetric Measures				
	Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Measure of Kappa Agreement	0,147	0,069	2,340	0,019
N of Valid Cases	93			

The Cooperation instruments of 93 employees assessed between rater 2 and rater 1 have approximate significance value of 0.019 is smaller than the value of α which is 0.05 (0.019 <0.05) then accept H1 and reject H0, which means there was agreement between the two rater with an agreement value 0.147, that of poor categories.

M. Rater 3 to rater 2

The collaboration instrument of 93 employees assessed between rater 3 and rater 2 have an approximate significance value of 0.901 is greater than the value of α, which 0.05 (0.901>

0.05) then accept H0 and reject H1, which means there was not agreement between the two rater.

Table 13
Kappa Coefficient Rater 3 to rater 2

Symmetric Measures				
	Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Measure of Kappa Agreement	0,002	0,011	0,124	0,901
N of Valid Cases	93			

N. Loyality Instruments Rater 2 to rater 1

Table 14
Kappa Coefficient Rater 2 to rater 1

Symmetric Measures				
	Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Measure of Kappa Agreement	0,280	0,095	3,010	0,003
N of Valid Cases	93			

The loyalty instrument of 93 employees assessed between rater 2 and rater 1 have an approximate significance value of 0.003 is smaller than the value of α which is 0.05 (0.003 <0.05) then accept H1 and reject H0, which means there are agreement between the two raters with a value of 0.280, that fair category.

O. Rater 3 to rater 2

Table 15
Kappa Coefficient Rater 3 to rater 2

Symmetric Measures				
	Value	Asymptotic Standard Error[a]	Approximate T[b]	Approximate Significance
Measure of Kappa Agreement	-0,071	0,045	-2,348	0,019
N of Valid Cases	93			

The Loyalty instrument of 93 employees assessed between rater 3 and rater 2 have an approximate significance value of 0.019 is smaller than the value of α which is 0.05 (0.019 <0.05) then accept H1 and reject H0, which means there are agreement between the two raters with agreement value -0,071, that in the poor category.

P. Discussion

Of the 7 (seven) instruments in the form of ratings, only disciplinary instruments have the coefficient of kappa in the excellent category, which means a very strong agreement between 1, 2 and 3. While the other 6 instruments have a simple, fair and poor category or agreement valid and low, riders. even though between 3 and 2 offenders on the instrument of knowledge about work, quality and quantity of work, the initiative and cooperation there is no agreement between the rulers. Disciplinary instruments have a very strong level of agreement, which is caused by disciplinary instruments is the assessment of attendance (absenteeism) employees who have a definite quantitative data and distributed to all farmers (1,2 and 3) before the performance evaluation is conducted. Whereas 6 other valuation instruments are qualitative despite certain

definitive assessment criteria, but the difference between objectivity and understanding between predators about these criteria results in an inevitable tendency in the evaluation.

5. Conclusion

By using Cohen's Kapa's approach to the overall agreement between rater in giving the performance of appraisal employees on performance assessment system at this company it can be concluded that it is still not good, and required improvement suggestions so that future performance appraisal system will be better and can give a sense justice among workers so as to increase performance motivation.

A. Sugestions

Referring to the results of the performance assessment analysis using the Cohen kappa approach and looking for phenomena as mentioned above, then to increase the approval stage among riders, it is recommended that improvements to each assessment instrument be made as follows:

- Knowledge of work, written and unwritten exams regarding items of knowledge related to work so that the exact value of the results of the exam and the results of the exam are circulated to all parties before the assessment.
- The quality and quantity of work, quality and quantity of work are two different cases so that in the future these two cases will be separated from the assessment with different measurement parameters and after obtaining definite values the distribution of the results of the assessment to each occupant is much needed.
- Regarding the order instrument, the agreement has very good agreement so that there is no improvement.

- Integrity, Initiative, Cooperation and Loyalty is a performance assessment based on behavior that requires intensive guidance to competitors so that the feasibility of the assessment can be minimized
- In general, to avoid an assessment bias that resulted in a low level of agreement between participants, it was reserved for the reduction of racists from 3 to 2 rater (direct supervisors and supervisor assistants). Because furthermore the position of workers' appraisers is estimated, the underachievement of information about workers, especially for the performance appraisal category based on behavior.

References

[1] Aguinis, H. (2013). Management of Third Edition Achievements. Kelly's School of Commerce at Indiana University.
[2] Denisi, A. S., & Murphy, K. R. (2017). Performance assessment and management achievements: 100 years of progress? Journal of Gunaan Psychology, 102 (3), 421-433. https://doi.org/10.1037/apl0000085
[3] Prof. Dr. H. M. Ma'ruf Abdullah, SH. MM (2014). Aswaja Pressindo, Yogjakarta. Management Assessment and Worker Achievement.
[4] Fuadi, Totok Sumaryanto, Wahyu Lestari (2015), Journal of Theme of Educational Investigation and Assessment: Development of Science Learning Instruments Psychomotor Assessment of Green Plant Materials Based on Conservation Experiments Approach. Education Study and Assessment Study Program, Post-Student Program, Universiti Negeri Semarang.
[5] Isti mayasari; professional psychology master's journal, Kristiana Haranti, Ferdinandus Noarto (2012), Competency Based Performance Assessment and KPI (main bookmark) in regional water supply companies in the Semarang area. Professional Psychology Bachelor in the Postgraduate Student Program at the Soegijapranata Catholic University Semarang
[6] Kappa'S Cohen Consistency Exam http://research-indonesia.blogspot.com/2012