The Development of Peer Assessment Instrument for the Learning Model of Peer Assessment

Hendro Permadi¹, Toto Nusantara², I. Nengah Parta³ Sisworo⁴
Department of Mathematics, Faculty of Mathematics and Science, Malang University

Abstract: A teaching model of peer assessment being developed based on the cooperative learning model brings an impact of active participation of the university students who are pre-service teachers in the learning process. Based on the researchers’ observation, there are some superiorities of the use of peer interactive learning over other cooperative learning models such as: 1) while having a discussion, the students are very active because they are assessed by their friends in the group, 2) the students are more self-prepared for the materials to be given (being less afraid of not understanding the contents), 3) lecturers or teachers can do assessment while they are teaching, 4) the assessment is done in the process of learning. The objective of this research is the development of peer assessment instrument in the evaluation of process and the result of teaching. The research subjects were the students who were pre-service teachers who took the course of Mathematical Statistics. The identification of distribution is obtained based on the result of weighted factor analysis by using five variables, which are accuracy, completeness, comment, suggestion, and the use of language. The identification result of this distribution is expected to provide some assessments which can be contributed by the students to the result of lecturers’ or teachers’ evaluation. The result of factor analysis showed the overall percentage of the five variables, i.e. 43.668% consisting of accuracy (27.5%), completeness (21.4%), comments (59.8%), suggestions (54.7%), and the use of language (54.9%). Through simulations of 837 score combinations and giving weight to each value, the suitable result of distribution identification which was in accordance with the peer value was found. It has a distributing lognormal (3p) with a location parameter of 1.002, scale parameter of 0.463, and threshold parameter of 2.178. While the average value was 5.75 and the standard of deviation was 1.483. The result was in the form of scores given by the students through peer assessment contributing to the evaluation of learning process and the results done by the lecturers or teachers.

Keywords: Teaching Model of Peer Assessment, Development of Peer Assessment Instrument, Analysis of Factor, Lognormal Distribution (3p)

I. Introduction

As pre-service teachers, the students of mathematics education are supposed to be able to master the mathematics contents and pedagogical knowledge so when they become teachers, they already can teach the contents well. In other words, teachers’ extensive knowledge of content and pedagogical content will influence their teaching quality which will have positive effect on their students’ achievement. However, in fact, teachers’ pedagogical knowledge related to mathematical contents is still far from the expectation, especially regarding the knowledge of how to design teaching strategies, materials, and to construct a lesson plan based on the chosen strategy, and conduct assessment in the process of learning so that the teaching process can be effective.

The development of cooperative learning is regarded to be appropriate for classes with large numbers of students and in the society that has a concept of communal work. Different from other commonly-used group discussion models, cooperative learning has specific basic elements and implementation procedures that enable a more effective class management (Beth A. H. et al., 2011). Cooperative learning model is included in the learning cluster of information and behavior processing (Lundgren, 1994), having principles on: 1) being positively interdependent, it means members in the group realize that they need to cooperate to achieve the goal, 2) all members are interacting, 3) the skills of cooperating and socializing, and in this case, their teachers’ guidance is necessary in order for the students to be able to work cooperatively and collaboratively, 4) the students need to assess the way they work effectively (Kellough & Kellough, 1999).

The facts found by the researchers in the implementation of teaching by using cooperative learning model are that there was not any self-responsibility because not all students got the opportunity to make a presentation, so the interaction process was dominated by a clever student, and the teachers performed the assessment separately. Therefore, it is necessary to develop a learning model which can increase student participation in discussions and presentations. Besides, the teachers should perform the assessments during the learning process. One of the efforts to increase students’ active participation in discussions and to have an assessment which is integrated with the learning process can be done by developing a learning model.
There is a paradigm change of teaching approach from teacher-centered into student-centered, and interaction shall functions as a catalyst. Garrison & Anderson, (2003), and Juwah, (2006), said that interaction is known as a basic element for a quality learning environment.

Interaction is conceptualized as a mechanism in which the students in a similar learning environment communicate one another and respond to each of their needs. Interaction can also be generalized to reflect various processes either intrapersonal, interpersonal, or to interact with technology (Hirumi, 2006). In the simple context, interaction can be defined as a form of involvement in learning (Topping, K. 1998). Interaction is also identified as a main component to create a more interesting learning atmosphere and as a catalyst to develop a learning community (Rhode, 2009). Peer assessment is appropriate to increase the interaction among students, between students and lecturers or teachers, and between students and media.

As an alternative method of assessment, the peer assessment is also used as a learning activity besides for assessing performance (Falchikove, N. Goldfinch, J,2000.). However, lecturers and students have some doubts about the validity level of the assessment made by the students with a similar level of knowledge (Sahin, S. 2008). Such issue seems to be the main barrier in the use of peer assessment. The results of research conducted by Mayes (2006) stated that if the peer assessment is correctly applied, it will show a result that is equal to the assessment performed by a lecturer.

The utilization of peer assessment learning model can be a supporting or complementary activity to increase the students’ active participation in discussions which are the main activities in cooperative learning (Silver H.F., Strong R.W.& Perini M.J., 2007). Peer assessment initiates with every student gives an assessment to their peers in one group and in another group. The assessment includes cognitive and affective domains when solving a given problem or topic. The cognitive domain assessment is done by assessing students’ accuracy and completeness in solving problems through a comprehensive evaluation on the students’ works. The evaluation is performed by the lecturer and the students’ peers, while the affective domain assessment is in the form of an assessment rubric for giving comments or opinions with the indicators of the quality of comments, quality of proposals, and the use of language.

The development of instrument for the learning model of peer assessment was expected to determine the level of peer assessment contribution in the evaluation of learning process done by the lecturers.

II. Research Method

This research is a part of learning model development through peer assessment focusing on developing the instrument for peer assessment. The instrument development procedure follows the procedure of learning model development, a set of prototypes and research instruments (Plomp, T. 2007). The phases of learning model development of peer assessment generally consist of (i) Introductory research, (ii) the making of prototype, (iii) final evaluation, and (iv) documentation and systematic reflection. To support the implementation of this learning model, it is necessary to develop a teaching set (materials, media/LKM, scenario, and assessment) and instruments (peer assessment rubric, scoring rubric, questionnaire and so on). The analysis of peer assessment was done in the following phases, a) conducting validity and reliability tests to score the peer assessment rubric, b) doing analysis of factors to determine the score weight of the peer assessment rubric, c) doing a regression analysis to the scores given from the lecturers and the score from peer assessment, d) identifying the distribution of weight scoring result. The subjects of this research were the Class G and H students who took Mathematical Statistics course, majoring Mathematics, from Academic Year of 2013/2014.

III. Result And Discussion

Validity and Reliability Test on Instrument

The results of validity and reliability tests to the score of peer assessment rubric from two classes based on the variables of problem solving, the completeness of problem solving, the quality of comments, the quality of proposals, and the use of language found in the instrument for peer assessment are as follows:

1. The result of validity test on class G showed a high level of validity with the lowest correlation value of 0.48 and the highest value of 0.866, but the reliability test result was still low with the value of Cronbach Alpha (0.238) until medium (0.361). It showed the necessity of initial explanation (short brief) about the procedure of scoring using a peer assessment rubric. The explanation of peer assessment was given to the other class (class H).

2. The result of validity test on class H showed a high level of validity with the lowest correlation value of 0.689 and the highest value of 0.961, again the reliability result shows the highest result with the value of Cronbach Alpha (0.850) until (0.865). It showed that the initial explanation (short brief) about the scoring procedure of a peer assessment rubric was already understandable.
The Analysis of Factor
The results of factor analysis by using five variables showed the overall value of the five variables, namely 43.668% (Table 1), consisting of accuracy (27.5%), completeness (21.4%), comments (54.8%), proposals (54.7%) and the use of language (54.9%) as shown in Table 3. Such results reflected that the students were still less competent (shown by the low percentage) in assessing their peers’ work (accuracy and completeness). It was considered to be reasonable because they did not have such experience before. However, in the aspects of comments, opinions and the use of language, the students tended to be able to do the assessment (the percentage is more than 50%). The value of KMO > 0.5 and the value of Bartlett’s test < 0.05 can be seen on Table 2. It means that the analysis of factors is feasible to be used in this research.

Table 1. The Eigen value and the percentage of factor analysis variance

<table>
<thead>
<tr>
<th>Component</th>
<th>Initial Eigenvalues</th>
<th>% of Variance</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.183</td>
<td>43.668</td>
<td>43.668</td>
</tr>
<tr>
<td>2</td>
<td>.982</td>
<td>19.634</td>
<td>63.302</td>
</tr>
<tr>
<td>3</td>
<td>.849</td>
<td>16.982</td>
<td>80.285</td>
</tr>
<tr>
<td>4</td>
<td>.555</td>
<td>11.109</td>
<td>91.394</td>
</tr>
<tr>
<td>5</td>
<td>.430</td>
<td>8.606</td>
<td>100.000</td>
</tr>
</tbody>
</table>

Extraction Method: Principal Component Analysis.

Table 2. KMO and Barlett’s Tests

<table>
<thead>
<tr>
<th>KMO and Bartlett’s Tests</th>
<th>Kaiser-Meyer-Olkin Measure of Sampling Adequacy.</th>
<th>Approx. Chi-Square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bartlett’s Test of Sphericity</td>
<td>.664</td>
<td>.000</td>
</tr>
<tr>
<td>Pred.</td>
<td>Df</td>
<td>Sig.</td>
</tr>
<tr>
<td>TOT_T</td>
<td>10</td>
<td>.734</td>
</tr>
<tr>
<td>TOT_L</td>
<td>10</td>
<td>.734</td>
</tr>
<tr>
<td>TOT_K</td>
<td>10</td>
<td>.734</td>
</tr>
<tr>
<td>TOT_U</td>
<td>10</td>
<td>.734</td>
</tr>
<tr>
<td>TOT_B</td>
<td>10</td>
<td>.734</td>
</tr>
</tbody>
</table>

The students’ less competence in assessing their peers is justifiable and is expected not to affect the assessments done by the lecturers. It is supported by the regression analysis result showing that there is not any influence of the objects assessed by the students to the lecturers’ assessment results (students’ competence in understanding the concept). It can be seen from the Anova and regression result with the F value of 0.555 (not significant) with the R² value of 5.4% (very low) (Table 4). Besides, the partial regression analysis result had significant variables. It means that there was not any relationship between the assessments performed by the students and the lecturers.

Table 4. The Value of R² and Anova analysis for regression

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.231</td>
<td>.054</td>
<td>-.043</td>
<td>19.68408</td>
</tr>
</tbody>
</table>

- Predictors: (Constant), TOT_B, TOT_T, TOT_L, TOT_K, TOT_U

- Regression Sum of Squares = 1074.740, df = 5, Mean Square = 214.948, F = .555, Sig. = .734

- Residual = 18985.696, df = 49, Mean Square = 387.463

- Total = 20060.436, df = 54

a. Dependent Variable: NILAI_G
b. Predictors: (Constant), TOT_B, TOT_T, TOT_L, TOT_K, TOT_U

Picture 1 presented that the average value of the peer assessment scores for all combinations (837 score combinations) was found to be 13.15. It means that the average value which can be contributed to the lecturers or teachers is only 13.15 (with the scale of 0 – 100). The lowest value was 8.67364 (The lowest contribution given to the lecturers or teachers was 9) and the highest score was 26.021 (The highest score given to the lecturers or teachers was 26). Based on Picture 1, it was also found that the scores from the students’ peers had a non-normal distribution with the value of Anderson Darling test of 13.57 (the p-value was less than 0.05, i.e. 0.005). The results of data distribution tests on some distributions (weibull, log-normal, exponential, log-logistic, weibul3p, log-normal 3p, exponential 2p, and log-logistic 3p) turned out that the scores from peer assessment followed the Log-Normal 3Pdistribution as can be seen in Picture 3 as it had the highest value correlation of 0.994.
The Development Of Peer Assessment Instrument For The Learning Model Of Peer Assessment

**Picture 1.** Descriptive Statistics results from 837 score combinations of peer assessment

**Picture 2.** Test Results of weibull, g-normal, exponential, and log-logistic distributions from the data of 837 score combinations for peer assessment

**Picture 3.** Test Results of weibull3p, log-normal 3p, exponential 2p, and log-logistic 3p distributions from the data of 837 score combination for peer assessment
Based on Picture 4, the results of log normal 3p distribution showed a location value of 1.002, scale parameter of 0.463, threshold parameter of 2.718 with the average value of 5.75 and standard deviation of 1.48. From such values, the difference between the lower class limit and upper class limit of the value of the peer assessment contribution to be combined with the value given by the lecturers or teachers.

IV. Closing

The result of factor analysis showed the overall value of the five variables amounted to 43.668% consisting of accuracy (27.5%), completeness (21.4%), comments (59.8%), suggestions (54.7%) and the use of language (54.9%). The regression analysis result pointed out that the students were still less competent (proven by the low percentage) in assessing their peers’ works (accuracy and completeness). However, in the assessment of comments, suggestions, and the use of language, the students tended to be able to do the assessment (the percentage was more than 50%). Furthermore, the results of the partial regression analysis demonstrated that there was not any relationship between the values given by the lecturers and the students’ variable assessment (no significant variable).

The results of simulation on 837 score combinations and weighting on each value showed a suitable distribution identification result from the value from peer assessment that was a lognormal (3p) distribution with the location parameter of 1.002, scale parameter of 0.463, and threshold parameter of 2.178. The average value was 5.75 and the standard deviation was 1.483. This result is used as the students’ contribution in the form of scores given through peer assessment to the evaluations of learning process and outcomes done by the lecturers.

References


DOI: 10.9790/7388-0703012025 www.iosrjournals.org
