# The Effect of Peer Assessment on Students' Performance in Writing Narrative Essays

Sonny Elfiyanto
Graduate School of Education – Hiroshima University
Hiroshima, Japan
elfiyanto@gmail.com

*Abstract*—**This study is an attempt to investigate the effect of peer assessment in improving students' performance in writing narrative essays at the university level. The participants were 95 first year students at the English department of a private university in Indonesia with mixed proficiency in English. During this study, they were divided into two different groups: control and experimental. Further, they are in the same semester and take the same course, which is ESP Writing. The control group received a teacher assessment of their writing, while the experimental group had peer assessment. Both groups were taught for eight meetings/weeks, including pre and posttest. For the duration of the treatment, the participants wrote one narrative essay every week. The students' writings in the pre- and posttest were analyzed to compare the relative effect of the two treatments, namely, peer assessment and teacher assessment only. The result revealed that peer assessment had a greater influence on improving students' ability in writing English essays than the teacher assessment method did.**

*Keywords: peer assessment, students' performance, writing*

## I. INTRODUCTION

Writing is now widely acknowledged that it has an essential role not only conveying information but also in transforming knowledge. Consequently, learning how to write is a challenging skill to master; therefore, some people, of course, still find that writing is burdensome. For some the first language (L1) students, writing is not an easy skill yet, even more to the second language (L2) or EFL students. The L2 and EFL students need more process than just finding the idea. They must think twice indeed: first, have to think of what to write, and then have to think of how to express their view in another language in which they still find difficulties in using it. Thus, [1] states that writing skill is an essential part of the curriculum in schools from the earliest grades onward and that most children in countries that have a formal education system will learn to write, at least at a basic level, in that setting. That is why writing is considered, still, an essential skill that must be mastered.

Writing in the second language is found to be different from the first language. The problems in second language writing are due to lack of several aspects. The first is about the correlation of understanding the source text or task instructions to writing. The lack of this matter will instigate poor performance in writing. The second is about linguistic proficiency. This matter plays a vital role in encoding ideas into a writing piece. So far, the second language writers still deal a lot with the searches for appropriate lexical and syntactic choices. Accordingly, the incorrect use of the proper linguistic components in writing will cause the differences between the content of the text in the writing products with the intention of the writers. Silva (as quoted in [2]) articulates that, in a review of differences between first and second language writing, writing in the second language tends to be "more constrained, more difficult, and less effective" than writing in a first language. The second language writers plan less, limited revision on content but more in grammar and write less fluently and accurately than first language writers.

Peer assessment tends to highly correlated with teacher assessment of the same students [3, 4] However, in addition to establishing comparability between teacher and peer assessment scores, it is important to determine whether peer feedback also has a positive formative effect on future academic performance.

Reference [6] propose peer assessment as a teaching approach which beneficial for lecturers to improve students' writing performance. Peer assessment is a process in which students can edit and respond to the writing of their peers. By applying peer assessment in the classroom, students can practice their writing and receive immediate feedback. Besides, they could interact with their classmate. Feedback from their peers can sometimes be more acceptable than the teacher's feedback.

Thus, this study aims to analyze the effect of peer assessment in improving students' performance in writing narrative essays at the university level. Then, by applying peer assessment, the problem in teaching English writing, especially in narrative essays, can be solved. Consequently, it requires a deeper research to identify whether there is any difference on students' performance in writing narrative essays between the students who are taught by using peer assessment and those by using teacher assessment.

### A. Writing

According to [7], writing is a form of communication to deliver through or to express through written form. Reference [8] said that writing is an activity to produce a sequence of sentences arranged in a particular order and linked together in

certain ways. Writing can be defined as a communicative act, the way of sharing observation, information, thoughts, or ideas with ourselves and others [9].

In short, writing is a form of communication that enables students/learners to express their feelings and ideas on a paper, to organize their understanding and views into reasonable reasons, and to convey meaning through well-constructed composition which has its own sets of the norm, rules, and conventions in a specific genre.

### B. Peer Assessment

According to [10], peer assessment is as an arrangement in which individuals consider the amount, level, value, and quality of success of the products or outcomes of learning through peer. In an EFL context, peer assessment has been commonly incorporated into English language writing instruction, because it involves the students in the learning process and develops their ability to reflect on and evaluate their learning and skill development. In this assessment, peers respond to and edit each other's written work. Thus, the purpose is helping with revision.

Further, [11] describes that peer assessment is an educational arrangement where students judge a peers' performance quantitatively and/or qualitatively and which stimulates students to reflect, discuss, and collaborate.

Peer assessment is defined as a collaborative learning activity in which peers participate in judging and assessing each other's work [12].

Reference [13] found that peer assessment training led to significantly more meaning changes and higher marks on L2 writers second drafts regardless of proficiency levels.

Peer assessment has been proved to have positive impacts on students' engagement in learning [14]. In peer assessment, peers use one another as a resource, by sharing ideas and evaluating the ideas of others, and providing feedback which can be qualitative (written or oral feedback) and/or quantitative (grades based on the criteria/scoring rubrics). Furthermore, peer assessment is a reliable vehicle of 'assessment for learning' because it actively involves students in evaluating their learning and allows them to participate in collaborative appraisal through the use of multiple perspectives when incorporating viewpoints from peers [15, 16].

The benefits of applying peer assessment have been discussed in several studies [17, 18]. Peer assessment is considered to enable students to develop their skills in writing. Also, some studies examining this type of assessment have revealed that it can work towards developing students' cognitive thinking. For example, peer assessment can increase students' interaction, confidence and critical thinking skills [19], help to nurture student-centered learning among undergraduate students [20], enhance relevant skills and interpersonal relationship between groups of students [21], and reduce the writing uneasiness level for the students [22].

In addition, peer assessment is beneficial for both lecturers and students. For the lecturers, peer assessment helps them save time in correcting their students' works. In one way or another, it helps them reduced their working load. According to [23], this type of activity saves lecturers' time on specific tasks, freeing them for more helpful instruction and is more at the

learner's level of development. However, it does not mean that lecturers are free from obligation for assisting students to learn and to write better. Nevertheless, this approach lessens their job, not eliminating.

### C. Narrative Essay

A narrative essay is a kind of text to retell the story in the past, which tells an interesting or imaginative story. Its social function is to tell stories or past events and entertain or amuse the readers. A narrative essay consists of the following structure: Orientation, introducing the participants and informing the time and the place, Complication, describing the rising crises which the participants to be involved in, and Resolution, showing the way of participant to solve the crises, better or worse, and Coda, the stepping back to evaluate the moral of the story or reorientation. However, this part is optional [24]. The examples of narrative texts are legend, fable, and fairy tales. There are some language features of a narrative essay: focus on specific participants, use of material for action processes, use of temporal conjunctions and temporal circumstances, and use Simple Past Tense.

## II. METHOD

### A. Participants

Those who took part in the study were 95 first year undergraduate students in English department at a private university in East Java, Indonesia and they had mixed proficiency level in English. There were 43 students in the experimental group and 52 students in the control group. They were grouped based on their own class. Therefore, the number could not be equal.

### B. Instruments and Procedure

The data were gathered through pre and posttest. The experiment was conducted for about eight times, including pre and posttest. Briefly, it was begun from February 19, 2019 to April 9, 2019. Every meeting lasted 100 minutes.

First, a pretest in a narrative essay was conducted at the beginning of the meeting. The pretest is intended to see the students' preliminary competence in writing. Besides, the pretest aimed to know that the experimental and control groups have no significant differences in their level of ability in writing a narrative essay. Students composed a narrative essay in 50 minutes about fairy tales or legends and could use a non-electronic dictionary.

In the experiment, the two different assessments were implemented during the treatment. The peer assessment was implemented to the experimental group and the teacher assessment was implemented to the control group.

During the treatment, the researcher and the writing lecturer were the teachers for both experiment and control group. In the experimental group, the first four meetings were taught by the researcher, while in the control group, the first four meetings were taught by the writing lecturer. After that, in the middle of the treatment, the researcher and the writing lecturer swapped with each other. The researcher taught the last four meetings of the control group while the writing lecturer taught the last four meetings of the experimental group. This

swap was based on the consideration that every change happening in the posttest for both experimental and control group did not have any relationship with the lecturer of the group. In other words, this was one of the researcher's efforts to control the possible threats during the treatment. Consequently, before conducting the treatment, the researcher and the lecturer had many discussions concerning the teaching and learning activities during the treatment for both experimental and control group.

The treatment for both groups was carried out in three steps: modeling, guiding, and practicing. For the second and the third meeting, the lecturer explained the generic structure of a narrative essay, gave the types of narrative essays, like fairy tales, fable, legend and science fiction and how to write a good essay in English. After that students were asked to practice by writing their narrative essays. The schedule of the treatment is presented in Table I.

### Table I. The Schedule of the Treatment

| MEETING | TREATMENT | DATE |
|---------|-----------|------|
| 1 | Pretest | February 19, 2019 |
| 2 | Modeling | February 26, 2019 |
| 3 | Modeling | March 5, 2019 |
| 4 | Guiding | March 12, 2019 |
| 5 | Guiding and Practicing | March 19, 2019 |
| 6 | Practicing | March 26, 2019 |
| 7 | Practicing | April 2, 2019 |
| 8 | Posttest | April 9, 2019 |

In the fourth and the fifth meeting, the lecturer explained the guiding step. In this step, the control group only learned how to understand teacher assessment and the scoring rubrics from Tribble which were applied during the study, while for the experimental group, the lecturer trained the students by giving models on how to assess writing essays and using the scoring rubrics. The researcher used Tribble's scoring rubric because this scoring rubric includes five areas, content, organization, vocabulary, language, and mechanics. Each area has different values, for content 20, organization 20, vocabulary 20, language 30, and mechanics 10. Thus, the total score will be 100.

Finally, a posttest of narrative essay was conducted in the eighth meeting for 50 minutes. The topic was about fairy tales or legends and the students could use a non-electronic dictionary.

### C. Data Analysis

The data, which were in the form of scores representing the students' writing achievement of the experimental and control groups, were analyzed. The scores were used to determine whether the mean scores obtained by the students in the two groups after treatment were significantly different. Therefore, ANOVA was applied to examine the effect of the peer assessment on students' performance in writing narrative essays.

### III.    FINDINGS AND DISCUSSION

### D. Giving the Pretest

The Pretest was conducted in the first meeting. Both groups were conducted on April 25, 2013, although the time was different because of the university's schedule. The time allotment for the pretest was scheduled as a class session that was 50 minutes. Based on the result of the pretest, the mean score for the experimental group was 56.6 while the mean score for the control group was 53.3.

### E. Giving the Posttest

The most important data in this study is the students' English narrative essay test score of the experimental and control groups from the posttest. After giving different treatments to both groups, a posttest was conducted to get the data of their writing test. The posttest of the experimental and the control group was held on April 9th, 2019. Shortly, the mean score of the experimental and control groups in the pretest and posttest score can be seen in Table II.

### Table II. The Results of the Posttest of the Experimental and Control Groups

| THE OBSERVED SCORE | Group | |
|---------------------|-------------|---------|
| | Experimental | Control |
| Highest observed score | 85 | 86 |
| Mean | 62.0 | 56.9 |
| Median | 59 | 55 |
| Mode | 58 | 52 |
| Lowest observed score | 40 | 32 |

Based on Table II, the highest score of the posttest of the experimental group was 85 and the control group was 86, the lowest score of the experimental groups was 40 and the control group was 32. Also, Table 4.1 indicates the mean score of the experimental group was 62.0 and the control group was 56.9. Related to the median of the posttest results, the experimental group was 59 and the control group was 55.

Briefly, the posttest results of the experimental group which received peer assessment and the control group which treated by teacher assessment should have different in terms of the highest score, the lowest score, the mean, the median, and the mode.

Figure 1 shows the difference of the posttest scores in terms of the mean score compared between the experimental group and the control group.
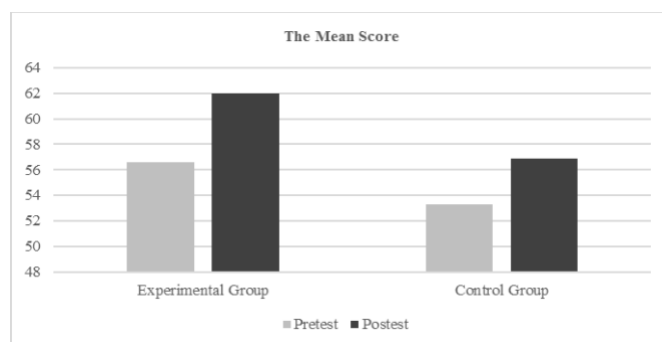


**Fig. 1 The Mean Difference Between the Experimental and Control Groups in the Pretest and Posttest**

## F. *Data Analysis Using ANOVA*

The data obtained in the posttest was computed by using Analysis of Variance (ANOVA). The data, which were in the form of, scores representing the students' writing achievement of the experimental and the control groups were analyzed. They were used to determine whether the mean scores obtained by the students in both groups after the treatment differed significantly.

**Table III. The Descriptive of the Posttest of the Experimental and Control Groups**

**Descriptives**

Score

| Group | N | Mean | Std. Deviation | Std. Error | 95% Confidence Interval for Mean | | Minimum | Maximum |
|---|---|---|---|---|---|---|---|---|
| | | | | | Lower Bound | Upper Bound | | |
| Pretest | 86 | 59.3605 | 15.05434 | 1.62335 | 56.1328 | 62.5881 | 25.00 | 88.00 |
| Posttest | 104 | 55.1346 | 14.43974 | 1.41593 | 52.3264 | 57.9428 | 21.00 | 88.00 |
| Total | 190 | 57.0474 | 14.83250 | 1.07606 | 54.9247 | 59.1700 | 21.00 | 88.00 |

From Table III, related to the descriptive of the posttest of both groups, peer assessment group got mean 59.36, while for teacher assessment group had mean 55.14. Before we go to ANOVA, the variants of the respondents must be equal. Next, it can be seen from Table IV below. Here, both groups have the same p-value (0.932). Therefore, since it was higher than the alpha 0.05, ANOVA can be conducted.

**Table IV. The test of Homogeneity of Variables**

**Test of Homogeneity of Variances**

Score

| Levene Statistic | df1 | df2 | Sig. |
|---|---|---|---|
| .007 | 1 | 188 | .932 |

To know the difference between the two groups, it was shown in Table V. If you see column Sig., the p-value is 0.050, with the level of confidence 95% ($\alpha = 0.05$). Thus, the p-value was smaller than the alpha 0.05. In other words, there is a difference in the students' writing English narrative essays taught by using peer assessment and teacher assessment.

**Table V. The Computation of the Statistical Computation by means ANOVA**

**ANOVA**

Score

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Between Groups | 840.633 | 1 | 840.633 | 3.879 | .050 |
| Within Groups | 40739.941 | 188 | 216.702 | | |
| Total | 41580.574 | 189 | | | |

The improvement between the two groups can be seen from the mean of their posttest, increased point from pretest to posttest, and how many percentages of the students who enhanced their score. The peer assessment group had better score compared to the teacher assessment. Before they are guided by peer assessment, the experimental group got mean 56.6; however, after they are taught by peer assessment, they could reach 62.0. Overall, their mean improved 5.4 points. Meanwhile, for teacher assessment group, they had 53.3, then for the posttest, they got 56.9. Thus, their score increased by 3.6 points. Next, their scores were also improved, since It was about 74.4% of the students could improve their score after the treatment (there were 32 of 43 students).

Finally, peer assessment has a contribution in improving students' performance in writing, although in EFL country like Indonesia in which the lecturers are considered to be persons who are "know everything", peers' opinion still cannot be neglected.

## IV. CONCLUSION

The result of this study showed that the students who were taught by peer assessment had a better score (performance) than the students who were taught by teacher assessment. Therefore, the use of peer assessment is effective to be used for improving students' performance in writing English narrative essays, especially in EFL country like Indonesia. It suggests that peer assessment also can be applied in other EFL countries which have a similar condition like Indonesia, for example, Japan or Korea.

Besides during the observation, peer assessment can help the lecturers save time and make them and the students to be more active in the teaching learning process.

Afterward, during the treatment, the main attention of the researcher is how to train the students to use the scoring rubrics and try to give the score to their classmates' work, since they were not so confident to check and score their friends' work. Future researchers should focus on making the students feel confident in scoring their classmates' work.

The result of this research can be used as a further consideration to conduct additional research dealing with the application of peer assessment in different kinds of education level, text types, language skills, research area and research design.

## REFERENCES

[1] S. C. Weigle, Assessing Writing. Cambridge: Cambridge University Press, 2002.

[2] K. Hyland, Second Language Writing. Cambridge: Cambridge University Press, 2010.

[3] N. Falchikov, and J. Goldfinch. "Student peer assessment in higher education: a meta-analysis comparing peer and teacher marks" in Review of Educational Research, vol. 70, no. 3, 2000, pp. 287-322.

[4] C. C. Liu, K.H. Lu, L. Y. Wu, and C. C. Tsai, "The impact of peer review on creative self-efficacy and learning performance in Web 2.0 learning activities", in Journal of Educational Technology & Society, vol. 19, no. 2, 2016, pp. 286-297.

[5] M. van Zundert, D. Sluijsmans, and J. van Merriënboer, "Effective peer assessment processes: research findings and future directions", in Learning and Instruction, vol. 20, no. 4, 2010, pp. 270-279.

[6] D. Spiller, "Assessment matters: self-assessment and peer assessment", Hamilton, New Zealand: Teaching Development, The University of Waikato, 2012, pp. 1-19.

[7] J. Harmer, The Practice of English Language Teaching. Essex: Pearson Education Limited, 2001.

[8] D. Byrne, Teaching Writing Skills. Longman Group Ltd, 1982.

[9] L. Cohen, and M. Riel, "The effect of distant audiences on students' writing", in American Educational Research Journal, vol. 26, no. 2, 1989, pp. 143-159.

[10] Topping, K. "Peer assessment between students in colleges and universities", in Review of Educational Research, vol. 68, no. 3, 1998, pp. 249-276.

[11] J. W. Strijbos, and D. Sluijsmans, "Unravelling peer assessment: methodological, functional, and conceptual developments", in Learning and Instruction, vol. 20, no. 4, 2010, pp. 265-269.

[12] I. Jones, and L, Alcock, "Peer assessment without assessment criteria", in Studies in Higher Education, vol. 39, no. 10, 2014, pp. 1774-1787.

[13] E. C. Berg, "Preparing ESL students for peer response", in TESOL Journal, vol. 8, no. 2, 1999, pp. 20-25.

[14] K. Topping, (2003). "Self and peer assessment in school and university: reliability, validity and utility", in M. Segers, F. Dochy, and E. Cascallar (Eds.) *Optimising new modes of assessment: In search of qualities and standards*, Dordrecht: Kluwer Academic Publishers, pp. 55-88.

[15] N. Falchikov, "Involving students in assessment", in Psychology Learning *& Teaching*, vol. 3, no 2, 2004, pp. 102-108.

[16] E. Panadero, (2016). "Is it safe? Social, interpersonal, and human effects of peer assessment: a review and future directions", in G. T. L. Brown and L. R. Harris (Eds.), *Handbook of human and social conditions in assessment*, New York: Routledge, 2016, pp. 247–266.

[17] K. P. Kwan, and R. Leung, "Tutor versus peer group assessment of student performance in a simulation exercise", in Assessment & Evaluation in Higher Education*, vol. 21, no. 3, 1996, pp. 205-214.

[18] D. Reinholz, "The assessment cycle: a model for learning through peer assessment", in *Assessment & Evaluation in Higher Education,* vol. 4, no. 2, 2016, pp. 301-315.

[19] R. Ruegg, "The relative effects of peer and teacher feedback on improvement in EFL students' writing ability", in *Linguistics and Education*, vol. 29, 2015, pp.73-82.

[20] K. Oldfield, and M. MacAlpine, "Peer and self-assessment at tertiary level – an experiential report", in Assessment & Evaluation in Higher Education, vol. 20, no. 1, 1995, pp. 125-132.

[21] T. Issa, "Promoting learning skills through teamwork assessment and self/peer evaluation in higher education" in IADIS International Conference on Cognition and Exploratory Learning in Digital Age. CELDA. 2012.

[22] J. H. Jahin, "The Effect of Peer Reviewing on Writing Apprehension and Essay Writing Ability of Prospective EFL Teachers", in Australian Journal of Teacher Education, vol. 37, no. 11, 2012.

[23] C. L. Keh,"Feedback in the writing process: a model and methods for implementation", in ELT Journal*, vol. 44, no. 4, 1990, , pp. 294-304.

[24] R. Hartono, Genres of Texts. Semarang: Semarang State University, 2005.