

## Does Student Self-Assessment Assess as Valid and Reliable as Teacher Assessment?

**Abdul Muth'im**

English Education Study Program  
Faculty of Teacher Training and Education  
Lambung Mangkurat University, Banjarmasin  
South Kalimantan, Indonesia

### Abstract

Assessment is an integral part of instruction. As one form of assessment, self-assessment is relatively rarely used in writing class. This might be because its validity and reliability is still in question. This research aims at finding (1) whether students assess their writing tasks the same as their teacher does, (2) if they assess differently, why they do it differently. Twenty four students who were taking Writing IV course were involved in this study. The research method used was a mix of quantitative and qualitative descriptive. For one semester the students were taught to write an essay in English. Strategies on how to get an idea, to organize idea, to implement proper language use, to choose appropriate vocabulary, and to use mechanics (e.g. punctuation, spelling and capitalization) were intensively discussed. Along with the activities they were encouraged to implement the feedback received in their writing tasks. Finally, at the end of semester, an assessment was done. The writing task was not merely assessed by the teacher. The students were also given trust and responsibility to assess their own writing tasks. The results of the two assessments were compared. It was found that, in general, the results of assessment represented by scores given by the students did not much differ significantly from the scores given by the teacher. This implied that self- assessment can be as valid and reliable as teacher-assessment if students were properly and adequately trained by teacher.

**Key words:** essay, score, self-assessment, teacher assessment

**Cite as:** Abdul Muth'im, A. (2017). Does Student Self-Assessment Assess as Valid and Reliable as Teacher Assessment? *Arab World English Journal*, 8 (1).

DOI: <https://dx.doi.org/10.24093/awej/vol7no1.9>

## Introduction

Assessment is an integral part of teaching and learning program. As an integral part of teaching and learning, assessment may be used for different purposes, such as (a) checking whether the goal or the objective(s) of teaching and learning is/are attained, (b) determining whether the teaching and learning materials are suitable for certain group of students, whether the methods and techniques of teaching and learning are appropriate for the students, and whether the learning experience provided are stimulating learning. In other words, assessment can be used as a means of determining whether instructional objective(s) or purpose(s) are achieved. Wiliam (2013) believes that “it is only through assessment that we can discover whether the instructional activities in which we engaged our students resulted in the intended learning” (p. 15). New Zealand Ministry of Education (2007, p. 39 in Earl & Giles, 2011) state that “the primary purpose of assessment is to improve students’ learning and teachers’ teaching as both student and teacher respond to the information that it provides” (Earl & Giles, p. 12). Boud (1995 in Spiller, 2012) argued that all assessment including self-assessment comprises two main elements: making decisions about the standards of performance expected and then making judgments about the quality of the performance in relation to these standards.

Assessment is understood in a variety of definitions. The New Zealand Ministry of Education, as reported in *The Literacy and Numeracy Secretariat Capacity Building Series*, define assessment as the process of gathering information ... from a variety of sources that accurately reflect how well a student is achieving the curriculum expectation in a subject (2012). Joice, et al. (2009) define assessment as collecting information on student learning or performance based on various sources of evidence. Miller (2008, p. 2) defines assessment “as a broader term than tests and encompasses the general process of collecting, synthesizing, and interpreting formal and informal measurement. Assessment provides information about individual or group over time”.

Assessment, in accordance with Gronlund and Waugh (2009) can be administered in different phases of instruction. It can be administered at the beginning of instruction, during instruction and at the end of instruction. The first is called *placement assessment*, the second one is called *formative and diagnostic assessment*, and the last is called *summative assessment*. At the beginning of instruction, the *questions* that need to be answered are to what extent do the students possess the skills and abilities that are needed to begin instruction, and to what extent have the students already achieved of the planned instruction. During instruction, the questions that must be answered are on which learning tasks are the students progressing satisfactorily? On which ones do they need help?, and which students are having such severe learning problems that they need remedial work? At the end of instruction, the questions that must be answered are which students have mastered the learning tasks to such a degree that they should proceed to the next course or unit of instruction, and what grade should be assigned to each student?

Earl & Giles (2011) group assessment into three categories: assessment of learning, assessment for learning, and assessment as learning. Assessment of learning, which is also called summative for assessment purposes, in accordance with Krause, et al. (2003, in Earl & Giles, 2011) refers to ‘formal checks of learning outcomes that are conducted at the end of a teaching program. Assessment for learning, according to Black & William (1998), is defined as a range of informal and formal procedures undertaken by teachers as an integral part of the normal teaching

and learning. The information obtained via these procedures is used to modify and enhance learning and understanding. Whereas assessment as learning is comprehended as learning concept underscores that students should be valued participants in their own learning, anticipate receiving and utilizing constructive feedback and feed-forward and be able to identify their own learning gaps and solve their learning needs, with teacher assistance (Black & William, 2008 in Earl & Giles, 2011),.

Seen from its purpose, assessment is intended to (1) evaluate and improve student learning, (2) identify student strength and weaknesses, (3) assess the effectiveness of a particular instructional strategy, (4) evaluate and improve the effectiveness of curriculum programs, (5) evaluate and improve teaching effectiveness, and (6) communicate with parents and guardians and involve them in their children's learning (Kellough, 1993).

### *Self-assessment*

Self-assessment, in accordance with Andrade & Du (2007), is defined as a process of formative assessment during which is a process students reflect on and evaluate the quality of their work and their learning, judge the degree to which they reflect explicitly stated goals or criteria, identify strengths and weaknesses in their work, and revise accordingly (in Spiller, 2012). McMillan & Hearn (2008), in the other hand, propose that self-assessment is more accurately defined as a process by which students 1) monitor and evaluate the quality of their thinking and behavior when learning and 2) identify strategies that improve their understanding". For Joyce, et al. (2009), self-assessment takes place if the student him/herself is involved in some or all aspects of assessment process. McMillan & Hearn (2008), suggest that self-assessment occurs when students judge their own work to improve performance as they identify discrepancies between current and desired performance (in Price, et al., 2011). Kayler & Weller, (2007); Mok et al., (2006) claim that self-assessment develops the self-judgmental ability of students through analysis of their own work in class and at home (in Amo & Jareño, 2011). Joice et al., (2009) argue that self-assessment is crucial for students because they not only have to gather evidence of their learning, but also analyze their work in terms of the goal/standard, make decisions about what they need to do to improve, know what to do to close the gap, and monitor their progress towards achieving this. Of course, the involvement of students in assessment depends very much on the teacher. If the teacher is committed to the learner-centered principles, he/she will be willing to get students involved in the assessment. Kayler & Weller (2007) state "Commitment to the learner-centered led us to involve students in the assessment of the pedagogy" (p. 136).

So far the most common means of assessment is the assessment done by teacher. That is why it is called teacher assessment. Yet, teacher assessment is not the only means which is claimed to be valid and reliable in assessing students' learning. There are other means of assessment which can also be used to assess students' learning, for instance, peer assessment and self-assessment.

Self-assessment, for example, is relatively rarely used in teaching and learning program. It is not known for sure why this kind of assessment is not used as a means of assessing students' learning. It might be related with its validity and reliability which are still in question. It is based on this phenomenon that this research is carried out. The questions that should be answered in

this study are: 1) Do students assess their writing tasks the same as their teacher does?, 2) if they assess their writing tasks differently, why do they do that?

### *Review of previous studies*

Self-assessment has proven itself as an effective means of assessing students' learning. The results of a number of studies show this effectiveness. Noonan and Duncan (2005), in their study, for example, uncovered that peer and self-assessment useful and there is a potential for greater classroom applicability. The study carried out by Zheng, et al. (2012) reveal the following: (1) students could perform self-assessment in writing reasonably well, (2) the instruction of scoring rubric contributed to the improvement of self-assessment in writing and the overall improvement was significant (3) students' overall composition performance was enhanced. Price, et al. (2011) quoted various results of research on self-assessment. For instance, they quoted the studies conducted by H. Andrade & Valtcheva (2009); by Klenowski (1995); by McMillan & Hearn (2008). The results of those quoted research reveal that self-assessment can have positive effect on achievement, motivation, self-perception, communication, and behavior (in Price, et al., 2011). The other result of study quoted was the one conducted by McDonald & Bound (2003) who found that high school students who were trained in self-assessment not only felt better prepared for their external examinations, they actually outperformed their peers who had not received the training. The third result of study quoted was the one conducted by Ross (2006) who found that students across grade levels and subject areas including narrative writing, mathematics and geography outperformed their peers in the control group who had not received self-assessment training (in Price, et al. (2011). McMillan & Hearn (2008) argue that if it is correctly implemented, student self-assessment can promote intrinsic motivation, internally controlled effort, a mastery goal orientation, and more meaningful learning". Azorín (1991) argues that self-assessment has a number of additional advantages related both to the effective implication of students in introspecting about their learning processes and to students' participation in class management.

In summary, from what have been discussed above, Spiller (2012) drew conclusions that self-assessment had a number of advantages: for example, (a) the natural tendency to check out the progress of one's own learning is built, (b) the recognition of what is needed to be learned is only possible after further learning, (c) if a student can identify his/her learning progress, further learning may be motivated, (d) reflection on one's own learning is encouraged, (e) learner responsibility and independence can be promoted, (f) student's ownership of learning tasks be encouraged, (g) the focus of the tasks is shifted from something imposed by someone else to a potential partnership, (h) the formative aspects of assessment is emphasized, (i) a focus on process is encouraged, (j) diversity of learners' readiness, experience and backgrounds can be accommodated, (k) align well with the shift in the higher education literature from a focus on teacher performance to an emphasis on student learning is practiced, and (l) students in the formulation of criteria for self-assessment tasks are engaged to help them to deeper their understanding of what constitutes quality outcomes in a special area.

### *Assessing writing*

Writing may be understood in different ways. Copying is the most basic and is considered as the simplest form of writing. Whereas composing an essay is considered to be the most difficult and complicated one. This can be implied from the definitions of writing given by

some experts for the latter. Byrne (1984), for instance, defines writing as the production of a sequence of sentences arranged in a particular order and linked together in certain ways. For Troyka (1987), writing is understood as a way of communicating a message to a reader for a purpose. Hammond (1989), on the other hand, defines writing as a complex task which requires everything from getting your spelling right to making your voice distinctive enough to be heard. Based on the definitions, it may be concluded that writing is a way of communicating ideas, opinions, feelings, wants, or wishes in the form of written language, called composition. To be more specific, it is the composition of essay.

In higher education context, the most common genre that second language learners usually produce is the genre of academic writing. Academic writing in accordance with Oshima & Hogue (2006) is the kind of writing used in high school and college classes. The kinds of writing that belong to this genre, according to Brown (2004), includes: papers and general subject report, essays, compositions, academically focused journal, short-answer test responses, technical reports, theses, and dissertation. The reasons why the teaching of writing in higher education context are emphasized on academic writing are (a) it does many of the things that personal writing does not such as its structure in which there should be beginning, middle, and end, (b) it is based on citation of published authors, (c) there are always rules of punctuation and grammar to be followed, and (d) academic topics traditionally focus on abstract things, like ideas and concepts (Bowker, 2007). In the context of genre-based writing teaching, Hyland (2007) claims “... people don't just write, they write something to achieve some purpose: writing is a way of getting things done” (p. 5). This means that there are certain social conventions that should be followed in organizing the messages. Hyland (2007) argues that these conventions can be described and taught.

There are three scales used to assess writing assignment, namely primary trait scales, holistic scales, and analytic scales (Weigle, 2002). Primary trait scoring, in accordance with Weigle (2002) “lies on the philosophy that understanding how well students can write can be assessed through a limited range of discourse, e.g. persuasion or explanation” (p. 110). She then continues “in primary trait scoring, the rating scale is defined with respect to the specific writing assignment and essay is judged according to the degree of success with which the writer has carried out the assignment”. For this purpose, Weigle suggests the rater to create a scoring rubric in which the writing task, a statement of the primary rhetorical trait, a hypothesis about the expected performance on the task, a statement of the relationship between the task and the primary trait, a rating scale which articulates levels of performance, sample scripts at each level, and explanation of why each score was given should be included (Weigle, 2002).

Holistic scoring scale in accordance with Weigle (2002) is the way of assigning of a single score to a script based on the overall impression of the script. In quite different formulation Ferris & Hedgcock (2005) define holistic scoring as the way to rate or rank writing proficiency as reflected in a given sample based on the scoring rubrics provided. There are some advantages in using holistic scoring scale. Weigle (2002) argues that holistic scoring is faster (and therefore less expensive) to read a script once and assign a single score than to read it several times, each time focusing on a different aspect of writing. In addition to the above advantage, holistic scoring scale has another strength, that is, it is intended to focus the reader's attention on the strength of the writing, not on its deficiencies. However, holistic scoring scale



also has some disadvantages. For example, in the second language context, a single score does not provide useful diagnostic information about a person's writing ability, whereas, this information is needed to improve students' learning. The reason is it does not allow raters to distinguish between various aspects of writing. The other disadvantage is that holistic scores are not always easy to interpret.

Analytic scoring in accordance with Weigle (2002) is a way of giving score to a script based on several aspects of writing or criteria rather than given a single score. Jacobs et al. (1981) suggest to use what they call ESL Composition Profile. This guideline suggests raters to focus on five components of writing, namely: content, organization, vocabulary, language use, and mechanics. In an analytic scoring system, the rater relies on a rating guide that separates and weights textual components a priori: criteria are prioritized before scoring begins (Ferris & Hedgcock, 2005). In other words, the instructor gives score based on the components of writing. Of course, this type of assessment is more complicated and takes longer time because the rater should give score to each of the components of writing. In the context of second language, the major benefit of this assessment is that it provides more useful diagnostic information about students' writing abilities (Weigle, 2002). Al Makhzoomi & Farheit (2011) in their study found that the analytic method has pedagogical advantages over the impressionistic one in that it lays the foundation of the relevant elements of good writing. The major disadvantage of analytic scoring in accordance with Weigle (2002) is that it takes longer time than holistic scoring because readers are required to make more than one decision for every script. Besides, if scores on different scales are combined to make a composite score, a good deal of information provided by the analytic scale is lost, Weigle (2002) continues the argument.

Based on the discussion above, considering the advantages and disadvantages of each type of writing assessment, the present research employs analytic scoring approach. Assessing students' writing by using analytic scoring approach gives many advantages for the students because they will know what their strengths and also their weaknesses are in their writing. By being corrected this way they will know in what component(s) they need to practice a lot and improve their weakness.

## **Method**

### ***Subject***

By employing quantitative and qualitative descriptive method 24 Indonesian university students of English education study program who were taking Writing IV course consisting of 9 male students and 15 female students were involved in this study were. They were purposively chosen because they met the pre-determined characteristics of the subjects, that was, the students had already passed some pre-requisite courses, namely, Writing III course, Structure III course, and Reading III course. Those were pre-requisite courses for the students to write thesis proposal. By having passed in these courses, the students are assumed to have adequate preparation to write an essay.

### ***Research procedure***

This research was conducted to find out (1) whether students assess their writing tasks the same as their teacher does, (2) if they assess differently, why they do it differently. Since this study was carried out in classroom context, for the purpose of study, the procedure of research

was relatively similar to the one usually carried out in normal situation. The procedure can be described as follows.

#### 1. Distributing questionnaire phase

By referring to Jacob's et al. (1981) writing profile, this questionnaire was used to get an overall picture of students' problem(s) in writing essay in English. Specifically, it was conducted to find out whether they had problem(s) in getting and developing idea as the content of their essay, in organizing idea(s) in terms of unity and coherency, expressing idea(s) in term of language use, in choosing appropriate word(s) for certain context and situation, and implementing their knowledge of mechanics (e.g. punctuation) in their essay correctly. (Appendix A)

#### 2. Introducing scoring guide phase

In this phase, a guide for giving score of writing tasks was introduced. The scoring guide which was adapted and adopted from Jacob's et al. (1981) covered five areas: (1) how to assess content, (2) how to assess organization, (3) how to assess language use, (4) how to assess vocabulary, and (5) how to assess mechanics. (Appendix B)

#### 3. Reviewing phase

In this phase, lead by the teacher, the students were asked to assess their essays. First, they had to assess the content to find out whether it was relevant with the task of writing. Second, they should assess the organization of the essay to make sure that the essay was well organized both in terms of sentence level and paragraph level. Third, they had to assess the grammar used in the essay to make sure that it reflected the meaning(s) intended. Fourth, they should assess the words used in the essay whether or not the choice was appropriate. Fifth, they were asked to assess whether the mechanics used (e.g. punctuation, spelling) were rightly implemented.

#### 4. Scoring phase

In this final phase, the students were asked to give score to their writing tasks by filling out scoring sheet provided. (Appendix C)

### Findings

From the question distributed it was found out that the problems faced by the students in writing essay in English can be described in the table 1.

Table 1. *Problems faced by students in writing essay in English*

No	Components	Rate										Total	
		Always		Most of the time		Rarely/ Seldom		Almost never		Never			
		5		4		3		2		1			
		f	%	f	%	f	%	F	%	f	%		
1	Content	2	8.33	12	50	7	29.17	3	12.50	0	0	24	100

2	Organization	2	8.33	12	50	8	33.33	2	8.33	0	0	24	100
3	Language use	6	25	12	50	4	16.67	2	8.33	0	0	24	100
4	Vocabulary	16	66.67	3	12.5	3	12.5	2	8.33	0	0	24	100
5	Mechanics	0	0	1	4.16	2	8.33	20	83.33	1	4.16	24	100

Table 1 revealed that those who always had problem in content were 2 students (8.33%), those who had problem most of the time in content were 12 students (50%), those who rarely/seldom had problem in content were 7 students (29.16%), those who almost never had problem in content were 3 students (12.50%), and those who never had problem in content was none (0%).

Table 1 also discovered that those who always had problem in organization were 2 students (8.33%), those who had problem in content most of the time were 12 students (50%), those who rarely/seldom had problem in organization were 8 students (33.33%), those who almost never had problem in organization were 2 students (8.33%), and those who never had problem in organization was none (0%).

Table 1 also found that those who always had problem in language use were 6 students (25%), those who had problem in language use most of the time were 12 students (50%), those who rarely/seldom had problem in language use were 4 students (16.67%), those who almost never had problem in language use were 2 students (8.33%), and those who never had problem in language use was none (0%).

Table 1 also uncovered that those who always had problem in vocabulary were 16 students (66.67%), those who had problem in vocabulary most of the time were 3 students (12.50%), those who rarely/seldom had problem in vocabulary were 3 students (12.50%), those who almost never had problem in vocabulary were 2 students (8.33%), and those who never had problem in language use was none (0%).

Table 1 also showed that those who always had problem in mechanics was none (0%), those who had problem in mechanics most of the time was 1 student (4.16%), those who rarely/seldom had problem in mechanics 2 students (8.33%), those who almost never had problem in mechanics 20 students (83.33%), and those who never had problem in mechanics use was 1 student (4.16%).

How the students scored their writing tasks and how the teacher scored his students' writing tasks can be seen in table 2 and table 3. (Appendix D & Appendix E).

From table 2 and table 3 it was found that in general both the students and the teacher scored the writing tasks relatively the same. The mean scores of the two assessments: student self- assessment and teacher assessment were consecutively as follow, 73.92 and 73.75. Though they were different, the difference was not significant. Both scores in the system of evaluation of the institution fell onto grade B. Out of 24 students, 22 persons (91.67%) gave the same score as



the teacher did. Only 2 students (8.33%) gave different scores. One student scored her writing task under the score given by the teacher. The other student scored his writing task above the score given by the teacher. When asked, the student (FAU) who scored her writing task lower than the score given by the teacher told that she did not deserve to get more than that score, i.e. 71 (B). She thought that she still found some problems in her writing essay in English. On the contrary, her teacher thought that she deserved to get higher score than that – she had the right to get  $\geq 80$  (A). The teacher believed that his assessment about FAU was not far from her real capability. This was proven by the fact that she has been chosen as one of the speakers representing her university in English language education seminar held in one private university in Bali, Indonesia, on November 2014.

On the other hand, the male student, MSA, when interviewed, honestly acknowledged that actually he did not deserve to get score 80 (A). He confessed that he scored his writing task that way because he needed to upgrade his grade point average (GPA, or *Indek Prestasi Kumulatif, IPK*, in Indonesian). He argued that high GPA is needed in finding job. However, this has justified the teacher's assessment which gave him score lower than 80 (A), i.e. 66 (C+). There has not been any more writing essays in English he produced. The complete comparison of scores given by the students and the teacher can be seen table 4. (Appendix F)

## Discussion

As discussed earlier that there are two questions that this study wanted to answer. They were: 1) Do the students assess their writing tasks the same as the teacher does?, 2) If they assess their writing tasks differently from what the teacher does, why do they score their writing tasks differently?

The results of study revealed that out of 24 students, 22 students (91.67%) gave the same scores as what the teacher did to their writing tasks. This implied that both the students and the teacher had the same understanding about the scoring guide and interpretation on students writing tasks. What Noonan and Duncan (2005) discovered about the usefulness and the potential for greater classroom applicability of peer and self-assessment was reflected in present study. What Zheng, et al. (2012) found that students could perform self-assessment in writing reasonably well, the instruction of scoring rubric contributed to the improvement of self-assessment in writing and the overall improvement was significant, and students' overall composition performance was enhanced were also reflected in present study. The report written by Price, et al. (2011) on the results of some studies on self-assessment may also support the results of the present study. The results of study conducted by H. Andrade & Valtcheva (2009); by Klenowski (1995) and by McMillan & Hearn (2008) as quoted by Price, et. Al., (2011) which found that self-assessment can have positive effect on achievement, motivation, self-perception, communication, and behavior also strengthened the result of present study. The results of other study conducted by McDonald & Bound (2003) as quoted by Price, et al. (2011) which found that high school students who were trained in self-assessment not only felt better prepared for their external examinations, outperformed their peers who had not received the training may also show the benefit of self-assessment. Also, the study conducted by Ross (2006) which found that students across grade levels and subject areas including narrative writing, mathematics and geography outperformed their peers in the control group who had not received self-assessment training also supported the current study (reported by Price, et al., 2011). The results of present

study may represent what McMillan & Hearn (2008) claimed that if it is correctly implemented, student self-assessment can promote intrinsic motivation, internally controlled effort, a mastery goal orientation, and more meaningful learning". The last results of study that supported the present study was the one carried out by Azorín (1991) which argues that self-assessment has a number of additional advantages related both to the effective implication of students in introspecting about their learning processes and to students' participation in class management.

### Conclusion

The findings and the theory which support the findings brings us to a conclusion: (1) in general, the students assessed their writing tasks relatively the same as what their teacher did, i.e. 73.92 compared to teacher assessment, i.e. 73.25 (2) there were two students who assessed their writing tasks differently, one student scored lower than the teacher's score and the other one scored his writing task higher than the score given by the teacher. The student who scored her writing tasks lower than the score given by teacher is actually an example of student who values knowledge and skill of writing are more important than the score. On the contrary, the student who scored his writing tasks higher than the score given by the teacher is also an instance of student who values knowledge and skill of writing is less important than score.

### Recommendations

Based on the findings it is recommended for the teacher to give responsibility for the students to assess their own learning tasks. The results of this study has proven that students assessment can be as valid and reliable as teacher assessment with one condition that students get proper and appropriate trainings done by teacher. For future researcher it is recommended to find out whether writing tasks assessed by other kinds of assessment, such as peer assessment improve student writing tasks.

### Biography of the Author

**Dr. Abdul Muth'im** got his Doctor's degree in English Education from State University of Malang, Indonesia. His main interest of teaching and research is TEFL especially writing. Until recently he has written three books: two books on writing and one book on TEFL.

### Reference

- Al Makhzoomi, K. & Freihat, S. (2011). Analytic or Impressionistic Evaluation of EFL Jordanian Students' Composition Performance? *AWEJ* Vol. 2, 3, 141-167.
- Amo, E. & Jareño, F. (2011). Self, Peer and Teacher Assessment as Active Learning Methods. *Research Journal of International Studies*. Issue 18.
- Azorín, María José Martínez. (1991). Self-Assessment in Second Language Teaching: Journals. *Revistá Alicantina de Estudios Englises*. 4, 91-101.
- Bowker, Natile (Ed). (2007). *Academic Writing A Guide to Tertiary Level Writing*. Massey: Massey University.
- Brown, H. Douglas. (2004). *Language Assessment: Principles and Classroom Practices*. White Plain, NY: Pearson Education. Inc.
- Byrne, Donn. (1984). *Teaching Writing Skills*. England: Longman Group Limited.
- Earl, K. & Giles, D. (2011). An-other Look at Assessment: Assessment in Learning. *New Zealand Journal of Teachers' Work*. Volume 8, 1, 11-20.
- Ferris, D. R. & Hedgcock, J. S. (2005). *Teaching ESL Composition*. Mahwah, New

- Jersey: Lawrence Erlbaum Associates, Publishers.
- Gronlund, N. E. & Waugh, C. K. (2009). *Assessment of Student Achievement*. New Jersey Columbus, Ohio: Pearson Education, Inc.
- Hammond, Eugene R. (1989). *Critical Thinking, Thoughtful Writing*. New York: McGraw-Hill, Inc.
- Hyland, Ken. (2007). *Genre and Second Language Writing*. Ann Arbor: The University of Michigan Press.
- Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F. & Hughey, J. B. (1981). *Testing ESL Composition: A Practical Approach*. Rowley, Massachusetts: Newbury House Publishers, Inc.
- Joyce, C., Spiller, L., & Twist, J. (2009). *Self-assessment: What Teachers Think*. Wellington: New Zealand Council for Educational Research.
- Kayler, M. & Weller, K. (2007). Pedagogy, Self-Assessment, and Online Discussion Group. *Educational Technology & Society*. 10, 1, 136-147.
- Kellough, R. D., Kellough, N. G., & Hough, D. L. (1993). *Middle School Teaching Methods and Resources*. New York: Macmillan Publishing Company.
- McMillan, J. H. & Hearn, J. (2008). Student Self-Assessment: The Key to Stronger Student Motivation and Higher Achievement. *Educational HORIZON*. 40-47.
- Noonan, B. & Duncan, C. R. (2005). Peer and Self-Assessment in High Schools. *Practical Assessment, Research & Evaluation*. 10, 17, 1-8.
- O'Malley, J. M. & Pierce, L. V. (1996). *Authentic Assessment for English Language Learners*. New York: Addison-Wesley Publishing Company.
- Oshima, A. & Hogue, A. (2006). *Introduction to Academic Writing*. Fourth Edition. White Plain, NY: Pearson Education, Inc.
- Price, J. K., Pierson, E. & Light, D. (2011). *Using Classroom Assessment to Promote 21<sup>st</sup> Century Learning in Emerging Market*. Paper presented at Global Learn Asia Pacific 2011, Melbourne Australia.
- Spiller, Dorothy. (2012). *Assessment Matters: Self-Assessment and Peer Assessment*. Waikato: The University of Waikato.
- Weigle, Sara Crushing. (2002). *Assessing Writing*. Cambridge: Cambridge University Press.
- William, Dylan. (2013). Assessment: The Bridge between Teaching and Learning. *Voices from the Middle*. 21, 2, 40-45.
- Zheng, H., Huang, J. & Chen, Y. (2012). Effects of Self-assessment on Chinese Students' Performance on College English Writing Tests. *Polyglossia*, Volume 23.

Appendix A

Assessment Students' Writing Problems

How do you rate the problem(s) of writing you have for each component?

No.	Component of writing	Rate				
		5	4	3	2	1
1.	Content					
2.	Organization					
3.	Language use					
4.	Vocabulary					
5.	Mechanics					

Meaning rating value:

- 1. Always
- 4. Most of the time
- 3. Rare/seldom
- 2. Almost never
- 1. Never

Appendix B

Scoring Guide

Com-ponents	Score	Weight	Mark/Grade	Criteria	Indicator
Content	5	5		Excellent	<ul style="list-style-type: none"> <li>▪ The essay is related to the topic chosen and is thoroughly developed.</li> <li>▪ The essay is related to the topic chosen but it is not thoroughly developed.</li> <li>▪ The essay is closely related to the topic chosen and is thoroughly developed.</li> <li>▪ The essay is closely related to the topic chosen but is poorly developed.</li> <li>▪ The essay is not related to the topic chosen and there is no development.</li> </ul>
	4			Good	
	3			Average	
	2			Fair	
	1			Poor	

<b>Organization</b>	5	5	Excellent	<ul style="list-style-type: none"> <li>▪ There are introductory, content, and concluding paragraphs; introductory paragraph contains general statement and thesis statement; content paragraph discusses what is stated in the thesis statement; and the shift from one paragraph to other paragraph moves smoothly</li> <li>▪ There are introductory, content, and concluding paragraphs; introductory has only thesis statement; content paragraph discusses what is stated in the thesis statement; the shift from one paragraph to other paragraph moves smoothly.</li> <li>▪ There are introductory, content and concluding paragraphs, introductory paragraph has general statement as background without thesis statement; discussion in content paragraph does not have focus; the shift from one paragraph to other paragraph moves smoothly.</li> <li>▪ There are introductory, content paragraph, and concluding paragraph; introductory neither has appropriate background and nor specific thesis statement; discussion in content paragraph does not focus; the shift from one paragraph to other paragraph does not move smoothly.</li> <li>▪ There is no paragraphing and organization is poor.</li> </ul>
	4		Good	
	3		Average	
	2		Fair	
	1		Poor	
<b>Language use</b>	5	7	Excellent	<ul style="list-style-type: none"> <li>▪ No deviation in using grammar (e.g., tenses, plural-singular form, concord and agreement, word order, etc.)</li> <li>▪ There is a slight deviation in using grammar, e.g. tenses, plural-singular form, concord and agreement, word order, etc.</li> <li>▪ There are many deviations in using grammar, e.g. tenses, plural-singular form, concord and agreement, word order, etc.</li> <li>▪ There are lots of deviation in using grammar, e.g. tenses, plural-singular form, concord and agreement, word order, etc.</li> <li>▪ There are too many deviations in using grammar, e.g. tenses, plural-singular form, concord and agreement, word order, etc.</li> </ul>
	4		Good	
	3		Average	
	2		Fair	
	1		Poor	



Vocabulary	5	2		Excellent	<ul style="list-style-type: none"> <li>▪ Using word appropriately, always using other word to refer the same thing</li> <li>▪ Using word quite appropriately, often other word to refer the same thing</li> <li>▪ Using word quite appropriately, other word is sometimes used to refer the same thing</li> <li>▪ Using other word inappropriately, other word is rarely used to refer to the same thing</li> <li>▪ Using word inappropriately, other word is never used to refer to the same thing</li> </ul>
	4			Good	
	3			Average	
	2			Fair	
	1			Poor	
Mechanics	5	1		Excellent	<ul style="list-style-type: none"> <li>▪ There is no error in spelling, punctuation or capitalization</li> <li>▪ There is one or two errors either in spelling, punctuation or capitalization</li> <li>▪ There are three to four errors either in spelling, punctuation or capitalization</li> <li>▪ There are five to six errors either in spelling, punctuation, or capitalization</li> <li>▪ There are more than seven errors either in spelling, punctuation, or capitalization</li> </ul>
	4			Good	
	3			Average	
	2			Fair	
	1			Poor	

Appendix C

EVALUATION SHEET  
FOR THE QUALITY OF STUDENTS' WRITING

Name : .....

Reg. Std. Num. : .....

No.	Component	Score		Obtained x Weight	Maximum Score
		Obtained	Weight		
1.	Content (1 – 5)		5		25
2.	Organization (1 – 5)		5		25
3.	Language Use (1 – 5)		7		35
4.	Vocabulary (1 – 5)		2		10
5.	Mechanics (1 – 5)		1		5
TOTAL SCORE				$\Sigma X =$	100

Banjarmasin, ....., 2014  
Scorer,

NIM.\* ..... (filled in with student's registration number)

Note:

$$\text{Final score} = \frac{\text{Obtained score}}{\text{Maximum score}} \times 100$$

Appendix D

Table 2. *How the students scored their writing tasks*

No.	Name (Code)	Score					Final Score	
		Cont.	Organ.	Lang. use	Voc.	Mech.	Score	Grade
1.	ASP	20	20	21	6	4	71	B
2.	APA	20	20	21	6	4	71	B
3.	AMU	20	20	28	6	4	78	B+
4.	CBN	20	20	28	8	5	81	A
5.	DRA	20	20	21	6	5	72	B
6.	EDM	20	20	28	8	4	80	A
7.	FRA	20	20	21	6	5	72	B
8.	FAU	20	20	21	6	4	71	B
9.	FNA	20	20	21	8	5	73	B
10.	IHA	20	15	21	6	4	66	C+
11.	IPK	20	20	21	8	5	74	B
12.	JAN	20	20	28	8	5	81	A
13.	MFA	20	20	28	8	4	80	A
14.	MLA	20	15	21	6	4	66	C+
15.	MAR	20	20	21	8	5	74	B
16.	MAD	20	15	21	6	4	66	C+
17.	MEL	20	20	28	6	4	78	B+
18.	MSA	20	20	28	8	4	80	A
19.	NIF	20	20	28	6	4	78	B+
20.	NOR	20	20	28	6	4	78	B+
21.	RDP	20	20	28	8	5	81	A
22.	RAH	20	15	21	6	4	66	C+
23.	RAM	20	20	21	6	4	71	B
24.	SHK	20	15	21	6	4	66	C+
Mean							73.92	

Appendix E

Table 3. *How the teacher scored students' writing tasks*

No.	Name (Code)	Score					Final Score	
		Cont.	Organ.	Lang. use	Voc.	Mech.	Score	Grade
1.	ASP	20	20	2	6	4	71	B
2.	APA	20	20	21	6	4	71	B
3.	AMU	20	20	28	6	4	78	B+
4.	CBN	20	20	28	8	5	81	A
5.	DRA	20	20	21	6	5	72	B
6.	EDM	20	20	28	8	4	80	A
7.	FRA	20	20	21	6	5	72	B
8.	FAU	20	20	28	8	5	81	A
9.	FNA	20	20	21	8	5	73	B
10.	IHA	20	15	21	6	4	66	C+
11.	IPK	20	20	21	8	5	74	B

12.	JAN	20	20	28	8	5	81	A
13.	MFA	20	20	28	8	4	80	A
14.	MLA	20	15	21	6	4	66	C+
15.	MAR	20	20	21	8	5	74	B
16.	MAD	20	15	21	6	4	66	C+
17.	MEL	20	20	28	6	4	78	B+
18.	MSA	20	15	21	6	4	66	C+
19.	NIF	20	20	28	6	4	78	B+
20.	NOR	20	20	28	6	4	78	B+
21.	RDP	20	20	28	8	5	81	A
22.	RAH	20	15	21	6	4	66	C+
23.	RAM	20	20	21	6	4	71	B
24.	SHK	20	15	21	6	4	66	C+
Mean							73.75	

## Appendix F

Table 4. Comparison between scores given by students and teacher

No.	Name	Students assessment		Teacher assessment		Differ	
		Score	Grade	Score	Grade	Yes	No
1.	ASP	71	B	71	B	-	v
2.	APA	71	B	71	B	-	v
3.	AMU	78	B+	78	B+	-	v
4.	CBN	81	A	81	A	-	v
5.	DRA	72	B	72	B	-	v
6.	EDM	80	A	80	A	-	v
7.	FRA	72	B	72	B	-	v
8.	FAU	71	B	81	A	v	-
9.	FNA	73	B	73	B	-	v
10.	IHA	66	C+	66	C+	-	v
11.	IPK	74	B	74	B	-	v
12.	JAN	81	A	81	A	-	v
13.	MFA	80	A	80	A	-	v
14.	MLA	66	C+	66	C+	-	v
15.	MAR	74	B	74	B	-	v
16.	MAD	66	C+	66	C+	-	v
17.	MEL	78	B+	78	B+	-	v
18.	MSA	80	A	66	C+	v	-

---

19.	NIF	78	B+	78	B+	-	v
20.	NOR	78	B+	78	B+	-	v
21.	RDP	81	A	81	A	-	v
22.	RAH	66	C+	66	C+	-	v
23.	RAM	71	B	71	B	-	v
24.	SHK	66	C+	66	C+	-	v
Mean		73.92	Mean	73.75			