Who Gains the Benefits of On-Task Versus Off-Task Self-Assessment? Higher or Lower Proficiency?

Herlinawati[1], M. Fadhly Farhy Abbas[2]
[1]Universitas Lancang Kuning
E-mail: herlinawati@unilak.ac.id
[2]Universitas Lancang Kuning
E-mail: fadhly@unilak.ac.id

Abstract

This study was attempted to investigate the effects of two different types of self-assessment: on-task versus off-task on EFL Indonesian students' writing performance across different proficiency level. The on-task self-assessment asked the students to self-evaluate their performance immediately after they completed their English writing tasks. The off-task self-assessment asked the students to evaluate their overall performance in a general and somewhat in de-contextualized way. The total number of forty-eight (48) students participated in this study and randomly assigned into two experimental groups. The data was gained from the post-test (essay writing test). The results indicated that the low proficiency students in on-task self-assessment group posed a significant effect and they could self-assess their writing performance more accurately than they had in an off-task format. In conclusion, on-task self-assessment is suggested to be implemented for lower achievers to consolidate their writing performance. However, further research is suggested to conduct to know the students' experiences in applying on-task and off-task self-assessment as an effective writing strategy.

Keywords: On-Task and Off-Task Self-Assessment, Writing Performance, Proficiency Levels

Tingkat Kemahiran Bahasa yang mana yang Diuntungkan dari Penilaian Diri Kontekstual dan Non-Kontekstual: Tinggi atau Rendah?

Abstrak

Penelitian ini bertujuan untuk mengetahui pengaruh dua metode penilaian diri: kontekstual dan non-kontekstual terhadap kemampuan menulis bahasa Inggris sebagai bahasa asing bagi mahasiswa Indonesia dengan latar belakang tingkat kemahiran bahasa Inggris yang berbeda. Penilaian diri kontekstual adalah suatu teknik penilaian dimana mahasiswa diminta menilai sendiri kemampuan menulisnya sesaat setelah mereka menyelesaikan tugas-tugas di kelas. Sedangkan penilaian diri non-kontekstual adalah bentuk penilaian diri dimana mahasiswa diminta untuk menilai kemampuan menulis secara umum dan adakalanya tidak ada hubungannya dengan penilaian pembelajaran menulis di kelas. 48 mahasiswa berpartisipasi dalam penelitian ini yang dibagi dalam 2 kelompok eksperimen dipilih secara acak. Nilai posttest tes menulis diambil sebagai data penelitian. Hasil analisis menunjukkan bahwa metode penilaian diri secara kontekstual berpengaruh secara signifikan bagi mahasiswa dengan tingkat kemahiran

bahasa yang rendah berbanding dengan penilaian diri non-kontekstual. Kesimpulan, metode penilaian diri kontekstual dapat dijadikan sebagai strategi yang efektif bagi mahasiswa untuk mengasah kemampuan menulisnya. Penelitian selanjutnya diharapkan untuk mempertimbangkan aspek pengalaman dosen dan mahasiswa dalam mengimplementasikan metode penilaian diri kontekstual.

Kata Kunci: Penilaian diri kontekstual dan non-kontekstual, kemampuan menulis, Tingkat kemahiran bahasa.

1. INTRODUCTION

It is generally recognized that assessing students' learning performance is one of the important tools used by the teachers to measure students' learning performance as well as to improve instructions, to provide feedback about learning progress and to motivate students. Assessment provides valuable information that allows teachers to adapt instructional procedures to the learning needs of the students. Harris & Brown (2013) state that assessments facilitate teachers in decision making about learning progress through systematic information gathering. In other words, the relationship between learning and assessment is very strong. Students learn more in classes where assessment is an integral part of instruction than in those where it is not. Therefore, teachers need to consider assessment as a part of teaching and instructions to reach the objectives of assessment.

In reality, one of the most basic and difficult task that teachers face in classroom teaching is the process of assessment because classroom assessment includes all the process involved in making decisions about students' learning progress. The classroom assessment process assumes that students need to receive feedback early and often, that they need to evaluate the quality of their own learning, and that they can help the teacher improve the strength of instruction. Prior to the students'

participation in the process of assessment, it can be a way providing an opportunity for students to assess their own learning while allowing the teacher/tutor to play their significant role in the process of assessment. This term is defined by co-assessment or collaborative assessment or cooperative assessment (Sluijsmans, Dochy, & Moerkerke., 1998).

Self assessment is defined as the students' procedures to evaluate their own learning by reflecting the quality of their own work, judge the degree to which it reflects explicitly stated goals or criteria and revise accordingly. Viewed from formative perspectives, self assessment is called student evaluation, by contrast, viewed from summative perspectives, it is often referred to students' grading or students marking which teachers using the scores from such assessments when they assign course grade (Esfandiari & Myford, 2013). Conceptually, self assessment is supported by theories of cognition, constructivism and learner autonomy of Piaget and Vygotsky theories which implies that knowledge and meaning are constructed. From these perspectives, meaningful learning is reflective, constructive and self regulated. Through self assessment, students learn to rely on themselves, take responsibility for their own learning, helping them to become more reflective and being engaged learners.

The benefits of using self assessment have been reported to

improve students' learning and to encourage students to become more actively involved in the learning process to have greater responsibility for setting goals and making decisions about their own learning (Ross, 2006; Matsuno, 2009; Lam, 2013; and Yoon & Lee, 2013). In L2 writing research, when self assessment is used with clear assessment criteria, it can promote personalized writing strategies and higher-order thinking skills (for example logic) that could navigate students towards their future learning of writing and integrating self assessment as an effective writing assessment and instruction could facilitate an appropriate assessment for students' better opportunities in writing progress (Schulz, 2009 & Lam, 2013, Matsuno, 2009, Kato, 2009, Rajab and Carol, 2013 & Suzuki, 2008).

In the process of developing L2 writing performance, different students have different progress cognitively and socially especially while doing the interactions. Some students tend to do self-corrections on their own works more while others tend to interact with one another or with their peer (Storch, 2005; Hanjani & Li, 2014). Some students consider teachers as the only legitimate agent for giving supports for their learning experiences and therefore are reluctant to assess themselves or their peers or negotiate with others over its use in their revisions and much prefer to have teachers assess them. It is explained mainly in terms of students perceived low/high level of EFL proficiency. Therefore, the factor of students' characteristics must be taken into account. As stated by Paleczek, Seifert, Schwab & Klicpera (2015), there is a difference among the performance of students possessing different types of proficiency; high and low. It is believed that higher proficiency students will have better initial scores. However, it is still unclear if initial proficiency upon a learning process affects growth rate.

From previous studies, little is known about the effects of different tasks in SA especially for L2 writing with different levels of proficiency. This is the focus of the study. Under the assumption that if differential effects of different task in SA on specific aspects of L2 writing are observed for various L2 learners, only a portion of the students will benefit from the implementation of SA. To come to an accurate answer, the present study is expected to be able to give scientific conclusions as new evidence of using different types of SA for different proficiency levels in developing students' L2 writing performance.

2. METHOD

The design of this study was a quasi experiment which aimed at investigating the effects of two types of self-assessment (on-task versus off-task) for different proficiency levels on students' argumentative essay writing. The total numbers of students involved were 48 students. Pre-test (writing test) was administered to know whether the students of the two experimental groups were equally homogenous in terms of writing performance. Before giving the treatment, the 48 students were classified into high and low proficiency level based on the test of TOEFL-Like score. The distribution of subjects in experimental groups is shown in the table below.

The experimental conditions

Methods	Proficiency Levels
On-Task SA	H(12) L(14)

Off-Task SA	H (9)	L (13)
	H (21)	L (27)
Total (48)		

Post-test was held at the end of experimentation. The students' writing was scored by three raters by using the rubric developed by the researcher. The rubric covered five different rating dimensions of writing quality with 100-point scale, each dimension having a different weight: introduction (15 points), argumentative points (40 points), organization (15 points), sentence structure and convention (20 points), and relevance (10 points).

To estimate the reliability of the scores, the measurement of inter-rater reliability was used in this study. Pearson Product Moment Correlation was used to estimate the internal consistency between the three raters. The correlation was performed in coefficient alpha.

3. FINDINGS AND DISCUSSION

Before testing the hypotheses, the statistical assumptions (independence of the dependent variable, normal distribution, and the homogeneity of variances) had been fulfilled. To maintain the independent test of all participants of this study, the test was administered once. In the same time, all participants were asked to write an argumentative essay independently based on the instructions in the writing prompt. Shapiro-Wilk test was used since the number of participants of this study was categorized into small/medium. The essay writing test scores gained by the low proficiency students who applied on-task self assessment was D (14) = .978, p = .964, and those low proficiency students who applied off-task self assessment was D (14)=.905, p= .134 were normally distributed. For those high proficiency students who applied on-task self assessment was D (12) = .886, p = .106, and those high proficiency students who applied off-task self assessment was D (8) = .952, p = .735 were both statistically normal. The output of test implied that the assumptions of homogeneity met the evidence that Levene statistic was (3, 44) = .628, p = .601. The result of homogeneity test was, p=.601 > alpha (.05). From the results of testing normality and homogeneity variances above, all of the assumptions were fulfilled. From the results of the test, data analysis using one-way ANOVA was preceded. From the result of analysis, there was significant effect of on-task SA and off-task SA on students writing. Statistically, the findings revealed that the quality of students' writing applying on-task and off-task SA was different. Based on the statistical output for the effects of on-task vs. off-task self assessment, the F value was higher than F table (3.639 > 2.816) and p < alpha (.020 < .05). Therefore, H0 was rejected which means that there was a significant different among the four means of the four experimental groups.

The follow up Tukey's test (5%), the result was shown as follows: the group A (low) was statistically significant different with group B (low) by the evidence that p (.016) < .05. With regards to the research question of the study, whether the effects of different types of self-assessment depend on students' proficiency level, the study revealed that low proficiency students' writing performance in on-task self-assessment group performed better than in off-task self-assessment. From the observation during the experimentation, the students in on-task

self-assessment group were provided with rich contexts for evaluating and rating themselves. This result was similar to the finding of Chalhoub-Deville (2001) that on-task assessment was used by the students to provide them with opportunity to negotiate by using their language in the classroom contexts. Therefore, this finding suggested that the self-assessment activity could be a reliable and valid measure only when it was applied in contextualized situations. This result was also similar to the previous findings of Butler and Lee (2006), Oscarson (2009) and Farahani (2009) indicated that the participants in on-task self assessment showed a great difference in motivation, sociability, and confidence as a direct effect of contextualized experiences.

In relating to the research question, the result was supported by Larsen-Freeman (2006). She found that different rates of change among her homogeneous learners of L2 ability. The learners of lower proficiency in her study had faster rates of growth because the students had lower starting point upon and enrollment but they had similar outcomes with higher proficiency students. It can be concluded that the students' experiences in on-task self assessment could be expected as an important factor to be considered to consolidating students' language skills. However, this study needs to be verified about why the lower proficiency students could have better performance compared with higher proficiency students. The reason was supported by Blanche and Merino (1989) that more proficient students tend to underrate themselves, while less proficient students tend to overrate themselves. Therefore, the next analysis about the finding needs to be verified.

4. CONCLUSION

When the statistical issues caused the different results between the two experimental groups, the other possible causes could be observed from different perspectives and aspects such as students and teacher's experiences on implementing on-task and off-task self-assessment as well as higher/lower students' expectations in implementing on-task/off-task self-assessment especially in EFL writing.

REFERENCES

Brown, H. D. 2004. *Language Assessment: Principles and Classroom Practices*. New York: Pearson Education Inc.

Butler, Y.G., & Lee, J. 2006. On-task versus off-task self-assessments among Korean elementary school students studying English. The modern language journal. Vol. 90 (4).

Butler, Y. G., & Lee, J. 2010. The effects of self-assessment among young learners of English. Language Testing. Vol. 27 (1) : 5-31.

Esfandiari, R., &Myford, C.M. 2013. Severity Differences among Self-assessors, Peer-assessors, and Teacher-assessors rating EFL essays. *Assessing Writing*. Vol. 18:111-131.

Farahani, K. & Khezrlou, S. 2009. On-Task versus Off-Task Self-Assessments: An Experience in Iranian Context. Journal of Teaching English as a Foreign Language and Literature. Vol. 3 (39-52).

Hanjani, A.M., & Li, L. 2014. Exploring L2 Writers'

Collaborative Revision Interactions And Their Writing Performance. *System.* Vol. 44 (101-114).

Harris, L.R., & Brown, G.T.L. 2013. Opportunities and obstacles to consider when using peer- and self-assessment to improve student learning: Case studies into teachers' implementation. *Teaching and Teacher Education.* Vol. 36 (101-111).

Kato, F. 2009. Student Preferences: Goal-Setting and Self-Assessment Activities in Tertiary Education Environment. *Journal of Language Teaching Research.*Vol.13 (2).

Lam, R. 2013. The Relationship Between Assessment Types and Text Revision. *ELT Journal* Vol. 67 (4).

Larsen, F. D. 2006. The Emergent of Complexity, Fluency, and Accuracy in the Oral and Written Production of five Chinese Learners of English.*Applied Linguistics*. 590-619.

Lee, I. 2007. Feedback in Hong Kong Secondary Writing Classrooms: Assessment for Learning or Assessment of Learning?.*Assessing Writing.* Vol. 12: 180–198.

Matsuno, S. 2009. Self-, Peer-, and Teacher-Assessments in Japanese University EFL Writing Classrooms.*LanguageTesting*.Vol.26 (1).

Oscarson, A.D. 2009. Self-Assessment of Writing in Learning English as a Foreign Language. Gateborgs Universitet.

Paleczek, L., Seifert, S., Schwab, S., & Klicpera, B. G. 2015. Assessing reading and spelling abilities from three different angles – correlations between test scores, teachers' assessment and children's self-assessments in L1 and L2 children.*Procedia Social and Behavioral Science*. Vol. 174 (2200-2210)

Ross, J., A. 2006. The Reliability, Validity, and Utility of Self-Assessment.*The Journal of Practical Assessment Research & Evaluation*.Vol.11 (10).

Schulz, M. M,. 2009. Effective Writing Assessment and Instruction for Young English Language Learners. *Early Childhood Education Journal*. Vol. 37: 57–62.

Sluijsmans, D., Dochy, F., &Moerkerke, G. 1998.The Use of Self-, Peer-, and Co-Assessment in Higher Education. Educational Technology Expertise Centre, OTEC Report. Open University of the Netherlands.

Storch, N. 2005. Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing*. Vol. 14 (153-173).

Suzuki, M. 2008. Japanese Learners' Self Revisions and Peer Revisions of Their Written Compositions in English. *TESOL Quarterly*.Vol.42 (2).