

Efficacy of Student Assessment as Part of English Writing Instruction for Japanese High School Students

Yoko Oi

Abstract

The study focused on the efficacy of student assessment for two reasons. Firstly, student assessment, such as self- and peer assessment, has long been used in English classes, from elementary level to adult learners, but senior high school students have not been targeted in previous studies. Secondly, writing classes seem to be stagnant and students' interest has diminished in recent years. Therefore, the study aimed to activate a writing class to develop students' writing skills to keep up with the globalized society. As a solution, the present study proposed that student assessment could trigger the activation of a writing class in Japan, because it is considered that student assessment enables students and teachers to share responsibility for learning, teaching, and assessment. However, the specific features of self- and peer assessment have not been thoroughly elucidated, especially for young learners. In particular, the effects of each student assessment method on writing performance and learner affect have not been well investigated. The study therefore implemented a convergent mixed-methods approach by adopting a comparative study of self- and peer assessment to clarify the specific qualities of each assessment type. Furthermore, the present study aimed to explore the relationship between formative assessment and student assessment to make student assessment effective for both teachers and students. The results of the study showed that both

assessment types had positive effects on grammatical accuracy and learner autonomy, while the severity of self-assessment ratings was close to that of teacher assessment, and only the self-assessment group improved the composite score in the writing test. In terms of the five key strategies of formative assessment (Leahy & Wiliam, 2012), the peer assessment group covered all five key strategies, while self-assessment covered three. Based on those findings, it was difficult to judge the hierarchy of the two assessment types, so it is important for teachers and administrators to know the features of both types and complement the strengths and weaknesses of each. The critical role of teachers in student assessment was found to make formative assessment successful.

Keywords: self-assessment, peer assessment, teacher assessment, formative assessment, writing performance, learner affect, high school students, convergent mixed-methods

Chapter 1 INTRODUCTION

The purpose of this study is to reveal the potential of student assessment to enhance formative assessment in writing classes in Japanese high schools. In the present study, student assessment, that is, self-assessment and peer assessment, is defined as the involvement of students as assessors in assessment. Student assessment has been implemented in class and considered to be effective in encouraging learners to be responsible for their learning and motivated to learn. However, the features of these two assessment methods are yet to be investigated thoroughly or understood well, especially for Japanese senior high school students. It is believed that student assessment has a positive effect on the development of writing ability and learner affect, so the current study aims to explore the efficacy of student assessment with special attention to these two issues.

As the background to this study, two points should be considered. Firstly, it is necessary to find a solution to break the deadlock of the status quo of English writing classes in secondary schools in Japan. Secondly, it is essential to make writing instruction effective and to improve the English writing proficiency of EFL learners in Japan to keep up with the rapid globalization. The present study proposes that writing classes need to be reformed and that effective adoption of student assessment could be key to achieving this goal. Another issue the present study highlights is the relationship between formative assessment and student assessment. EFL writing classes are substantially organized by both teachers and students. Illuminating and enhancing the role of students as the agents of classroom assessment might facilitate the development of writing instruction and writing ability. With the above as the background, the following research questions were

addressed in order to explore the efficacy of student assessment:

- (1) How do self- and peer assessment compare with each other in terms of:
 - a) the reliability of scores against teacher assessment;
 - b) the effects on writing ability; and
 - c) the effects on writing anxiety and learner autonomy?
- (2) How can student assessment work as formative assessment in the classroom?

Chapter 2 LITERATURE REVIEW

This chapter mainly discusses previous studies on the framework of assessment in the classroom in terms of classroom-based assessment and formative assessment. With respect to student assessment, the details are explored by taking a meta-analysis approach. Here, the research themes and methodologies of previous studies are summarized, as well as previous findings concerning various aspects of self- and peer assessment, including score reliability compared to teacher assessment, as well as their effectiveness and limitations. In addition, previous studies about learner affect, in other words, writing anxiety and learner autonomy, are discussed. The scales developed in previous studies to measure writing anxiety and learner autonomy are also compared, focusing on usefulness and validity. At the end of this chapter, study hypotheses are postulated based on the findings of the previous studies.

Chapter 3 METHODOLOGY

Chapter 3 outlines the methodology adopted in the present study. In this study, 293 Japanese senior high school students aged 15 to 18, two Japanese English teachers, and two native-speaking English teachers participated in the survey. The study adopted a convergent mixed-methods research design because this makes it possible to integrate quantitative and qualitative data, compensating for the weakness of each research method. Before the study, two types of writing assessment rubric were developed, one for self-assessment and the other for peer assessment. The Second Language Writing Anxiety Inventory (SLWAI; Cheng, 2004) and Chang's questionnaire (2007) were adapted as the measures of writing anxiety and learner autonomy. These questionnaires were translated into Japanese and modified based on the results of a pilot study conducted in regular English classes at a high school in Japan prior to this study. The study comprised three phases: (1) student assessment training sessions, pre-test, and pre-questionnaire; (2) intensive writing and student assessment sessions, with observation by teachers; and (3) post-test, post-questionnaire including open-ended questions, and semi-structured interviews. In the second phase, four teachers observed students' attitudes toward writing and student assessment.

In the quantitative analyses, the reliability of student assessment compared to teacher assessment was analysed in terms of the results of the many-faceted Rasch measurement (MFRM; Linacre, 1989; Linacre & Wright, 1993; McNamara, 1996). The severity and consistency of rating scores were analysed by employing the MFRM. Concerning Hypothesis 1.2, MFRM was again employed to examine the improvement in writing ability by focusing on the composite scores and analytic rating scores between

pre- and post-tests assigned by the teachers. The effects of student assessment on learner affect was investigated through a multivariate analysis of covariance (MANCOVA). MANCOVA was conducted to analyse the covariates of difference between writing anxiety and learner autonomy levels in the self- and peer assessment groups for the pre-test and post-test. The covariates were pre-somatic anxiety, pre-cognitive anxiety, pre-avoidance, and pre-learner autonomy.

On the other hand, two types of qualitative data were analysed: (1) self- and peer assessment group students' responses to open-ended questions conducted in regular English classes just before and after the experiment; and (2) protocols of one-to-one semi-structured interviews with two teachers and 12 students. In order to analyse the qualitative data, opinion words were extracted from the students' responses to open-ended questions, and grounded theory (Glaser & Strauss, 1967) was employed to analyse the transcription of semi-structured interviews with 12 students who were selected from the self- or peer assessment group and two teachers who observed the intensive student assessment sessions. The results of both the quantitative and the qualitative analysis were synthesized in order to explore the relationship between student assessment and formative assessment

Chapter 4 QUANTITATIVE STUDY

Chapter 4 presents the results of the quantitative analyses of the difference between self- and peer assessment from three perspectives: a) the reliability of scores compared to teacher assessment; b) the effects on writing ability; and c) the effects on writing anxiety and learner autonomy.

Firstly, the reliability of scores compared to teacher assessment is analysed in terms

of the severity and consistency of ratings in the MFRM analysis of the composite pretest and post-test scores. It was found that the self-assessment group students were stricter than those in the peer assessment group, while the severity of the self-assessment group was close to that of teacher assessors. However, neither the self- nor the peer assessment group was as consistent in assigning scores to their own or peers' writing products as the teacher assessors. In short, neither self- nor peer assessment showed the same level of reliability as teacher assessment, but self-assessment did show a level of rating severity comparable to teacher assessment.

Secondly, the results of MFRM analyses of the data obtained from pre- and post-test composite scores and four pre- and post-analytic scores of self- and peer assessment groups suggested that both student assessment methods had some effect on the improvement of writing ability. In particular, both assessment types improved grammatical accuracy, while only the self-assessment group improved the composite score. In short, both assessment methods had an influence on the improvement of grammatical accuracy, and self-assessment in particular is considered to have an effect on overall writing improvement.

Finally, the relationships with writing anxiety and learner autonomy were analysed through a MANCOVA of learner questionnaire responses. It was found that both types of student assessment resulted in learner autonomy improvement from the pretest to the post-test. On the other hand, there was a change in balance between subscales of writing anxiety, but this did not reach a statistically significant improvement from the pre-test to the post-test.

To sum up, the results above suggest that the two types of student assessment have differential effects on the reliability of assessment, the improvement of writing ability,

and learner affect.

Chapter 5 QUALITATIVE STUDY

This chapter reports the results of the qualitative analyses that investigated the effects of self- and peer assessment on writing ability and learner affect. The qualitative data from the responses to an open-ended questionnaire and semi-structured interviews were explored in terms of term extraction and grounded theory analyses.

According to the frequency of words that appeared in the open-ended responses (Table 5.1, p.188), about half of the students in the self-assessment group mentioned self-reflection, self-awareness of the limited vocabulary and grammar, and lack of confidence in conducting the assessment by themselves. On the other hand, about half of the students in the peer assessment group referred to positive learning effects from peers and enjoyment of the peer assessment activities. The common response obtained from both assessment groups was a lack of confidence in their own English ability and student assessment, although the ranking of this item differed from each other between the two assessment types. Lack of confidence was ranked higher in the self-assessment group than in the peer assessment group.

The results of the analysis of data from semi-structured interviews with students and teachers in respect of grounded theory showed that the two assessment types shared the same core category: development as a writer, although the makeup of the other categories differed between the self- and peer assessment groups (Figure 5.3, p.229). To be specific, except for the category of learners' perception of English study (Category C for the self-assessment group, Category G for the peer assessment group), different

categories emerged across the groups. For instance, the student interviewees in the self-assessment group more frequently commented on the effects of self-assessment on the development of consciousness of assessment criteria and writing content, which positively influenced the enhancement of self-reflection and a habit of self-revision, while the student interviewees in the peer assessment group more frequently stated the enjoyment of peer assessment and interaction with peers, which led to a positive learning effect from peers and metacognition. Those differences are displayed in a contrastive schematic representation of learner development as a writer between the self- and peer assessment groups (Figure 5.3, p.229).

Chapter 6 DISCUSSION

This chapter synthesizes the results of the quantitative and qualitative analyses obtained from the mixed-methods study in order to discuss the two research questions. In particular, the relationship between formative assessment and student assessment is addressed based on the integrated results of quantitative and qualitative analyses.

With respect to the first research question, the results of the quantitative and qualitative analyses were integrated and the two types of assessment were compared in terms of reliability, writing ability, and learner affect. The synthesized data presented commonalities and differences between the two assessment types. With respect to commonalities between the two assessment types, the results of the quantitative analyses based on the MFRM suggested that the reliability of scores assigned in self- and peer assessment was not as high as that of teacher assessment. At the same time, the self-assessment group presented a level of severity that was close to teacher assessment, while

the peer assessment group was more lenient than the self-assessment group and teacher assessors. Moreover, scores assigned by students in both assessment groups were inconsistent. The results above were supported by the results of qualitative analysis of student and teacher semi-structured interview results. The student interviewees in both groups and teacher interviewees suggested doubt as to the reliability of student assessment, the need for teacher guidance to confirm the accuracy of student assessment, and the importance of a reliable standard for their study.

With respect to the improvement of writing ability in terms of the quantitative analysis, both assessment types could improve grammatical accuracy, while only the self-assessment group improved the composite score. The results of qualitative analyses of student responses to the open-ended questionnaire and student and teacher semi-structured interviews also suggested that both types of student assessment had a positive effect on writing performance. This is because self-assessment enabled students to enhance self-reflection, while peer assessment nurtured metacognition. It was considered that those things helped students to improve their writing performance. As for the effects of student assessment on learner affect in terms of the results of quantitative MANCOVA analysis, learner autonomy was positively influenced, while writing anxiety was not changed after the intensive student assessment sessions.

With regard to the second research question, how student assessment works as formative assessment in the classroom was discussed. The results of qualitative analysis of the responses to the open-ended questionnaire and semi-structured interviews considered that the commonalities and differences between self- and peer assessment were related to distinct qualities of both types of assessment. The students in the self-assessment group enhanced self-reflective attitudes toward English writing and were

involved in deepening a sense of learner autonomy, while the students in the peer assessment group had positive effects on the enhancement of metacognition and social interaction with peers. It was considered that those positive effects were related to the improvement of grammatical accuracy and learner autonomy.

The five key strategies proposed by Leahy and Wiliam (2012) were applied to analyse the relationship between formative assessment and student assessment based on the qualitative data: (1) clarifying, sharing, and understanding learning intention and criteria for success; (2) drawing evidence from learning; (3) providing feedback that promotes learning; (4) encouraging learners to take on the role of instructional resource for one another; and (5) promoting learners to become owners of their own learning (p. 7). The peer assessment group moderately covered all five key strategies, while the self-assessment group conformed to only three of these strategies, because self-assessment could not provide external feedback or exchange instructional resources with others. Though all the key strategies of formative assessment (Leahy and Wiliam, 2012) did not apply to self-assessment, it was difficult to judge which assessment type was superior.

In addition, the synthesized results of the quantitative and qualitative analyses suggested the importance of teachers in relation to student assessment and formative assessment. As student and teacher interviewees mentioned, the absence of teachers in the present study illuminated the critical role of teachers in formative assessment, and teacher involvement in student assessment was desirable to make formative assessment more fruitful. Therefore, it is important for teachers and school administrators to adopt student assessment in class, based on the qualities of each assessment type, complementing each other's strengths and weaknesses.

Chapter 7 CONCLUSION

The present study aimed to explore the efficacy of student assessments for Japanese high school students. The comparative study identified important commonalities and differences between self- and peer assessment. Also, a mixed-methods analysis allowed the quantitative and qualitative data to be brought together, making it possible to analyse the recondite relationship between student assessment and formative assessment. The results of quantitative analysis suggested that student assessment was not a substitute for teacher assessment. The results of qualitative analysis showed that self- and peer assessment had different qualities. The study illuminated the importance of student assessment as well as the critical role teachers play in the classroom context in order to revitalize writing class. The key stakeholders in classroom assessment are students and teachers. Therefore, it is concluded that teachers and students should share responsibilities for teaching, learning, and assessment so that formative assessment can be successfully demonstrated in class.