

Perceptions and Use of AWE Facilitated Peer Feedback for Improving Writing

Dragana LAZIC, Saori TSUJI

Introduction

Feedback is a method often used in L2 classrooms to help students develop their writing skills. It is “information provided by an agent (e.g., teacher, peer, book, parent, self, experience) regarding aspects of one’s performance or understanding” (Hattie & Timperley, 2007). The purpose of feedback is to help learners reduce the gaps between current performances and desired learner outcomes. Even though literature on the perceived usefulness and effectiveness of feedback in EFL writing has clearly demonstrated that EFL students still prefer teacher feedback to other types (Baier Schmidt, 2012; Chang, 2016; Chen, 2014; Hyland & Hyland, 2006; Otoshi, 2012), peer feedback and automated feedback provided by computers, are also beneficial for students’ learning and motivation, albeit serving different purposes. Peer feedback helps learners as they write for more authentic audiences, which in turn raises awareness of the communicative purpose of writing. According to Hyland and Hyland (2006), it also helps writers, especially beginners, to see how others understand their writing while, at the same time, they can negotiate meanings, practice language use against peer comments, and revise.

An area that has gained a considerable research interest in recent years is technology reinforced human feedback in EFL writing classes. One research avenue looks at the use of automated writing evaluation (AWE) with electronically

delivered peer feedback. Vojak, Kline, Cope, McCarthy & Kalantzis (2011) found some common issues inherent to these programs: feedback is formulaic and non-specific with some incorrect error flagging; there is overemphasis on surface-level text errors; the length of the text is more important than the quality. In other words, AWE fails to recognize writing as “socially-situated practice,” which is “functionally and formally diverse activity; and that it is increasingly multimodal” (Vojak et al., pg. 108).

Many agree (e.g., see Chen & Cheng, 2008; Huang, 2014) that in order to address the shortcomings of computerized feedback, it is necessary to provide human feedback in computer-assisted learning environments as these often fail to recognize global errors in writing, namely, discourse context or rhetorical aspects of writing. In addition, if writing is conceptualized as a communicative, meaning-making process, to be meaningful, it needs human attention, or as Godwin-Jones (2018) noted, it involves negotiation with readers and awareness of appropriate writing approaches and strategies.

The purpose of the current study is to explore how an AWE system, ETS Criterion®, and peer feedback could be used to support writing in an EFL classroom at a public university in Japan. Previous studies in Japanese educational context focused on comparing teacher and AWE feedback (Heffernan & Otsoshi, 2015; Long, 2013; Otsoshi, 2012), using an AWE classroom setting and its perceived usefulness (Ohta 2008; Tsuda 2014; Wakabayashi, 2013), and examining the validity argument of using an AWE system in assessing writing (Koizumi et al., 2016). However, to the authors knowledge, little research has been done on electronically delivered peer feedback (e.g., one study, Heffernan and Otsoshi, 2015). Besides, while studies among Japanese learners looked at writing improvements over time and from an assignment to assignment (Heffernan and Otsoshi, 2015; Otsoshi, 2012), less attention was paid to the potential use of an AWE system in relation to peer feedback for revision. In general, revision in the context of EFL academic writing is an under-researched topic, and this is addressed here through the study design. Thus, the following research questions guided the current study:

1. What are students' perceptions about the usefulness and effectiveness of

feedback provided by peers in combination with an AWE system for learning English writing?

2. How can students use peer feedback effectively in combination with an AWE system?

Literature Review

Peer Feedback

Peer feedback is applied to facilitate the writing process as an instructional strategy. Authors use different terms to describe the same or similar processes and related activities, for example, ‘peer response’ (Hyland and Hyland, 2006), ‘peer revision,’ ‘peer tutoring’ and ‘peer critiquing’ (examples in Hu, 2005, pg. 321). The terminology might differ; however, these definitions have several points in common. First, peer feedback is a collaborative activity involving pairs or groups of learners who read and interact with each other in order to exchange spoken, written, or mixed comments in the process of writing. The interaction goes in both directions (Hyland & Hyland, 2006), and both sides are “helping to extend each other’s writing competence” (pg. 91). Second, learners take the role and responsibilities of a trained teacher or tutor (Yu & Lee, 2016), focusing on both global and local writing issues (Chang, 2016). The goal is to help other learners to make immediate changes and develop stronger writing competences over time (Hu, 2005). Finally, this is a process without grades or formal evaluations (e.g., see Chang, 2016). In this paper, the term ‘peer feedback’ denotes both the process and the product of this scaffolding activity (for a further distinction and terminology, see Yu & Lee, 2016).

Research on peer feedback in EFL has been steadily growing in the past 30 years, as shown in systematic reviews (Chang, 2016; Hyland & Hyland, 2006; Yu & Li, 2016). These reviews suggest that peer feedback is effective as it provides students with opportunities to raise audience awareness (Hyland and Hyland, 2006; Chang, 2016), and acquire a new language, good sentences, and organization by reading peers’ assignments (Hyland & Hyland, 2006). Findings consistently demonstrated that training played an important role in making the peer activity

useful as trained peer reviewers provided higher-quality feedback on both global and local issues (reviewed in Hyland & Hyland, 2006; Yu & Lee, 2016). Finally, feedback provided in L2 is more likely to focus on the form, while L1 comments referred to both form and content of writing (Chang, 2016).

When compared to other types of feedback, according to Chang (2016), peer feedback was perceived as a more useful, specific, and higher quality than feedback provided by an AWE system. However, despite the perceived effectiveness of peer review, the reviews mentioned above affirmed that teacher's feedback was preferred to and was more prominent in students' revisions than peer feedback. Peers were perceived as lacking language and communicative skills in L2 to detect errors or focusing too much on local or surface-level errors, e.g., grammar (Chang, 2016; Hyland & Hyland, 2006; Yu & Li, 2016).

Peer Feedback in Japanese Education

Research on peer feedback in a Japanese context looked at the advantages and disadvantages of peer feedback as an instructional tool in writing classes. Learners had positive attitudes towards peer feedback as it led them to become more aware of their audiences; they took more time to plan for writing; and improved grammatical and organizational skills (Kohoro, 1995; Yoshikawa, 2016; Baierschmidt, 2012; Wakabayashi, 2013; Fujii et al., 2016). In addition, students noticed that negotiating and feedback could contribute to language learning (Fujii et al., 2016; Yakame, 2005) as they needed to realize by themselves how to improve composition (Yakame, 2005). Several studies showed that learners perceived reading their peers' writing as being particularly useful as it enabled learners to learn composition, compare their writing to that of their peers (Yoshikawa, 2016), and use peer's texts as sources of information (Wakabayashi, 2013). In addition, reading was a mediation tool for increasing learning among lower proficiency writers (Allen and Mills, 2016).

Studies in Japan also looked at the relationship between proficiency levels, types, and the amount of feedback provided and received. Kohro (1995) found that one of the disadvantages of peer feedback is that highly proficient and motivated students benefit more from peer feedback, while lower proficiency

students can hardly improve. Others found that higher proficiency learners gave more suggestions, whereas lower-level learners used less meaning-related suggestions (Allen and Mills, 2016). Additionally, students with less confidence gave and used less feedback (Allen and Katayama, 2016).

Even though feedback motivated students to include what they learned in their subsequent writing and to be more aware of not repeating similar mistakes when writing later (Yoshikawa, 2016), not all agree that peer feedback has more benefits than other types of feedback. For example, Suzuki (2008) found that students negotiated more during peer revisions, but more revisions were made during self-revision. Baierschmidt (2012) reported that students preferred teacher feedback to that provided by their peers and suggested that lack of confidence was one of the reasons why students did not include most of the peer suggestions in their reviews. Studies also demonstrated that students questioned whether their peers had linguistic and other competencies to provide meaningful and accurate feedback (e.g., in Yoshikawa, 2016).

Despite mixed findings in terms of the usefulness of the peer process, it may be said that peer feedback has its place in the Japanese EFL writing classroom. Basically, more training, especially metacognitive instruction training (Fujii et al., 2016) before a peer feedback activity, would benefit students as a lack of training might have affected their actions and attitudes (Bierschmidt, 2012).

ETS Criterion® in Japanese Educational Context

Automated writing evaluation or AWE, in literature also known as automated essay evaluation or AEE (e.g., in Elliot & Klobucar, 2013), automated essay scoring or AES (Burstein, Tetreault, & Madnani, 2013; Vojak et al., 2011), is the ability of a program to score and evaluate writing. In other words, it is “an attempt to model human essay scoring, with its assignment of scores or grades based upon a rubric” (Deane, 2013). However, it is important to note that if AWE programs are to be effectively used as instructional tools, they need to be used in a way to provide formative and not only summative assessment (e.g., in Yu, 2015), meaning providing the feedback is more important than showing the score.

In terms of automated writing evaluation in the Japanese context, ETS

Criterion® is one of the most often used AWE tools, and the current study also used this software. ETS Criterion® is a web-based AWE service that uses natural language processing to “identify construct-relevant linguistic properties in text” (Burstein et al., 2013). Previous research looked at validity argument and how this program can be used for assessing L2 writing proficiency (Koizumi et al., 2016), compared the effects of teacher feedback and ETS Criterion® (Heffernan and Otoshi, 2015; Long, 2013; Otoshi, 2012), surveyed classroom use and perceived usefulness (Ohta, 2008; Tsuda, 2014), explored ambiguities in the Criterion’s topic categorization method and how this affected its use (Taoka, 2015), and examined learner gains in writing skills (Ohta, 2008; Otoshi, 2012). In Tsuda’s research (2014), Japanese EFL learners used this AWE tool to write at home. Their perceptions were positive as they appreciated having many opportunities to write essays; consequently, this led to improved writing, more writing as they got used to writing and more planning before writing. Ohta (2008) also found that students have positive perceptions of using ETS Criterion due to the program’s immediacy and availability.

However, it is also reported that the progress made by students by using ETS Criterion® had some limitations. For example, improvements depended on student’s fluency and the time students spend on writing tasks (Ohta, 2008). Koizumi et al. (2016) found that students who used ETS Criterion® over 28 weeks showed improvements but only in some aspects of writing, e.g., length and syntactic complexity. Additionally, students who used ETS Criterion® also expected teachers to provide feedback on content and organization, as well as vocabulary (Otoshi, 2012). Heffernan and Otoshi (2015) demonstrated that rhetorical features of writing, e.g., thesis statement, showed no improvements when ETS Criterion® was used without the teacher’s feedback. Long (2013) found that the majority of feedback instances provided by ETS Criterion® focused on surface structures, whereas teachers focused on meaning. The same study showed minimal improvements in terms of usage and mechanics, while there was an increase in the number of errors for grammar and style.

Methodology

Study Context and Participants

The study took place at a public women's university in Japan. In total, 31 students participated by signing a consent form previously approved by the university's ethics committee and/or by giving their oral consent before the interviews. All students were either 18 or 19 years old at the time of the study, all were Japanese native speakers and were taking one of the compulsory writing classes: *Academic Writing 1* for the first-year students and *Academic Writing 3* for the second-year students. *AW1* course taught the basic structure of a sentence and a paragraph, while *AW3* taught writing multiple coherent paragraphs and essays. Both courses were taught by the first author and lasted 16 weeks (divided into two quarters). At the beginning of the school year, all students were introduced to a process-oriented approach to writing (for an explanation of this approach to teaching writing, check Yu & Lee, 2016). These groups of students could be considered as having lower proficiency (TOEFL ITP between 400 and 450).

Table 1 presents demographic data for the participants in the study. Most of the participants studied English for up to ten years.

Table 1 *Participants' Demographics*

Measure	Item	Frequency	Percentage
English language experience	5-10 years	28	90%
	>10 years	3	10%
Year at the university	1 st	21	68%
	2 nd	10	32%
Did you use any online tool to help you with writing in English?	Yes	14	45%
	No	17	55%
Did you use ETS Criterion® before?	Yes	10	32%
	No	21	68%
What is peer feedback?	No answer	13	42%
	I don't know	10	32%
	Answered	8	26%

Note. n =31.

When asked to define peer feedback, most participants did not provide an answer or did not know what it was. Among those who answered (8), only three students provided acceptable answers: “check in pairs with classmate,” “to check what I wrote with a classmate,” “to write better essays by exchanging ideas with classmates.” Regardless of their answers being simple, it can be considered that these students understood the concept of peer feedback, up to some degree, prior to the commencement of the research.

As shown in Table 1, the majority of students (55%) did not use ETS Criterion® before participating in the study. However, when asked which aspect of their writing ETS Criterion® can help them with, students had some ideas about what they needed to be helped with: grammar (16 answers), organization/logic (13), expressivity (9), writing speed (7), and content (6) (n=51, multiple choice answers). To capture students' initial interest in academic writing and writing in English in general, we asked several questions before the study commenced (Table 2).

Table 2 *Participant's Perceptions about Writing in English in General and Academic Writing*

Item*	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I read native English writing***	5 (17%)**	11 (37%)	11 (37%)	2 (7%)	1 (3%)
Academic writing skills are important in my current studies			2 (6%)	19 (61%)	10 (32%)
Academic writing skills are important for my future career			4 (13%)	10 (32%)	17 (55%)
I am interested in academic writing	2 (6%)		10 (32%)	11 (35%)	8 (26%)
I like to write in English	4 (13%)	7 (23%)	8 (26%)	8 (26%)	4 (13%)
I'm good at writing in English	7 (23%)	16 (52%)	5 (16%)	3 (10%)	

Note. n=31. *Likert type question on a scale 1-5 (1- strongly disagree and 5 - strongly agree).

** Frequencies (percentages). *** n=30, one answer was not valid

Study Design

Figure 1. shows the research design. Training on peer feedback and how to use ETS Criterion® was provided during the first quarter (first seven weeks of the study) on three different occasions and was guided by the suggestions and findings outlined in previous research (e.g., Hyland & Hyland, 2006; Diab, 2011). The main, combined feedback activity took place during the second quarter: weeks two, five, and seven. The first quarter was used for regular class teaching so that students, especially the first-year students, could learn about writing and get familiar with writing in English.

In the study, we used ETS Criterion® as one of the AWE programs which can facilitate peer's written feedback. This AWE is based on the *e-rater*® scoring engine and provides "both annotated diagnostic feedback and holistic scoring based on level-specific models built from essays pre-scored by ETS-trained readers" (ETS, n/d). It provides diagnostic feedback in four categories: grammar, usage, mechanics, and style. It also flags organization and development aspects of writing; however, it only provides formulaic, built-in comments, i.e., it does not 'recognize' the content/context of each essay.

The holistic score ranges from 1 to 6, or 1 to 4. Essays used for writing were taken from ETS Criterion® topics' library. When choosing the essay's topics, we tried to use topics that were related to what has been taught in the classroom on the day when the homework (paragraph to be used in the study) was assigned. Each student had access to her account on this platform, which she could access at any time during quarters one and two.

Data Analysis

To answer both research questions, how students perceive combined peer and AWE feedback, and how they did peer review activity, we employed a survey and semi-structured interviews. The survey, administered at the end of the second quarter, consisted of closed and open-ended questions and Likert item type questions, which were used to learn about learner's perceptions. Since we had a relatively small number of participants, we did not combine questions to create scales; thus, the data from the responses were treated as ordinal data (Lavrakas, 2008, pg. 429).

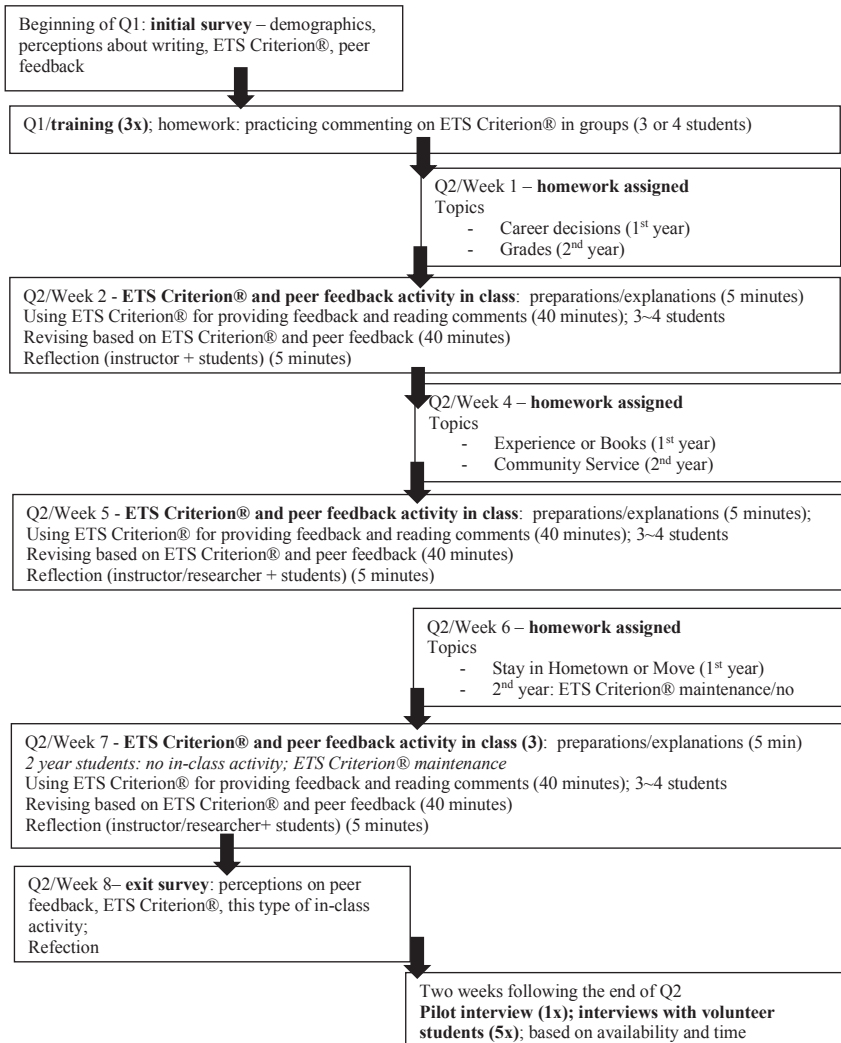


Figure 1. Study Design

As we wanted to further probe responses collected by surveying the participants and get more detailed answers as well as ‘maximize the quality of the data collected’ (Lavrakas, 2008, pg. 259), we used semi-structured interviews. We interviewed six students, whereas the first interview was used to pilot our questions (these answers were not included in the analysis). All the interviewees were volunteers, and we only talked to first-year students. All interviews were done within two weeks after the course in order to ensure that students still had fresh memories about the combined feedback activity. The interviews lasted between 30 and 45 minutes and were conducted by two interviewers. The questions were asked in English and were then translated to Japanese by the second author. Students answered mostly in Japanese, and sometimes in English. All the interviews were translated into English before analysis. The second author, who is a native Japanese speaker and an English language instructor, coded the interview transcript to identify the initial emerging themes. After that, the principal researcher coded the transcripts to ensure that both researchers agree on the constructs.

Results

Perceptions of AWE and Peer Feedback

The students’ overall perceptions of ETS Criterion® and peer feedback can be summarized as follows:

Table 3 *Perceptions of Usability of ETS Criterion® and Peer Feedback (n=31)*

	Strongly disagree (%) (n)	Disagree (%) (n)	Neutral (%) (n)	Agree (%) (n)	Strongly agree (%) (n)	Median (IQR)
Q1 I know how to revise my paragraph based on the feedback provided by ETS Criterion.	0% (0)	23% (7)	10% (3)	52% (16)	16% (5)	4 (0)
Q2 I know how to revise my writing (paragraph) based on the feedback provided by my peers	0% (0)	6% (2)	3% (1)	71% (22)	19% (6)	4 (0)
Q3 In my revised paragraph I used all ETS Criterion® feedback/suggestions/ flagged errors.	0% (0)	10% (3)	23% (7)	45% (14)	23% (7)	4 (0)
Q4 In my revised paragraph, I used all peer’s suggestions	0% (0)	3% (1)	16% (5)	61% (19)	19% (6)	4 (0)

Students viewed peer feedback as easier to understand than ETS Criterion® feedback. This means that the situation in which students did not know how to change their writings happened more often when writing based on ETS Criterion® comments, as evident from answers to Q1 and 2 in Table 3. This result was probably reflected in their answers to the amount of feedback used in revision: students used more peer suggestions (80% combined answers) than ETS Criterion® comments (68% combined answers).

Table 4 *Perceptions of Efficacy of ETS Criterion® and Peer Feedback (n=31)*

	Strongly disagree (%) (n)	Disagree (%) (n)	Neutral (%) (n)	Agree (%) (n)	Strongly agree (%) (n)	Median (IQR)
Q5* I think my English writing ability has improved after using Criterion.	0% (0)	7% (2)	27% (8)	43% (13)	23% (7)	4 (0)
Q6 Reading and evaluating peers' paragraphs helped me to improve my own composition/writing.	0% (0)	6% (2)	3% (1)	71% (22)	19% (6)	4 (0)
Q7 I think that using ETS Criterion® feedback was a waste of time and it did not help me improve my writing.	13% (4)	52% (16)	26% (8)	6% (2)	3% (1)	2 (0)
Q8 I think that using peer feedback was a waste of time, and it did not help me improve my writing	23% (7)	65% (20)	13% (4)	0% (0)	0% (0)	2 (0)

Note. *n=30 (one answer to this question was not valid)

Peer feedback is accepted as not only more user-friendly, but also more effective than ETS Criterion® in the process of improving their English writing ability. In addition, three students (9%) answered that using ETS Criterion® had no influence on improving their English, whereas nobody answered that peer feedback is a waste of time, thus confirming positive perceptions of these two types of feedback.

Table 5 *Preference of ETS Criterion® and Peer Feedback (n=31)*

	Strongly disagree (%) (n)	Disagree (%) (n)	Neutral (%) (n)	Agree (%) (n)	Strongly agree (%) (n)	Median (IQR)
Q9 I prefer peer feedback to ETS Criterion® feedback	3% (1)	16% (5)	61% (19)	13% (4)	6% (2)	3 (0)

Although, as shown in Tables 3 and 4, peer feedback seems to be accepted comparatively more favorably in its utility and efficacy, most students could not decide which they prefer: the ratio of students who agreed (strongly agree + agree) and disagreed (strongly disagree + disagree) to the question was the same as shown in Table 5.

The Use of Peer Feedback

The second research focus of this paper is to see how students can use the peer review effectively in a class when this is combined with AWE. By focusing on this topic, we wanted to investigate unique roles or influences of peer feedback, and which, in nature, are different from those of AWE.

Tables 6 and 7 show whether giving and receiving feedback had different impacts on learning.

Table 6 *Perceptions of Efficacy of Giving and Receiving Peer Feedback (n=31)*

	Strongly disagree (%) (n)	Disagree (%) (n)	Neutral (%) (n)	Agree (%) (n)	Strongly agree (%) (n)	Median (IQR)
Q10 I learned most from writing and providing feedback	0% (0)	6% (2)	42% (13)	45% (14)	6% (2)	4 (1)
Q11 I learned most from receiving peer feedback	0% (0)	6% (2)	45% (14)	39% (12)	10% (3)	3 (1)

Table 7 *Perceptions of Efficacy of Giving and Receiving Peer Feedback (n=31)*

Q12 Which aspect of the peer review activity did you learn the most from? (Choose one answer only)	
Answer Choice	(%) (Number of Answers)
Giving feedback	10 % (3)
Receiving feedback	37 % (11)
Both giving and receiving peer feedback	53 % (16)
Neither giving nor receiving peer feedback	3 % (1)

It seems that students think they learned from both writing and receiving comments. However, in Q12, only three students answered they learned from giving feedback, compared to 11 students who thought receiving peer feedback was a fruitful learning experience. This small number may be connected to their confidence. In another question, 65 % of the students answered they did not feel

confident to provide comments, while only 9 % said they were confident (26 % was neutral). It can be said that students faced difficulties when making comments on peers' paragraphs as they had no previous experience, or they still did not get used to this sort of activity.

Table 8 shows what students focused on when they provided feedback.

Table 8 *Students' Focus to Give Feedback*

Q13 When giving feedback, what did you focus on? (Check all that applies)	
Answer Choice	Number
Grammar	22
Punctuation and spelling	11
Word usage	4
Topic sentence	13
Providing ideas/examples	3

Grammar gained the highest interest when writing peer feedback, and topic sentence and punctuation/spelling are chiefly focused on as the second and third interest respectively. When giving comments, the least attention is paid to the content, e.g., providing ideas/examples.

This result can be compared with the types of feedback students were provided by AWE and were found to be useful when using ETS Criterion® feedback in revising.

Table 9 *Students' Responses of ETS Criterion*

Q14 I think ETS Criterion® helped me with ... (check all that apply)	
Answer Choice	Number
Grammar	27
Punctuation, spelling, capitalization	7
Word usage	25
Organization (topic sentence and supporting sentence)	29
Ideas	9

Students thought that ETS Criterion® was useful especially in terms of grammar and organization. Word usage also gained a high ratio as ETS Criterion® highlights the repeatedly used words and encourages writers to use synonyms (Table 9).

Interview Excerpts

Perceptions of ETS Criterion. Questions asked during the interview were similar to those asked via the survey as we wanted to get a more detailed explanation about some aspects of this combined feedback process. As we used semi-structured interviews, some of the questions were not as same as those asked in the survey because these were guided by students' responses. For some probe questions, check Appendix B. Students sometimes made comments on both ETS Criterion® and peer feedback in one answer, so such students' answers were divided into different categories. Answers in Japanese or unclear English comments were translated or edited by the co-author whose native language is Japanese.

Table 10 shows the excerpts of students' answers on the uptake of ETS Criterion® feedback during revision.

Table 10 *How Much of the ETS Criterion® Feedback Did You Use to Revise?*

Student	Interview excerpts
[Student A]	I used all the feedback except the comments on repetition.
[Student B]	I used most. When I used the same words many times, such as 'I', I didn't change them because I didn't know how. I used all other comments.
[Student C]	To get a higher score, I used all the comments from Criterion. However, if my peers did not understand my English, I changed that part even though my score got worse.
[Student D]	I used it 60 to 70%. As for the rest, 30-40%, I didn't know how to revise. It's not because I wasn't convinced, but just I didn't know how to revise my essay.
[Student E]	I tried to revise every part that Criterion pointed out as much as possible.

Similar to the result of the survey (see Table 3), it can be observed that students tried to use ETS Criterion® comments as far as they understood how to rewrite.

Table 11 *Do You Think Criterion Is Useful to Improve Your Writing?*

Student	Interview excerpts
[Student A]	In my first writing, I didn't know how to write a topic sentence, and while I was writing the supporting sentences and the details, I became able to plan [the paragraph].
[Student B]	I practiced a lot, and Criterion pointed out my errors, [so] I became able to rewrite by myself when I thought I made mistakes.
[Student C]	... it was helpful that Criterion pointed out my mistakes about the topic sentence and supporting sentences. I couldn't realize such mistakes by myself. ...so Criterion was more useful for this (i.e., organization).

[Student D]	Criterion told me to write more than 3 sentences to support the main ideas. It was helpful. ... Criterion pointed out that I repeated words ... so I came to use synonyms carefully.
[Student E]	In the comments, by Criterion, I was sometimes pointed out that I have a subject-verb-agreement mistake. So when I make the same mistakes in google classroom, I tried to check by myself. And editing, finding mistakes became easier for myself.

From the excerpts above, in summary, we can see that ETS Criterion® helped them with (1) organization (topic sentence and supporting sentence), (2) using synonyms, and (3) finding errors by themselves.

Perceptions of Peer Feedback. Table 12 shows excerpts from the interviews related to the uptake of peer feedback during rewriting.

Table 12 *How Much of the Peer Feedback Did You Use?*

Student	Interview excerpts
[Student A]	I rewrote every part if I got comments from peer editing.
[Student C]	My friends sometimes pointed out that they did not understand my essay when they read it. So I thought it was better to rewrite and use easier words to understand. So I used my friends' comments 100% percent.
[Student D]	As for the comments which indicated that I should rewrite, I rewrote almost every part based on peer comments.
[Student E]	As for peer feedback, I read comments by my classmates, and I rewrote when I thought I was wrong. If I couldn't understand their comments, I asked them, and checked the dictionary, and then tried to write. So I used about 80% comments by peers. The other 20% was almost comments about the content.

One student did not give any estimates on how much of peer feedback she took and used during the revision. All other students accepted almost all of the peer comments in subsequent revisions. In addition, as *Student E* mentioned, unlike the ETS Criterion®, if students did not understand how to rewrite, they were able to ask the peers directly. This may be one of the reasons why students used peer feedback more than ETS Criterion® (see Table 3).

The following excerpts provide valuable insight into the students' responses to the utility of peer feedback.

Table 13 *Do You Think Peer Feedback Is Useful to Improve Your Writing?*

Student	Interview excerpts
[Student A]	I don't think it was helpful that much. <i>The classmates checked only word mistakes, I thought I was able to find such mistakes by myself</i> [emphasis added by the authors]. ... The class activity (i.e., teacher feedback) helped more to improve my writing, not the feedback.
[Student B]	<i>It became easier to think of an idea than before...</i> Because I read my classmates' ideas, and it affected me...
[Student C]	I think peer feedback was useful because ... I was able to understand what my classmates didn't understand about my writing. So I was able to realize my mistakes. And I learned something from my classmates' essays. ... <i>I came to think what my friends think when they read my writings, so I was able to reduce word mistakes. And I came to think what the best way to convey the idea is. So, I came to be more careful when I write...</i>
[Student D]	<i>There were some expressions I wanted to use. Comments from classmates were not so useful. I thought, I didn't have enough to write to give feedback to my classmates. And I couldn't receive enough comments. ... I could learn from reading other people's essays, but not from reading their comments.</i> I am not good at giving comments, and I am not used to it. And comments from my classmates were not so substantial.
[Student E]	<i>I tried to remember the good expressions.</i> So the peer feedback activity was interesting. ... <i>Essays written by other classmates were intriguing for me. They motivated me to use better expressions</i> ... Peer review was helpful to check my organization so I could add more ideas.

In the survey (see Table 6 and 7), students thought that both receiving peer feedback and giving feedback were effective, but receiving was perceived as slightly more effective compared to giving feedback. The interview answers above, however, depicted a different picture. *Students A* and *E* answered that peer comments were not useful for them to learn writing as they did not get enough comments; also, they expected these comments to be more substantial both in quality and quantity. What can be observed from the interviews with *Students B, D,* and *E* is that one of the most important aspects for the students in this activity was to read others' 'good' examples or ideas. In addition, *Students B* and *C*'s answers show that they were aware of their role of readers in this process.

Finally, the excerpts in Table 14 reveals what aspects students focused on as reviewers.

Table 14 *When Giving Feedback, What Did You Focus on?*

Student	Interview excerpts
[Student A]	Only spelling and punctuation. <i>I was not so confident to point out the connection between the topic sentence and supporting sentences.</i> So I indicated only the part where I was sure of mistakes.
[Student B]	Grammar. It was difficult to comment on other aspects as I was not confident if I was right.
[Student C]	<i>I was not confident to point out about topic sentence or supporting sentences to my classmates' essays, but it was easier to realize grammar mistakes.</i> So when I commented on my friend's essay, I focused on grammar mistakes.
[Student D]	Topic sentence and paraphrasing*. As for topic sentence, I only commented 'good' as I thought it was the writers' idea.
[Student E]	Giving ideas. It was difficult for me to point out grammar mistakes. So I focused on commenting the content. When the flow of the story was difficult to understand because the writer didn't give enough examples, I commented that it's better to insert some explanations.

Note. *The student mistakenly used the word 'paraphrasing' in the interview and the word she intended to use was 'synonym' as the interview context shows.

Overall, students showed less confidence in giving comments. Therefore, they were prone to give feedback on what was easier for them to correct, such as spelling/punctuation, grammar, topic sentence or content. In the survey (see Table 8), many students focused on comments related to the organization (topic sentence) when giving comments, but in the interview, only *Student D* stated that she focused on the topic sentence. However, the comments of *Students A* and *C* reveal that they might have paid attention to a topic sentence, but they could not make any comments as they felt unconfident in their writing ability.

Discussion and Conclusion

To answer our first research question about the perceptions of usefulness and effectiveness of peer feedback provided by peers in combination with AWE for learning English writing, we used survey results and interviews with five students. As for the usability, according to the result of the survey, although there were no significant differences between ETS Criterion® and peer feedback (see Table 3), comments from peers were accepted more favorably compared to those provided by AWE. This is consistent with previous research where students perceived humans as a more preferred source of feedback.

Studies comparing teacher feedback and AWE feedback (Chang, 2016; Hyland & Hyland, 2006; Yu & Lee, 2016; Baierschmidt, 2012), as well as studies comparing AWE and peer feedback (examples in Chang, 2016) provided similar results.

As the interview excerpts indicate (see Table 10), one possible explanation for this may be that students faced difficulties in understanding the AWE feedback. Similar results were reported in previous studies, e.g., Huang (2014), Chen and Cheng (2008). Learners in our study did not know how to use the AWE feedback; consequently, they did not know how to make revisions. It can be speculated that this is due to the nature of ETS Criterion® feedback, which is pre-determined and formulaic, especially in terms of organization or rhetorical aspects of writing. Also, ETS Criterion® uses metalinguistic explanations and terminology that students might not be able to understand due to their proficiency levels and/or previous English language education. Since we did not look at the relation between the types of feedback which students could understand and use, or the type of feedback students could give and proficiency levels, this should be examined in a new research project.

On the other hand, when engaging in peer review aspect of this combined feedback activity, students were able to interact and orally exchange the ideas. This may explain why peer feedback was comparatively easier to understand and use for revising. Furthermore, after discussion, as *Student E* said in the interview (see Table 12), if the writer thought rewriting was unnecessary, she could retain that part and made no changes to her text. This is an important finding, as this indicates that these students understood the underlying premise of the writing process: it's a communicative practice intended for an audience, and negotiation of meaning is an integral part of it. Additionally, it may be interpreted that this kind of activity had some positive effects on learner autonomy: students were in control, up to a certain degree, of what they changed or not. Moreover, they reflected on the language they had to use and how this can be used later. For example, at least one of the students, *Student E*, indicated that this careful language consideration and checking for mistakes by herself was something she started to use in assignments she completes on *Google*

Classroom for other AEP classes (see Table 11).

To conclude, although the comments both from ETS Criterion® and peers were not sometimes accepted, the reason seems to be different: in the case of ETS Criterion®, students did not know how to change it, but for peer comments, students intentionally decided not to change.

Students felt that both ETS Criterion® and peer feedback have an effective influence on learning writing, but as shown in the questionnaire (see Table 4), peer feedback has a more favorable reception than AWE, though the difference is not remarkable. The survey (see Table 9) and interview excerpts (see Table 11) clarify that by using ETS Criterion®, students felt that they improved in terms of writing skills, especially in organization, vocabulary, and grammar. This is somewhat different from previous studies; for example, in Ohta (2008), students also perceived ETS Criterion® positively, but they said they expected more explicit rhetorical feedback. Also, learners who participated in Otsoshi's study (2012) acknowledged possible effects of ETS Criterion® on their writing, but still perceived teachers as responsible for providing feedback on content and organization, as well as vocabulary. As our findings differ from the previous studies, it would be interesting to see why this happened. One way to do so is to look at students' paragraphs and accompanying revisions they submitted after the feedback, and see if there was any improvement in organization of their paragraphs, for example, topic sentences and supporting ideas, and how much was due to peer feedback and how much can be attributed to the use of AWE.

Concerning the efficacy and utility of peer feedback, results are mixed. The main issues are that students believed that feedback received from their peers was not substantial both in quantity and quality, and their classmates seemed to focus on surface-level mistakes and not global writing issues such as rhetorical aspects of writing. *Students A* and *D*'s responses (see Table 13) indicate that for students who felt peer comments were not substantial enough, receiving peer feedback was not helpful. This finding is consistent with previous research, which states that student reviewers whose skills are lower tend to provide less feedback (e.g., in Allen and Katayama, 2016). It is interesting to note that some students felt not only that comments they received were not sufficient, but they also

acknowledge they did not give enough comments either, e.g., *Student D*, Table 13.

However, participants did find usefulness in the peer feedback aspect of this combined activity. Both giving and receiving comments were useful to improve writing competence (see Table 6 and 7). What can be claimed from the interviews is that reading the paragraphs written by classmates was a meaningful experience for most students. Perceived benefits are expanding expressivity, generating more, or deepening the existing ideas. Other studies also found that the act of reviewing or reading peer's writing was potentially more helpful as learners could learn a new language, sentence structure, and organization, and then compare their writing to that of their peers (Yoshikawa, 2016; Hyland & Hyland, 2006). It can be speculated that reading was prominent for our participants because they just started to write in English and their level of proficiency was relatively low. This is in line with what Allen and Mills (2016) suggested about reading as a mediation tool for lower proficiency writers. Reading English compositions written by other students, which is an initial part of the peer-review procedure, can stimulate learners' thoughts and provide profound insight, even if they cannot give any tangible feedback. These effects cannot be brought about by the activity using only AWE, and this means that peer feedback can be utilized effectively in combination with AWE in the way it supplements the shortcoming of AWE.

The interviewees' answers in Table 14 reveal that when students tried to give feedback, they made comments on what they were comparatively confident in. On the whole, as Table 8 shows, their chief interest was drawn to grammar, topic sentence and punctuation/spelling. On the other hand, few students tried to make suggestions on content, which is an area where AWE is the least helpful according to the previous research as machines cannot understand the context.

What is noteworthy here is that these two aspects (grammar and topic sentence) are the same aspects that students thought they learned the most from the ETS Criterion® feedback (see Table 9). This might be the result of students' using ETS Criterion® feedback as a model of peer feedback with or without intention. That is, as students probably were constantly getting comments on these points from ETS Criterion®, they 'mirrored' this when they

were giving feedback. As a natural learning process, it became easier for them to pay attention to and find errors about the same aspects of writing as pointed by AWE and not others as they were exposed to these aspects more. In this respect, one implication is that, when peer feedback is conducted together with AWE, to be beneficial, the difference between these two methods in its characteristics and roles should be emphasized. For example, in the early phases of the course, instructors should explicitly explain and teach students about the limitations of an AWE system to enable learners to become aware of and alleviate those shortcomings. Students need to be trained to focus on the points which AWE does not comment on, such as coherence, logic or content. Other studies also highlighted the importance of training (e.g., in Hyland and Hyland, 2006; Diab, 2011) in order to prepare learners for giving feedback, as this potentially can lead to more high-quality feedback, better revision and more learning from this process.

To conclude, students perceive both AWE and peer feedback useful and effective when it comes to learning how to write. Thus, combined feedback is beneficial as an instructional tool for writing and can help lower proficiency students. Unfortunately, against our expectations, students did not pay attention to the rhetorical aspects of writing during peer review activity. Therefore, in order to use this combined feedback effectively, learners need to be taught how to recognize the different advantages and limitations which both AWE and peer feedback have.

There are several limitations to this study. The number of participants was relatively small, so it is difficult to apply conclusions beyond the given context. However, the results could apply to Japanese learners of English in a similar educational background, such as proficiency levels and approaches to teaching writing. Also, results and conclusions presented here are part of a larger study, which looks at the type of comments given by the students and received from AWE, and how that affected revision. It is important to look at this as it will go beyond students' perceptions and we will be able to see changes made in texts. This, in turn, will help instructors to tailor more precise training and have a more successful use of the AWE and peer review activity in the classroom.

References

- Allen, D., & Katayama, A. (2016). Relative second language proficiency and the giving and receiving of written peer feedback. *System*, 56, 96–106. doi:10.1016/j.system.2015.12.002
- Allen, D., & Mills, A. (2016). The impact of second language proficiency in dyadic peer feedback. *Language Teaching Research*, 20(4), 498–513. doi:10.1177/1362168814561902
- Baierschmidt, J. (2012). Japanese ESL learner attitudes towards peer feedback. *Kanda University Bulletin*, 24, 101-114. Retrieved from <http://id.nii.ac.jp/1092/00000605/>
- Burstein, J., Tetreault, J., & Madnani, N. (2013). The E-rater® automated essay scoring system. In M.D. Shermis & J. Burstein, (Eds.), *Handbook of automated essay evaluation: Current Applications and New Directions* (pp. 77-89). Routledge.
- Chang, C. Y. (2016). Two decades of research in L2 peer review. *Journal of Writing Research*, 8(1), 81–117. doi:10.17239/jowr-2016.08.01.03
- Chen, C.-F., & Cheng, W.Y. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning and Technology*, 12. Retrieved from https://scholarspace.manoa.hawaii.edu/bitstream/10125/44145/12_02_chencheng.pdf
- Chen, T. (2014). Technology-supported peer feedback in ESL/EFL writing classes: A research synthesis. *Computer Assisted Language Learning*, 29(2), 365–397. doi:10.1080/09588221.2014.960942
- Diab, N. M. (2011). Assessing the relationship between different types of student feedback and the quality of revised writing. *Assessing Writing*, 16(4), 274–292. doi:10.1016/j.asw.2011.08.001
- Deane, P. (2013). Covering the construct: An approach to automated essay scoring motivated by a socio-cognitive framework for defining literacy skills. In M.D. Shermis & J. Burstein, (Eds.), *Handbook of automated essay evaluation: Current Applications and New Directions* (pp. 298-312). Routledge.
- Elliot, N., & Klobucar, A. (2013). Automated essay evaluation and the teaching of writing. In M. D. Shermis & J. Burstein (Eds.), *Handbook of automated essay evaluation: Current applications and new directions* (pp. 16–35). Routledge.
- ETS Criterion® (n/d). How does the Criterion® online writing evaluation service work? Retrieved from <http://www.ets.org/criterion/about/how/>
- Fujii, A., Ziegler, N., & Mackey, A. (2016). Peer interaction and metacognitive instruction in the EFL classroom. In M. Sato & S. Ballinger (Eds.), *Language Learning & Language Teaching*, 45, 63–89. doi:10.1075/llt.45.03fuj
- Godwin-Jones, R. (2018). Second language writing online: An update. *Language Learning & Technology*, 22(1), 1-15. Retrieved from https://scholarspace.manoa.hawaii.edu/bitstream/10125/44574/1/22_01_godwin-jones.pdf
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*,

- 77(1), 81-112. doi:10.3102/003465430298487
- Heffernan, N., & Otsoshi, J. (2015). Comparing the pedagogical benefits of both Criterion and teacher feedback on Japanese EFL students' writing. *JALT CALL Journal*, 11(1), 63-76. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1107988.pdf>
- Hu, G. (2005). Using peer review with Chinese ESL student writers. *Language Teaching Research*, 9(3), 321-342. doi:10.1191/1362168805lr169oa
- Huang, S.J. (2014). Automated versus human scoring: A case study in an EFL context. *Electronic Journal of Foreign Language Teaching*, 11, 149-164. Retrieved from <http://e-ft.nus.edu.sg/v11s12014/huang.pdf>
- Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching*, 39(2), 83-101. doi:10.1017/S0261444806003399
- Kohro Y. (1995). An approach to the instruction of paragraph and essay writing using peer correction to enhance learners' awareness of readers. *Daigaku eigo kyoikugakai kiyou Japan Association of College English Teachers bulletin*, (26), 47-60. Retrieved from http://dl.ndl.go.jp/view/download/digidepo_10500542_po_ART0005017057.pdf?contentNo=1&alternativeNo=
- Koizumi, R., Asano, K., & Agawa, T. (2016). Validity evidence of Criterion® for assessing L2 writing proficiency in a Japanese university context. *Language Testing in Asia*, 6(1), 1-26. doi:10.1186/s40468-016-0027-7
- Long, R. (2013). A review of ETS's Criterion online writing program for student compositions. *The Language Teacher*, 37(3), 11-18. Retrieved from: http://jalt-publications.org/files/pdf-article/37.3tlt_art2.pdf
- Lavrakas, P. J. (Ed.). (2008). *Encyclopedia of survey research methods*. Thousand Oaks, California: SAGE Publications.
- Ohta, R. (2008). Criterion: Its effect on L2 writing. In K. Bradford Watts, T. Muller, & M. Swanson (Eds.), *JALT2007 Conference Proceedings* (pp.718-723). Tokyo: JALT.
- Otsoshi, J. (2012). Use of ETS Criterion to prepare for the TOEFL iBT independent writing task. *The Bulletin of the Writing Research Group, JACET Kansai*, 10, 13-24.
- Suzuki, M. (2008). Japanese Learners' self revisions and peer revisions of their written compositions in English. *TESOL Quarterly*, 42(2), 209-233. doi:10.1002/j.1545-7249.2008.tb00116.x
- Taoka, C. (2015). Recategorizing ETS's Criterion® Topics for More Effective Topic Analyses. *Kobe College Studies*, 62(2), 189-199. doi:10.18878/00002524
- Tsuda, N. (2014). Implementing Criterion (automated writing evaluation) in Japanese college EFL classes. *Language and Culture: The Journal of the Institute for Language and Culture*, 18, 25-45. doi:10.14990/00000561
- Vojak, C., Kline, S., Cope, B., McCarthey, S., & Kalantzis, M. (2011). New spaces and old places: An analysis of writing assessment software. *Computers and Composition*, 28(2), 97-111. doi:10.1016/j.compcom.2011.04.004

- Wakabayashi, R. (2013). Learners' roles in a peer feedback task: Do they view themselves as writers or reviewers? *Journal of Asia TEFL*, 10, 31–57. Retrieved from http://www.asiatefl.org/main/download_pdf.php?i=55&c=1404265236&fn=10_3_02.pdf
- Yakame, H. (2005). The role of peer feedback in the EFL writing classroom. *ARELE: Annual Review of English Language Education in Japan*, 15, 101-110 doi:10.20581/arele.16.0_101
- Yoshikawa, O. K. (2016). Student perception on giving peer feedback in EFL writing and its implications. *Journal of foreign education*, 13, 85-105. Retrieved from <https://pdfs.semanticscholar.org/3314/735c8725ef5be0a19c84b46b4c3ba0351486.pdf>
- Yu, B. B. (2015). Incorporation of automated writing evaluation software in language education: A case of evening university students' self-regulated learning in Taiwan. *International Journal of Information and Education Technology*, 5(11), 808-813. doi:10.7763/IJiet.2015.V5.616
- Yu, S., & Lee, I. (2016). Peer feedback in second language writing (2005-2014). *Language Teaching*, 49, 461–493. doi:10.1017/S0261444816000161

Appendices

Appendix 1: Exit Survey (End of Q2)

On a scale from 1 (strongly disagree) to 5 (strongly agree) decide about yourself:
1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree

- Q1 I know how to revise my paragraph based on the feedback provided by ETS Criterion®.
1 2 3 4 5
- Q2 I know how to revise my writing (paragraph) based on the feedback provided by my peers
1 2 3 4 5
- Q3 In my revised paragraph, I used all ETS Criterion® feedback/ suggestions/ flagged errors.
1 2 3 4 5
- Q4 In my revised paragraph, I used all peer's suggestions.
1 2 3 4 5
- Q5 I think my English writing ability has improved after using ETS Criterion®.
1 2 3 4 5
- Q6 Reading and evaluating peers' paragraphs helped me to improve my own composition/writing.
1 2 3 4 5
- Q7 I think that using ETS Criterion® feedback was a waste of time, and it does not help me improve my writing.
1 2 3 4 5
- Q8 I think that using peer feedback is a waste of time and it does not help me improve my writing.
1 2 3 4 5

Q9 I prefer peer feedback to ETS Criterion® feedback

1 2 3 4 5

Q10 I learned most from writing and providing feedback

1 2 3 4 5

Q11 I learned most from receiving peer feedback

1 2 3 4 5

Q12 Which aspect of the peer review activity did you learn the most from?
(Choose one answer only)

Giving feedback

Receiving feedback

Both giving and receiving peer feedback

Neither giving nor receiving peer feedback

Q13 When giving feedback, what did you focus on? (Check all that applies)

Grammar

Punctuation and spelling

Word usage

Topic sentence

Providing ideas/examples

Q14 I think ETS Criterion® helped me with ... (check all that applies)

Grammar

Punctuation, spelling, capitalization

Word usage

Organization (topic sentence and supporting sentence)

Ideas

Appendix B: Interview Questions/Prompts (End of Q2)

How much of the ETS Criterion® feedback did you use to revise?

Do you think ETS Criterion® is useful to improve your writing?

How much of the peer feedback did you use?

Do you think peer feedback is useful to improve your writing?

When giving feedback, what did you focus on?