DIFFERENT EFFECTS OF SAMPLE PERFORMANCE OBSERVATION BETWEEN HIGH AND LOW LEVEL ENGLISH LEARNERS

Yasuko Okada

(okada92025@gmail.com)

Seisen University, Japan

Takafumi Sawaumi

(t.sawaumi@gmail.com)

The University of Tokyo, Japan

Takehiko Ito

(take@wako.ac.jp)

Wako University, Japan

Abstract

The present study explores the effects of using sample videos of previous students' performance on the improvement of presentation skills, comparing effects of the observational learning in high versus low English proficiency classes in the English as a Foreign Language (EFL) classroom. Sample performance videos were shown to two classes alternately. Both classes first practiced delivering oral presentations without using videos, and then the high proficiency class ($n = 12$) watched the videos prior to the second presentation while the low proficiency class ($n = 17$) watched the videos prior to the third presentation. Peer and self-evaluations were conducted on each occasion and the ratings of the first and the second presentations were statistically analyzed to determine the effects of using the sample performances. Additionally, two open-ended video reviews and feedback about the sample observation were qualitatively analyzed using text mining. The results showed that there was a statistically significant difference in the effects on content between the two classes. Additionally, it was revealed that for the higher proficiency class, it was effective in enhancing the motivation to develop language and presentation skills, whereas the low proficiency class had difficulty taking advantage of the sample video observation.

1 Introduction

The purpose of this paper is to demonstrate the effects of social modeling in a Japanese EFL context, comparing high English proficiency and low English proficiency classes in which learners' oral presentation skills are developed. For language learners, it is essential to review their own performance and understand the strengths and weaknesses of their own presentations. The strengths can be maintained, and the shortcomings can be improved for the following presentation. Recent development of video technology enables teachers to easily and clearly record students' performance with a digital video camera and to analyze the performance in order to explore methods of teaching oral presentation skills. This study applies a theoretical framework of the social modeling theory (Bandura, 1986, 2005), where

human beings watch and model others when learning a new behavior. By fully understanding the effects of observational learning with the sample performance video, learners may not only be able to analyze what is missing in the previous students' performance, but they may also be better equipped to enhance their overall performance by developing the various aspects of verbal and non-verbal behavior as well as feeling a sense of accomplishment. In this paper, the terms observational learning and social modeling will be used interchangeably.

2 Literature review

2.1 Theoretical framework

Drawing upon Bandura's (1986, 2005) theory of social modeling, Grez, Valcke, and Roozen (2014) investigated how observational learning improves oral presentation skills. Bandura states that the learning process begins with watching and imitating other persons' behavior that would be most likely to result in a positive outcome (positive reinforcement). Based on the learning method by means of observation, a few studies have delved into the effects of observational learning on oral presentation skills. Nonetheless, no research has empirically compared those who learn by social modeling with those who learn by mere practice until Grez et al.'s publication. Grez et al. applied the theoretical framework to the Dutch-speaking context at a Belgian university and contrasted the modeling condition with the practice only condition.

Results revealed the following patterns. First, seven out of nine oral performance criteria improved at the third presentation as compared with the first presentation. Second, there was no overall significant differential effect at the final point of testing between the two conditions, implying that overall progress achieved by those who started with observational learning and then practice was not significantly different in the end from the progress of those who started practicing and then observational learning. Third, participants in the "observational learning first" condition improved on the content of their presentation between the first and the second presentation phases. They therein went through observational learning, and improved only little thereafter whereas participants in the "practice-based learning first" condition improved modestly between the first and the second presentation phases, and improved strongly after taking part in observational learning. Fourth, no significant effect was revealed as to the interaction between learners' characteristics and conditions on oral presentation performance.

The above analyses partly supported Grez et al. (2014) key hypothesis that learners beginning with observational learning would perform better than those starting with only practice by themselves. They concluded that improvement in oral presentation was greater at the third presentation phase, where all participants experienced both observational and practice learning. The improvement was notable on all criteria while the progress was greater on the content level (e.g., content of the conclusion) rather than on the delivery level (e.g., eye contact).

2.2 Video Recordings as Educational Tools

Due to the development of digital technology, video recordings of learners' performance have been commonly implemented in various educational contexts. Earlier studies have demonstrated that using digital video cameras has a positive effect on learners' presentation

skills. For example, video recordings can be effective to enhance learners' oral presentation skills, such as verbal and non-verbal communication skills, structure of the presentation, and engagement of the audience (Guo, 2013, p. 99). Having opportunities to review learners' recorded performances can also increase learners' engagement in presentation, successful communication, and career-related skill development (Tugrul, 2012). These studies suggest that reviewing the recordings can develop learners' presentation skills that are practical not only in the classroom but also in the world outside.

Video offers several possibilities to language learning. For one thing, digital video recordings allow students to reflect the performance on the development of their language and cognitive skills. Hung (2009) suggests that video recordings can help learners become cognitively reinforced and affectively engaged in language learning through video as mediation (p. 186). To explore learners' attitudes toward the use of video and the advantages and disadvantages of using video in the EFL classroom, Hung conducted a study with 26 native speakers of Chinese at a Taiwanese university. The analysis of learners' reflective journals, audio-recorded interviews, and course evaluation revealed that an instructional video can be an aid in language learning, that it can focus on their strengths and weaknesses through visual information, and that learners can eventually expand their knowledge of the language, enhancing their language awareness (p. 182). McNulty and Lazarevic's (2012) study highlighted ways to develop learners' pronunciation and presentation skills by using video technology at a secondary school in the U.S. They posit that the video may motivate learners to practice pronunciation and presentation skills more when they view their own video-recorded performances.

In previous studies of video technology in Japanese EFL contexts (Okada, 2011, 2012, 2013; Okada & Ito, in press), the benefits of using video recordings were discussed, for example, discovering the weaknesses of the learners' performance, such as pausing and stressing words (Okada, 2011, 2012), learning a positive attitude towards viewing and evaluating students' video recorded performance (Okada, 2013), and focusing more on visual information than audio information (Okada & Ito, in press). The results of these studies suggest that the use of video recording may be applicable to the Japanese EFL context. Although there have been some initial studies exploring the use of video recordings, there has been little research as to the effects of video technology based on Bandura's (1986, 2005) social modeling theory. This provides a fertile opportunity to include this study and indicates the need for both qualitative and quantitative research.

3 Research question

In light of the review of Bandura's (1986, 2005) social modeling, in which he stresses the importance of watching and imitating other persons' behavior to bring about a positive influence on a person's future attitude and actions, the purpose of this study is to examine whether there is an observational effect on learners' own performance in the foreign language. This study will be a refined replication of the Grez et al. (2014) study in order to incorporate the foreign language setting. We will first examine whether the sample performance observation improves learners' oral presentation performances.

Additionally, the previous study did not take into consideration learners' language proficiency; this study, by contrast, examined two classes representing two levels of English

proficiency. Therefore, it is meaningful to examine how the high and the low level learners of English view the sample performance videos and what they learn from observing the previous students' performances in the foreign language classroom.

4 Method

4.1 Participants

Twenty-nine Japanese students participated in the study. They were enrolled in two different English communication classes in spring 2014, taught by the first author at a university in the Tokyo metropolitan area. All were freshmen majoring in economics at the time of the study and had been studying EFL for at least six years. Although the study started with 31 students in two classes, one student had insufficient data, and another did not consent data use for the research. Therefore, data from these two students were discarded in the subsequent analyses.

At the beginning of the semester, the participants were placed into each class based on their scores on the Test of English for International Communication (TOEIC) Bridge test. The average scores were 142 for the high level class (8 males and 4 females) and 116 for the low level class (12 males and 5 females). There was a statistically significant difference in test scores between the high and the low English proficiency classes, t (16.42) = 6.40, $p < .001$. For gender and class size, there was no significant difference between the two groups.

4.2 Setting

The course was a mandatory class for graduation, with an emphasis on the development of oral communication skills through group and pair work. The ninety-minute long class met twice a week during the 14-week semester. Both the high and the low level classes covered the same content and used the same textbook. After being explained the nature of the study by the instructor, the students clearly understood that their performance in the classroom was going to be part of a research project, though it included 30% of the course grade. The students were also informed that participation in the research would not affect their grade, nor would the study require any additional work or time.

4.3 Procedures

In the course, students were strongly encouraged to give three memorized oral presentations in Weeks 3-1, 8-1, and 13-1, as shown Figure 1.[1] The topic of each presentation was "Self-introduction," "What I want to recommend," and "What I would like to accomplish during my college life" respectively. For each oral presentation, students' performances were video recorded while presenting them to the class, and then the recorded performance was self- and peer evaluated during the following lesson. Each student was instructed to evaluate all the students in the class. It should be noted that training for peer rating was not conducted since the focus of the study was to examine effects of observational learning with videos.

---

[1]   Week 3-1 indicates that the session was held in the first class of the third week, since the class met twice a week.

High Proficiency Class	Low Proficiency Class
($n = 12$)	($n = 17$)

Week 3-1 — 1st oral presentation

Week 3-2 — 1st self- and peer assessment (pretest)

Week 7-2 — Sample video observation (intervention) / Sample video reviews

Week 8-1 — 2nd oral presentation

Week 8-2 — 2nd self- and peer assessment (posttest 1)

Week 12-2 — Sample video observation (intervention) / Sample video reviews

Week 13-1 — 3rd oral presentation

Week 13-2 — 3rd self- and peer assessment (posttest 2)

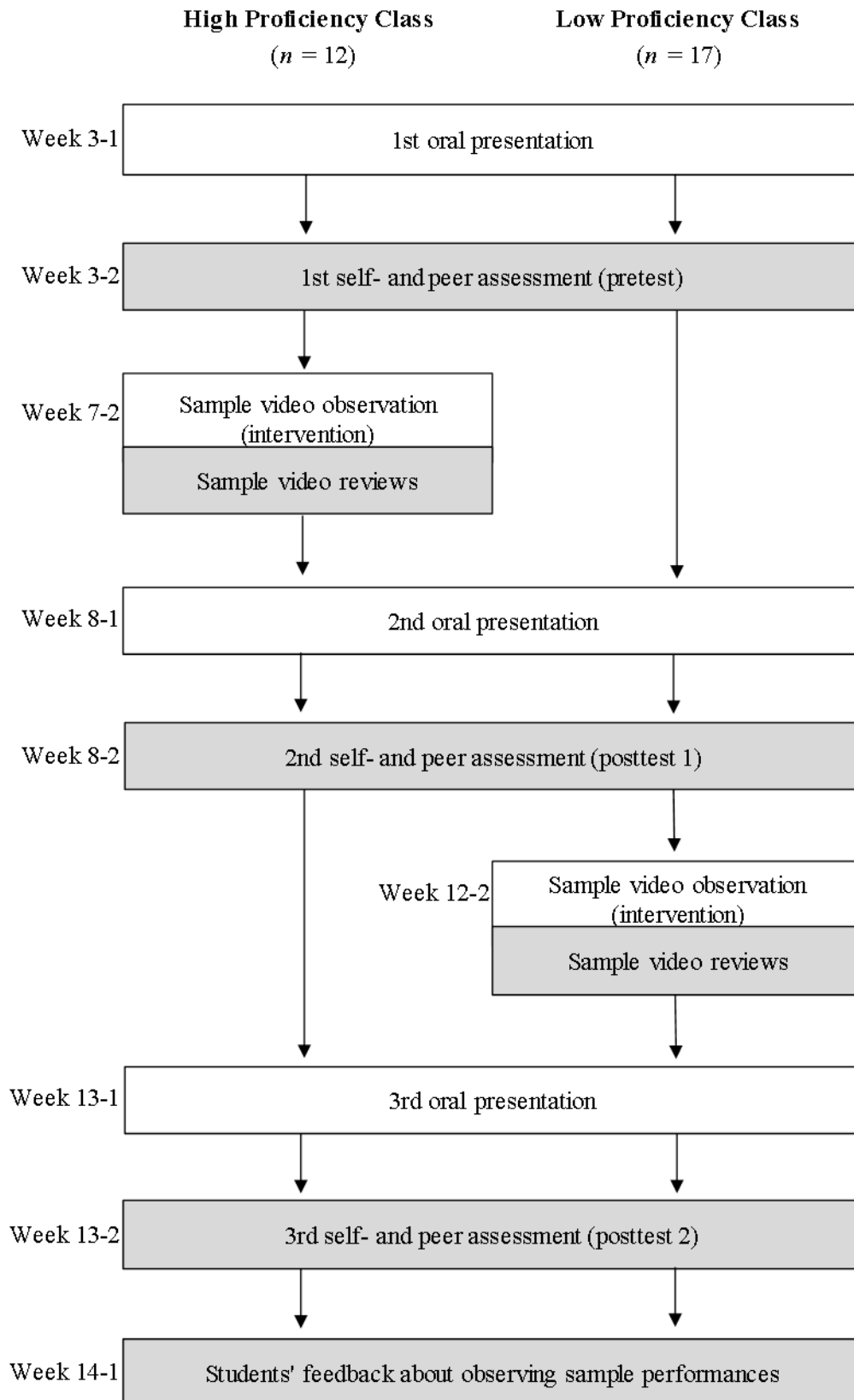Week 14-1 — Students' feedback about observing sample performances

Figure 1. Overview of the Research Procedure

Additionally, students were asked to fill in the two open-ended question forms, one of which was administered after students completed watching the sample performances of the past students in Week 7-2 for the high proficiency class and in Week 12-2 for the low proficiency class. For each class, students were given approximately one to two minutes to fill in the form, describing the strengths and weaknesses of the speaker on the video after viewing each sample video.

Sample videos were chosen from the previous students' performances, and nine samples (four males and five females) were shown to the participants in the study. The length of each sample video was two to three minutes, and the topic was slightly different from that of the study. After viewing and commenting on each of the three videos, students were divided into groups of four or five to discuss the videos for approximately two minutes. It is noted that the order of videos was either male–female–male or female–male–female so students could easily identify the sample videos they were talking about during the small group discussion.

The second open-ended question form was conducted in Week 14-1 after students completed their third oral presentation and self- and peer evaluation. Students were given 5 to 10 minutes to fill in the form.

4.4 Instruments

4.4.1 Evaluation form

The first author developed an evaluation form (see Appendix A) to make a judgment of voice control, body language, speech content, and effectiveness of the presentations on a four-point scale from 1 ("strongly agree") to 4 ("strongly disagree"). The evaluation items that appeared in Okada & Ito (in press) were used in both self- and peer evaluations in the study. The items were extracted from an earlier study (Yamashiro & Johnson, 1997). Both self- and peer evaluation forms were explained in Japanese. There was a total of 14 variables used in the study: Items 1 to 4 concerned voice control such as projection, pace, intonation, and diction; Items 5 to 8 asked about body language such as posture, foot and hand positions, eye contact, and facial expression; Items 9 to 11 asked about the content regarding introduction, body, and conclusion; Items 12 to 14 were related to effectiveness such as topic choice, language use, and vocabulary. A feedback section (Item 15) on the evaluation provided learners with the opportunity to elaborate on what they noticed in their peers' and their own oral presentations.[2]

4.4.2 Open-ended question forms

The two different open-ended question forms were sample video reviews and student feedback about sample video observation, both of which were employed to explore the effects of sample performance observation and differences in impressions of observing sample videos between the high English proficiency and the low English proficiency classes. Sample video reviews contained two questions: (1) What are the strengths of the speakers on the videos? and (2) What are the weaknesses of the speakers on the videos? Student feedback about

---

[2]   Analysis of comments in Item 15 was not included in this paper because it focuses on the qualitative analysis of video sample reviews and student feedback about the samples.

sample video observation included the following two questions: (1) What did you think about observing sample videos of the past students' work? and (2) Was there any difference in students' performances before and after observing sample performance videos? Students were asked to respond to the forms in Japanese.

4.5 Data analysis

The research questions were examined through quantitative and qualitative analyses. For quantitative analysis, a $2 \times 2 \times 2$ multivariate analysis of variance (MANOVA) was performed to examine the mean differences in terms of the sample video effect for the low proficiency and the high proficiency groups (Group), for those of self- and peer evaluations (Rater), and those of the first oral presentation and the second oral presentation (Time) as independent variables.[3]

To simplify the analysis, 14 items on the evaluation form were aggregated and referred to as the following subscales: voice control (Items 1 to 4), body language (Items 5 to 8), speech content (Items 9 to 11), and effectiveness (Items 12 to 14). The four types of subscales were dependent variables. It should be noted that, as the scale originally ranged from 1 (strongly agree) to 4 (strongly disagree), we have reversed it so that higher scores indicate strong agreement. The Cronbach alpha coefficient of each variable was .74, .68, .80, .69 for the first presentation and .76, .79, .85, .73 for the second (both self-evaluated); the Cronbach alpha coefficient was .88, .69, .86, .95 for the first presentation and .80, .36,[4] .84, .62 for the second (both peer evaluated). An alpha level of $p < .05$ was used for statistical significance in all inferential statistics. These quantitative analyses were conducted using IBM SPSS 20.0.

For qualitative analysis, students' responses to the two open-ended question forms were analyzed by means of text mining with the help of Text Mining Studio 4.2 by NTT DATA Mathematical Systems Inc. and the results between the high level and the low level classes were compared.

5 Results

The results are divided into four sections: (1) quantitative results of self- and peer evaluation, (2) results of text mining of sample video reviews, (3) results of text mining of the student feedback about observational learning, and (4) results of content analysis of the student feedback.

---

[3] It should be noted that the mean differences of self- and peer evaluations for the third presentation were not analyzed because we focused on the effects of observing the sample performance video in the two groups in this study.

[4] While the Cronbach alpha coefficient scale is ideally above .70, the scale for body language in the second peer presentation was relatively low. Mean scores of Items 5 to 8 for the second peer evaluation were 3.82, 3.66, 3.33, and 3.56 respectively. The mean scores were 3.84, 3.68, 3.32, and 3.66 for the high proficiency class and 3.80, 3.64, 3.34, and 3.50 for the low proficiency class. The Cronbach alpha coefficient scale for each class was .07 and .58 respectively. There were no large correlations among the four variables. Therefore, the aggregated mean was used in the following analyses for consistency with other means.

5.1 Quantitative results

The means for self- and peer assessment of the first and the second oral presentations are displayed in Table 1.[5] A glance at the means suggests that the high proficiency group made modest gains in all variables, and so did the low proficiency group except for content for self-evaluation, in which the means for the first and the second presentations were the same ($M = 3.14$).

Table 1: Results of Self- and Peer Evaluation for the Two Presentations

	Self				Peer			
	First		Second		First		Second	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
High proficiency group (*n* = 12)								
voice control	2.65	0.60	3.06	0.57	3.66	0.21	3.77	0.12
body language	2.79	0.65	3.04	0.68	3.58	0.18	3.63	0.13
content	2.89	0.69	3.31	0.87	3.85	0.13	3.93	0.09
effectiveness	3.03	0.63	3.31	0.76	3.86	0.09	3.88	0.05
Low proficiency group (*n* = 17)								
voice control	2.66	0.62	2.91	0.54	3.38	0.23	3.58	0.22
body language	2.63	0.47	2.85	0.52	3.39	0.13	3.57	0.12
content	3.14	0.62	3.14	0.47	3.53	0.13	3.73	0.08
effectiveness	2.98	0.56	3.20	0.51	3.46	0.08	3.78	0.08

A three-way MANOVA, with Group as a between-participants factor and Time and Rater as within-participant factors, revealed that there was a significant main effect of Time, $F(4,24) = 4.49$, $p = .008$, Pillai's trace = .43, *partial eta*^{2} = .43, and a significant main effect of Rater, $F(4, 24) = 22.67$, $p < .001$, Pillai's trace = .79, *partial eta*^{2} = .79. This indicates that there was a statistically significant increase in scores and that the high level and the low level classes scored higher on peer evaluation than self-evaluation. There was also a significant interaction effect between Time and Rater for voice control, $F(1, 27) = 4.34$, $p = .047$, *partial eta*^{2} = .14. Simple main effects analysis showed that for self-evaluation, there was a significant improvement in voice control for the high proficiency level class ($p < .05$), and for peer evaluation a significant improvement was observed in voice control for the low proficiency level class ($p < .001$).

A significant interaction effect was observed among Time, Rater, and Group for content, $F(1,27) = 5.13$, $p = .032$, *partial eta*^{2} = .16. Simple main effects analysis showed that for peer evaluation, there was a significant difference in content between the high level and the low level classes ($p < .001$), for the high level class between self- and peer evaluation for the first presentation ($p < .001$) and the second presentation ($p = .024$), and for the low level class

---

[5] While the analysis of the mean scores for self- and peer assessment of the third oral presentation was not included here, the results are shown in Appendix B. The results indicate that the high proficiency group scored higher in both self- and peer evaluation, whereas the low proficiency group scored lower in some variables for the third presentation, which will not be discussed further in this paper.

between self- and peer evaluation for the first presentation ($p = .016$) and the second presentation ($p < .001$). Any other effects were not significant, Fs(4,24) < 2.18, $ps > .102$.

The main effect of Time was further analyzed by univariate ANOVAs, reaching statistical significance for voice control, $F(1, 27) = 19.45$, $p < .001$, *partial eta*^{2} = .42, body language, $F(1, 27) = 5.32$, $p = .029$, *partial eta*^{2} = .17, content, $F(1, 27) = 6.15$, $p = .020$, *partial eta*^{2} = .19, and effectiveness, $F(1, 27) = 6.94$, $p = .014$, *partial eta*^{2} = .21. Likewise, the main effect of Rater was further analyzed, reaching statistical significance at $p < .001$ for voice control, $F(1,27) = 78.92$, *partial eta*^{2} = .75, body language, $F(1,27) = 82.18$, *partial eta*^{2} = .75, content, $F(1,27) = 39.04$, *partial eta*^{2} = .59, and effectiveness, $F(1,27) = 52.83$, *partial eta*^{2} = .66.
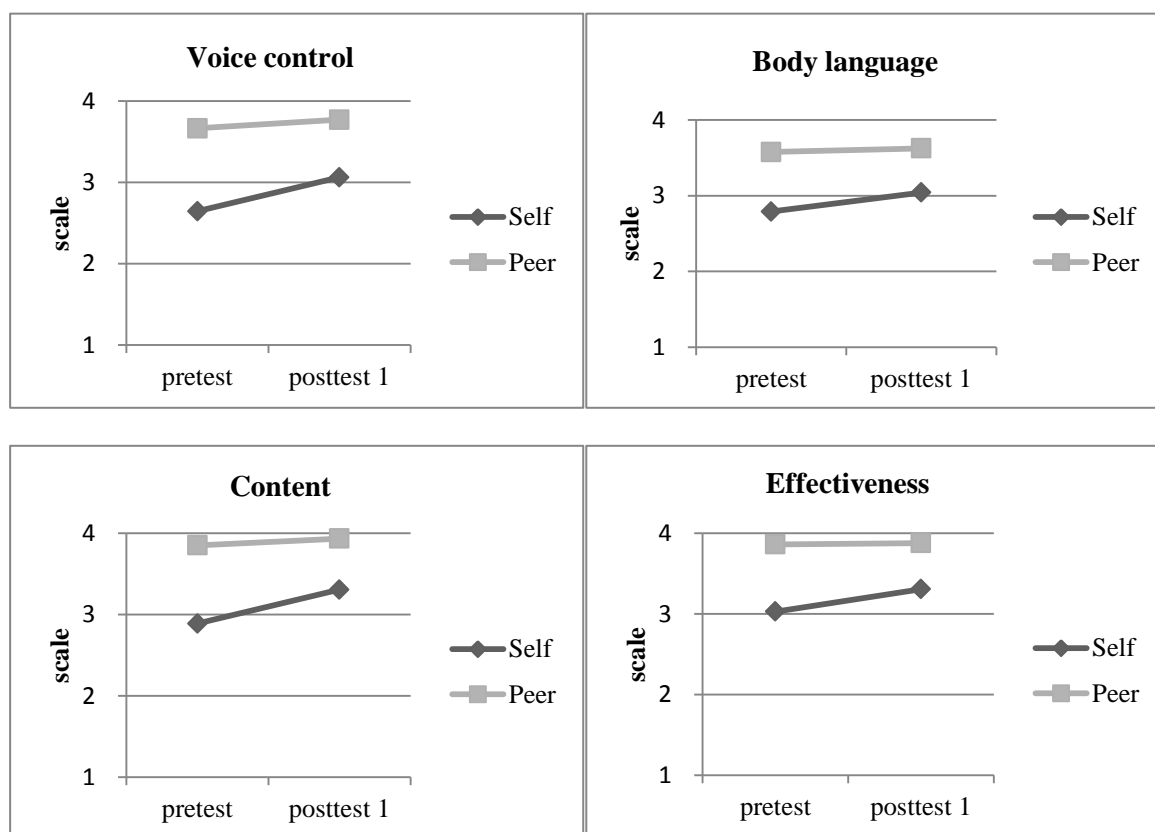

Fig. 2. Changes of the means in high proficiency group

5.1.1 Results of the high proficiency group

Figure 2 demonstrates the changes in mean scores of the four variables for the high proficiency group: voice control, body language, content, and effectiveness. It is shown that peer evaluation was scored higher than self-evaluation in the first and the second presentations in all variables. As comparing the mean scores at the different points of time, the most improved score was seen in content (0.42 point increase) for self-evaluation and voice control (0.11 point increase) for peer evaluation. These results suggest that watching the sample performance video may have contributed to giving a higher score on voice control to peers and on content to their own presentation.

5.1.2 Results of the low proficiency group

The changes in mean scores in the low proficiency group are shown in Figure 3. As observed in the high proficiency group, all the mean scores of peer evaluation were higher than those of self-evaluation in the low proficiency group. The most increased variable was voice control (0.25 point increase) for self-evaluation and effectiveness (0.32 point increase) for peer evaluation. Though almost all mean scores increased in their presentations, the mean scores remained the same in the content area for self-evaluation. These results suggest that it may be possible for the low proficiency group to improve all the variables for peer evaluation and the three variables other than content for self-evaluation of their oral presentation without watching the sample performance videos; the low proficiency students might have scored higher because of having practiced speaking and listening of the language in class, as those of the high proficiency group did, and therefore, it is necessary to consider what the students had learned between the two presentations.
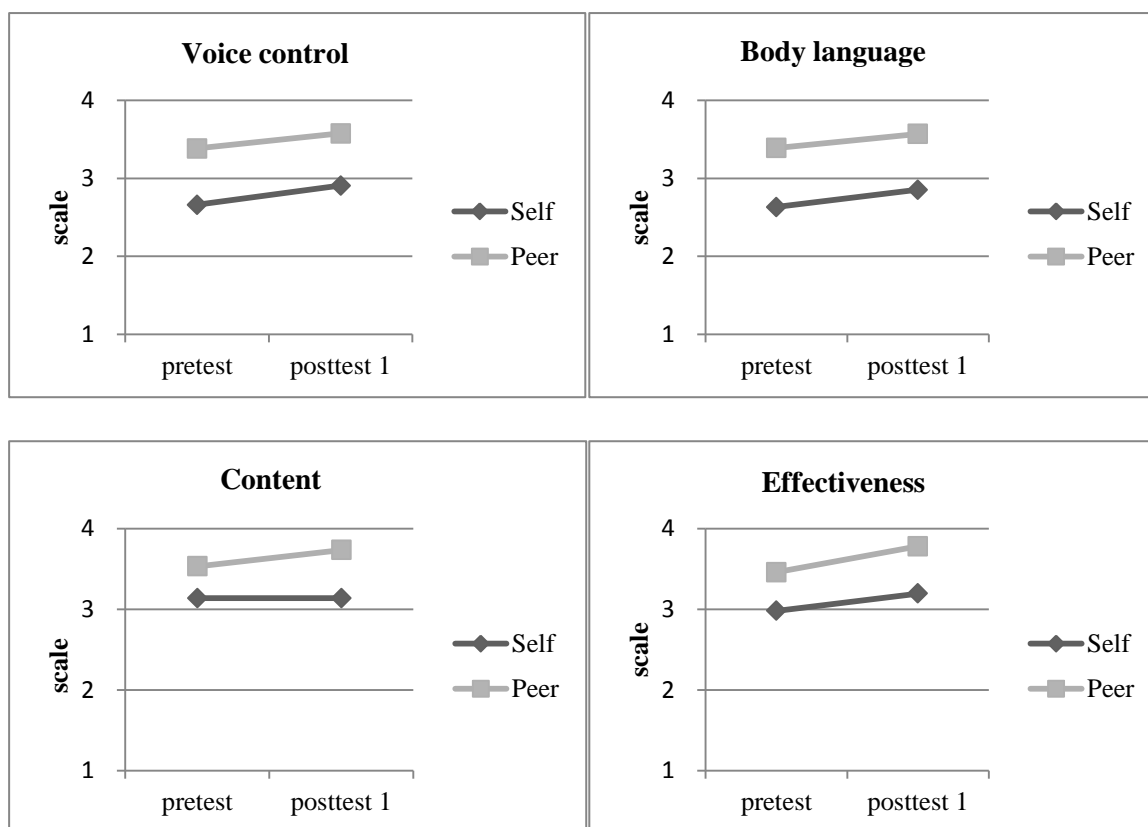

Fig. 3. Changes of the means in low proficiency group

5.2 Results of text mining: Sample video reviews

This section shows the results of text mining of students' responses to sample video reviews, in which the two questions were addressed as follows: (Q1) What are the speakers' strengths? and (Q2) What are the speakers' weaknesses? To examine the text data, the following three analysis methods were employed: word frequency analysis, contrastive analysis, and sentiment analysis. Frequency word analysis allows us to determine the frequency of words in the text. Contrastive analysis is used to identify words that occurred distinctively between the high proficiency and the low proficiency classes. Sentiment analysis determines whether a piece of text is positive or negative by examining the words that appear together.

Table 2: Top-10 High Frequency Words in Sample Video Reviews (*N* = 29)

	Word	Meaning	Part of speech	Frequency
1	yoi	*good*	adjective	77
2	hatsuon	*pronunciation*	noun	54
3	hanasu	*speak*	verb	36
4	kikitoru	*catch the meaning*	verb	23
5	koe	*voice*	noun	22
6	te	*hand*	noun	20
7	ookii	*big*	adjective	19
8	naiyou	*content*	noun	19
9	kiku	*listen*	verb	18
10	miru	*watch/look*	verb	6

5.2.1 Frequency word analysis

Table 2 shows the list of words that were frequently used in the sample video reviews. Responses to the two questions were added up and analyzed together instead of being analyzed for the two sections separately. The result shows that the adjective *good* appeared the most frequently in the text. This may be because students watched only good samples so they wrote positive comments more than negative ones on the samples. In addition, various words that are relevant to audio information occurred in the video reviews such as *pronunciation*, *speak*, *catch the meaning*, *voice*, *big*, and *listen*.

5.2.2 Contrastive analysis between the high and the low level classes

According to the results of contrastive analysis, the top-five words used particularly by the high English proficiency group were as follows: *good* (56/77),[6] *memorization* (12/14), *content* (14/19), *pronunciation* (31/54), and *place* (9/10). Some of these words appeared in the following contexts:

---

[6] In this case, 56 means the frequency with which the word appeared in one of the groups and 77 indicates the total appearance in both groups, and so forth.

- Speaking pace was moderate. I think the hand position was good. (Q1)

- It was easily understood because it was memorized well. (Q1)

- I can't understand the speech because of the pause after each word. (Q2)

For the low English proficiency group, the high frequency words used distinctively were: *pace* (13/15), *appropriateness* (9/9), *understand* (10/12), *speak* (22/36), and *vocabulary* (10/13). Following are examples of comments with these words:

- With simple English and appropriate pacing, I could understand the content easily. (Q1)

- The speaker stammered or mispronounced the words several times. (Q2)

5.2.3 Sentiment analysis

Further analysis revealed that the top–three nouns that were accompanied by the words with positive meanings were *pronunciation* (35), *pace* (12), and *voice* (11). On the contrary, the ones that appeared with the words containing negative implications were *pronunciation* (6), followed by *gesture* (2), and *posture* (2). The words that occurred in this analysis were used in the following contexts:

- The speech was articulated clearly enough to understand. (Q1)

- Since the body was moving, the posture should have been straighter. (Q2)

5.3 Results of text mining: Student feedback

This section shows the results of analysis of students' feedback about sample performance observation using text mining. In the student feedback, students were instructed to answer the following two questions: (Q1) What do you think about observing the sample performance videos? and (Q2) Were there any effects of the sample videos on your own presentations?

5.3.1 Frequency word analysis

Student feedback about observing the sample video performances was examined through text mining, adding up all responses to the first and the second questions. Table 3 shows the high frequency words that occurred in the student feedback. The result indicates that nouns were the most frequently used part of speech among the top-10 words.

Table 3: Top-10 High Frequency Words in Student Feedback (*N* = 29)

	Word	Meaning	Part of speech	Frequency
1	supiichi	*speech*	noun	40
2	miru	*look*	verb	34
3	hito	*person*	noun	34
4	jibun	*self*	noun	25
5	yoi	*good*	adjective	25
6	hatsuon	*pronunciation*	noun	11
7	kawaru	*change*	verb	11
8	ki	*attention*	noun	10
9	joozu	*skillfulness*	noun	9
10	umai	*good at*	adjective	9

5.3.2 Contrastive analysis between the high and the low level classes

Contrastive analysis was also performed to examine the words that appeared distinctively in the two classes. For the high proficiency class, the words that appeared distinctively were: *look* (21/34), *self* (16/25), *attention* (9/10), *good* (15/25), and *pay* (*attention*) (5/5). For the low proficiency class, the top-five words used by students were *change* (11/11), eight of which were accompanied with negations, *person* (22/34), *English* (7/7), *exist* (5/5), and *pronunciation* (8/11).

5.3.3 Sentiment analysis

A further analysis of students' feedback revealed that the words that co-occurred with the words carrying positive connotations were: *person* (13), *pronunciation* (5), and *speech* (4); meanwhile, the words that emerged with those having negative connotations were *person* (3), *English* (1), and *manner of speaking* (1).

5.4 Results of content analysis

In addition to analyzing the open-ended question forms with the text mining method, the student feedback form was further examined in terms of the overall responses. For Q1, it was revealed that almost all students in both groups agreed that the sample performance observation would be useful and helpful:

- I think that whatever you do, it is important to imitate those who do well, and that there is a lot to learn from observation. Because I can find something I should pay attention to, I think it is good to watch sample performances. (high proficiency learner)

- It was really helpful. Since I wanted to understand the difference in features between the good speaker and "not-good" speaker, it was very beneficial. (high proficiency learner)

- All the speakers on the video looked vividly and with confidence, thereby they did not get upset, improvised their speech, and kept smiling when making mistakes. Furthermore,

there were a variety of personalities and ways of speaking. It was good to watch the video since there are some that I can make use of. (low proficiency learner)

For Q2, students answered whether they could take advantage of social modeling when performing their own presentations. For the high proficiency group, 10 out of 19 students agreed that the sample video observation helped improve their language and oral presentation skills, whereas two students stated it did not make any difference. Here are examples of positive comments:

- A day before the presentation, we discussed that we should work hard to pronounce English well and memorize the speech. I also practiced my speech a lot right after observing the samples. I think that the presentation after the sample video observation was the best.

- I came to pay more attention to the hand position and posture. Though I did not care about these aspects of the presentation when watching my classmates' performances, I tended to focus on the speakers' manners and posture on the video since they were not my friends. I became aware that other students felt the same way as I did.

For the low proficiency group, only four out of 17 students mentioned the sample videos were effective to develop their performance. However, nine expressed difficulties in imitating the speakers on the sample videos, as shown below:

- As a result, those who are good at English improved their oral presentations and those who are not made less progress.

- Students who do not like English would think that it is great, but they would also be discouraged by noticing the gap between the sample and their own performances.

- Whether we watch sample performance videos or not, it does still make a difference. However, I didn't think my presentation improved greatly. It would be meaningless if speakers show no desire to improve themselves.

6 Discussion

The present study used self- and peer evaluation of EFL learners' oral presentations and two open-ended question forms to demonstrate how effectively observational learning works in developing learners' language skills as well as oral presentation skills, comparing high and low English proficiency classes.

6.1 Quantitative analysis

The quantitative data showed that the high and low proficiency classes gained increased scores in self- and peer evaluations, and that peer evaluation scores were higher than those of self-evaluation, regardless of when the sample performance videos were observed. For Time, since learners in both groups had observed their own videos of the first presentation, they might have already become aware of the weaknesses of their performance and looked for

ways to improve them for the second presentation. Therefore, it is suggested that the scores of the second presentation were higher than those of the first one.

As for Rater, the result that peer evaluation scored higher than self-evaluation seems consistent with the previous studies (Falchikov, 1995; Guo, 2013; Peng, 2009), where the notion of "friendship effects" in peer assessment was pointed out. According to Falchikov (2005), students are aware of the problem of giving higher marks to their peers, knowing why they do so. Since students are not confident of their own objectivity, their judgment tends to be lenient to those with whom they have good relations (Cheng & Warren, 1997). In this study, the students had been taking several courses together and were well acquainted with each other at the time of the study, which suggests that they may have been friendly enough to give a higher score to their peers, regardless of the effects of sample performance observation. Additionally, as for peer evaluation, the means on content yielded high scores on the first presentation ($M = 3.85$, $SD = 0.13$) and second presentation ($M = 3.93$, $SD = 0.09$) in the high proficiency group. It is more likely that ceiling effects occurred in the content area for the high proficiency group, probably because friendship effects might have worked in the first presentation, resulting in participants giving a higher score to those whom they were already familiar with, and consequently the mean score became higher.

It is further important to note that a significant interaction effect on content was observed among the three independent variables, Time, Rater, and Group. With regard to self-evaluation, the mean score progressed moderately from the first presentation ($M = 2.89$, $SD = 0.69$) to the second presentation ($M = 3.31$, $SD = 0.87$) for the high proficiency group, while there was no gain on content from the first presentation ($M = 3.14$, $SD = 0.62$) to the second presentation ($M = 3.14$, $SD = 0.47$) for the low proficiency group. It is assumed that observing the sample performance could enhance high proficiency learners' awareness of such content as introduction, body, and conclusion of the presentation, and as a result, the mean score on the content area was higher in the high proficiency group.

The topic of the second presentation and learners' listening skills might have also facilitated their understanding of the presentation content for the high proficiency group. First of all, the second presentation topic concerned what learners wanted to recommend to peers. Since they talked about what they were already familiar with and might have used the vocabulary that was relevant to the topic, it is more likely that they had understood the content well, delivering their own presentation with confidence. Second, the learners' English proficiency level might have mattered. Students were divided into the classes according to the scores of the TOEIC Bridge test, which measures learners' listening and reading skills. Because of this, it is more likely that learners in the high proficiency group already had stronger listening skills and thus could understand the content of the speech better than the low proficiency group. It is plausible to conclude that speaker and listener factors may have affected the results concerning content of the oral presentation.

6.2 Text mining and content analyses

The results of text mining of the sample video reviews showed that observational learning with videos was effective for students to enhance awareness of the essentials of language skills as well as presentation skills. In particular, observational learning with videos enabled students to pay careful attention to the audio information on the sample performances by

using words such as *pronunciation*, *speak*, *catch the meaning*, *voice*, and *listen*, according to the results of word frequency analysis in Table 2. Though students were able to obtain both visual and audio information through the videos, they seemed to pay less attention to the visual aspects of the sample performances. This may be because the study was conducted in EFL classes where the focus of learning was speaking and listening skills; therefore students' interest seemed to lie in what and how they spoke the language rather than how they looked.

Additionally, the high and low English proficiency classes experienced the sample performance observation differently. With respect to the content analysis, the present study indicates different effects of the sample performance videos between the high and the low level English proficiency classes. Namely, the high level class in this study tended to raise students' awareness of the essentials of effective presentation skills by observational learning since the sample performances were observed as a new behavior and modeled as a positive outcome, as Bandura (1986, 2005) suggests in social modeling theory. Conversely, it is likely that students of the low level class were discouraged by observing good samples that made them feel intimidated since there was a huge gap in the English proficiency of the sample performers and the low level learners. Accordingly, the same sample performance videos yielded different effects on the high and the low English proficiency classes.

6.3 Theoretical implications

This study was an adjusted replication of the Grez et al. (2014) study, in which effects of observational learning were examined in the Belgian university context. The result of our study was consistent with that of their study in terms of the improvement in the content area of the oral presentations; however, there were several differences between our study and theirs. First, the participants in their study made their presentations in their first language, whereas our participants made theirs in the foreign language. Second, the proficiency of oral presentation skills for the experimental and control groups were equivalent in their study, but our study compared two classes of learners whose English proficiencies were not equivalent. Nevertheless, the differences in the effects on observational learning between the two classes were clearly demonstrated. For further research, it is necessary to compare groups of learners who have the same level of English proficiency.

6.4 Educational implications

One purpose of using videos in educational contexts is to show them to learners and enable them to imitate them as a positive consequence. Our result indicates that the sample performances in this study reinforced learners' demotivational attitudes toward modeling the sample videos. This may be because the sample performances were so good that the low level learners might have felt it impossible to imitate them, even though they had observed them. As Shrosbree (2008) points out, it is necessary for teachers to select a negative model performance that "students can analyze and use to improve their awareness of common pitfalls and errors when communicating in a second language" (p. 76). A sample video that would represent an attainable performance can also be shown to them. Therefore, it is possible even for the low level students to improve their language and oral presentation skills through social modeling with videos so long as their English proficiency is considered.

6.5 Limitations

As a limitation of this study, we should point out that, first of all, the number of participants was small. If there had been more participants for each class, the study might have clearly demonstrated effects of observational learning with sample videos, aside from on the content area. Second, it would be necessary to reduce the number of evaluation items. In Grez et al.'s (2014) study, there were nine criteria in the evaluation form they employed, whereas our study used 14 items along with overall comments. It might depend on students' English proficiency level, but it seemed that the students in this study, particularly low level students, had difficulty focusing on various aspects of the oral presentations in a short period of time. Fewer items might give students more time to judge each item, and the reliability and validity of the evaluation instruments would be enhanced accordingly.

7 Conclusions

This study attempted to shed light on effects of observational learning using sample videos on EFL learners' development of language and presentation skills, comparing two classes representing two different English proficiencies in the Japanese university context. The present study found there was a significant gain in terms of the content of the speech by watching the videos from the previous students' work. Comparison of effects of sample videos between the high and the low proficiency classes suggests that students with different proficiency levels need different sample performance videos to learn the language and oral presentation skills. This study will hopefully take one step towards social modeling by Japanese EFL learners and offer a fuller comprehension of the use of video recordings as samples.

References

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (2005). The evolution of social cognitive theory. In K. Smith & M. Hitt (Eds.), *Great minds in management* (pp. 9–35). Oxford: Oxford University Press.

Cheng, W., & Warren, M. (1999). Peer and teacher assessment of the oral and written tasks of a group project. *Assessment and Evaluation in Higher Education, 22*(2)*,* 233–239.

Falchikov, N. (1995). Peer feedback marking: Developing peer assessment. *Innovations in Education and Training International, 32*(2), 175–187.

Falchikov, N. (2005). *Improving assessment through student involvement.* New York: Roughtledge/Falmer.

Grez, L. D., Valcke, M., & Roozen, I. (2014). The differential impact of observational learning and practice-based learning on the development of oral presentation skills in higher education. *Higher Education Research and Development, 33,* 256–271.

Guo, R.X. (2013). The use of video recordings as an effective tool to improve presentation skills. *Polyglossia, 24,* 92–101.

Hung, H. (2009). Learners' perceived value of video as mediation in foreign language learning. *Journal of Educational Multimedia and Hypermedia, 18*(2), 171–190.

McNulty, A., & Lazarevic, B. (2012). Best practices in using video technology to promote second language acquisition. *Teaching English with Technology, 12*(3), 49-61. Retrieved from http://www.tewtjournal.org/VOL%2012/ISSUE3/ARTICLE%204.pdf

Okada, Y. (2011). Reflecting on students' videotaped presentations. *Journal of Teaching English, 20,* 41–54.

Okada, Y. (2012). Video feedback comparison with EFL students. *Studies in Language Education Seisen University, 4,* 17–33.

Okada, Y. (2013, August). EFL learners' positive perceptions with regard to reviewing their classroom performance. Paper presented at *18th Pan-Pacific Association of Applied Linguistics Conference*, Suwon, South Korea.

Okada, Y., & Ito, T. (in press). The effects of using video recording in peer and self-assessment. *Research Bulletin Liberal Arts Edited by Nihon University College of Economics, 76.*

Peng, J.F. (2009). *Peer assessment of oral presentation in an EFL context.* (Unpublished doctoral dissertation). Indiana University. (UMI No. 3380148)

Shrosbree, M. (2008). Digital video in the language classroom. *The JALT CALL Journal, 4*(1), 75–84.

Smith, K., & Hitt, M. (Eds.) (2005). *Great minds in management.* Oxford: Oxford University Press.

Tugrul, T.O. (2012). Student perceptions of an educational technology tool: Video recordings of project presentations. *Procedia – Social and Behavioral Sciences, 64,* 133-–40.

Yamashiro, D.A., & Johnson, J. (1997). Public speaking in EFL: Elements for course design. *The Language Teacher, 21*(4), 13–17. Retrieved from http://jalt-publications.org/tlt/articles/2124-public-speaking-efl-elements-course-design

Appendix A: Evaluation Form

		Rating (1: strongly agree、4 : strongly disagree)				Description
1	**Projection**	1	2	3	4	Spoke loud enough for the audience.
2	**Pace**	1	2	3	4	Spoke at a good rate.
3	**Intonation**	1	2	3	4	Put appropriate stress and pausing.
4	**Diction**	1	2	3	4	Spoke clearly.（Did not mumble; Did not use inappropriate stress.）
5	**Posture**	1	2	3	4	Stood straight.
6	**Foot & Hand Positions**	1	2	3	4	Placed the foot shoulder-width apart and set the hands together, keeping around waist high.
7	**Eye Contact**	1	2	3	4	Looked at the audience.
8	**Facial Expression**	1	2	3	4	Showed a relaxed facial expression.
9	**Introduction**	1	2	3	4	Included greeting, who the speaker was, and what s/he would talk about.
10	**Body**	1	2	3	4	Conveyed a clear explanation.
11	**Conclusion**	1	2	3	4	Restated and summarized the key points.
12	**Topic Choice**	1	2	3	4	Selected an interesting topic.
13	**Language Use**	1	2	3	4	Used simple sentence structures.
14	**Vocabulary**	1	2	3	4	Used easy vocabulary words.

15. Please comment on the overall performance.

●

●

●

●

●

Appendix B: Results of Self- and Peer Evaluation for the Third Presentation

	Self		Peer	
	M	*SD*	*M*	*SD*
High proficiency group (*n* = 12)				
voice control	3.38	0.38	3.81	0.10
body language	3.35	0.53	3.71	0.14
content	3.64	0.46	3.97	0.05
effectiveness	3.56	0.52	3.93	0.07
Low proficiency group (*n* = 17)				
voice control	2.74	0.76	3.58	0.21
body language	2.88	0.62	3.67	0.10
content	2.93	0.71	3.74	0.08
effectiveness	3.00	0.83	3.75	0.10