



# Effects of Observing Model Video Presentations on Japanese EFL Learners' Oral Performance

**Yasuko Okada**

([okada92025@gmail.com](mailto:okada92025@gmail.com))

Seisen University, Japan

**Takafumi Sawaumi**

([t.sawaumi@gmail.com](mailto:t.sawaumi@gmail.com))

Ryutsu Keizai University, Japan

**Takehiko Ito**

([take@wako.ac.jp](mailto:take@wako.ac.jp))

Wako University, Japan

---

## Abstract

As observing model videos can develop learners' speaking (e.g. pronunciation and word/sentence stress) and oral presentation skills, this study explores its effects on Japanese university students' performance with respect, in particular, to model videos of presentations. Two types of model videos were shown to 27 participants in two classes: in one class, more-proficient model videos were shown prior to less-proficient model videos; in the other, the order was reversed. In both cases, the videos were shown prior to the students' second and third presentations. To examine any observational learning effects, the first, second, and third presentation scores obtained through self- and peer evaluation were analyzed quantitatively, while student responses to open-ended questions were analyzed qualitatively using text mining. The results indicate that there was no significant effect of any factor on the self-evaluation scores. However, peer ratings show that the third presentation was rated significantly higher than the first for the class in which students watched the less-proficient model videos before the more-proficient. The findings indicate that the observation sequence of the model videos may affect the development of learners' performance, although the use of both more-proficient and less-proficient model videos in any order can improve learners' cognitive, linguistic, and presentation skills.

---

## 1 Introduction

Since the late 1990s, the Japanese business community has been emphasizing the importance of developing communication skills in order to develop human resources and contribute to the globalized society (Ikeda & Erikawa, 2016). A study by Koike, Takada, Matsui, Terauchi, and IIBC (2010) showed that 85.1% of more than 7,000 Japanese business people feel that, in addition to sufficient English proficiency, presentation skills are also necessary to negotiate in the international context. With view to the 2020 Tokyo Olympic and Paralympic Games, the Ministry of Education, Culture, Sports, Science and Technology (MEXT) has issued the English Education Reform Plan (2014), with the help of which students at elementary, junior, and high schools are expected to improve their

English communication skills. Therefore, it is essential for Japanese students to acquire English presentation skills while at the university and before becoming members of society.

In English as a foreign language (EFL) education, a number of studies have shown that learners can increase their oral proficiency in the target language by delivering presentations (e.g. Adams, 2004; Hincks, 2010; Kibler, Salerno, & Palacios, 2014). To improve their presentation skills effectively, students can have their presentations video-recorded, ascertain their strengths and weaknesses while watching the videos, and then develop these skills in subsequent presentations. Observing other non-native speakers' videos can provide useful models and motivate students to give presentations, as well as engage in self-reflection on their own videos (Okada, Sawaumi, & Ito, 2014). One advantage of using video-recorded students' performances as models is that, if students' presentations are video-recorded with their permission, they can be shown to other students, at different points in time, by different teachers, and in different places. Although model videos may not be identical to live models, students can learn presentation techniques from video-recorded student presentations in classroom settings.

The authors investigated the effects of non-native model videos on EFL learners' presentation skills in the Japanese university context. Okada, Sawaumi and Ito (2014) tested model presentations that were considered to represent more-proficient speaker performances with two groups of students whose English proficiency levels were high and low, respectively. An analysis of open-ended responses revealed that more-proficient model observation was effective for the high-proficiency group but intimidated the low-proficiency group. The authors suggested that students with different English proficiency levels needed different models customized to their language aptitude. Thus, in addition to more-proficient model videos, students were shown less-proficient model videos prior to their own presentations. The aim of this mixed methods study is to examine the effects of using two different types of non-native model videos on the development of oral presentation skills in the Japanese EFL context.

## **2 Literature on oral presentations**

Oral presentations are a significant part of college students' academic experiences. Yet, instructional studies on oral presentation skills are fragmented, and a theory of oral presentation instruction has not been established (De Grez, Valcke, & Roosen, 2009). However, the following three key topics are necessary to discuss oral presentation instruction. First, performance expectations are explicitly stipulated with students model good presentations (De Grez et al., 2009; De Grez, Valcke, & Roosen, 2014; Okada, Sawaumi, & Ito 2014). In observational learning theory, people observe others and acquire new human behaviors through modeling (Bandura, 1971, 1977). If people observe an appropriate model, they will consider it in a positive light, which will thus trigger modeling behaviors. However, if people observe an inappropriate model, they may notice its negative effects and, accordingly, will not imitate the behavior. Thus, people can identify positive and negative patterns based on the rewards or punishments they receive through any imitated behaviors. In foreign language learning, language proficiency necessarily varies among students, and good models do not always bring about a positive effect. Thus, either more-proficient or less-proficient models could be either appropriate or inappropriate depending on the learners' proficiency levels. For example, Okada, Sawaumi and Ito (2014) used more-proficient non-native speaker presentation videos to instruct high- and low-proficiency groups. The results indicated that more-proficient model video observation was effective for the high-proficiency group but intimidated the low-proficiency group, suggesting that students with different English proficiency levels need different models that are customized to their language aptitude.

Second, it is essential that students' presentations are evaluated according to presentation criteria (e.g. De Grez et al., 2009, 2014; Yamashiro & Johnson, 1997), and that the criteria focus on content and delivery. De Grez et al. (2009, 2014) used quality of introduction, structure, and conclusion as content criteria, and eye contact, vocal delivery, enthusiasm, interaction with the audience, and body

language as criteria for the nature of delivery. On the other hand, emphasizing students' acquisition of language function, skills, and cross-cultural awareness, the elements identified by Yamashiro and Johnson consist of (a) projection, pace, intonation, and diction as elements of voice control; (b) posture, eye contact, and gesture as elements of body language; (c) introduction, body, and conclusion as elements of oral presentation content; and (d) topic, language use, vocabulary, and purpose as elements of effectiveness. To avoid subjective evaluation, it is necessary to set clear evaluation criteria elements when teachers or peers evaluate students' presentations.

Finally, students need multiple opportunities to deliver presentations. They should be able to receive feedback from teachers and peers, and based on this feedback, refine their skills in subsequent presentations (e.g. Castañeda & Rodríguez-González, 2011; Okada, Sawaumi & Ito, 2014). Specifically, in studies in which students have given multiple presentations in the target language over a semester, their presentations were video-recorded and viewed repeatedly to allow them to reflect on their performance. Doing so benefited not only the students but also the teachers, since students' strengths and weaknesses were explicitly shown, providing teachers with the opportunity to consider what oral presentation instructions students really needed in order to deliver successful presentations.

### **3 Methods**

#### ***3.1 Purpose and design***

Based on previous studies, we aimed to explore the effects of model videos on learners' own presentations, and, more specifically, non-native speaker models. In Okada, Sawaumi and Ito's (2014) study, there was a significant improvement in students' presentation content after they were shown more-proficient model video presentations, although the two groups of learners had different impressions of the videos; it is therefore important to use different model videos depending on learners' language proficiency to develop their linguistic and presentation skills. Both more-proficient and less-proficient video presentations may help learners develop these skills and techniques to some extent. Viewing either of the two model qualities could improve learners' cognitive skills; students' awareness of common language errors and pitfalls will be enhanced during communication in the target language (Shrosbree, 2008).

The authors conducted the current study as part of a university's regular English lessons. As in previous studies, it had the following limitations: (a) the number of participants could not be controlled because those students who were enrolled simply became the participants; (b) random assignment was not possible because students had already been assigned according to their placement test scores (i.e. English proficiency); and (c) since the study was part of their regular lessons, it was necessary to complete the task within the allotted class time, although these video observations are uncommon in the Japanese EFL context and therefore require more extensive instruction (e.g. Okada & Ito, 2014; Okada & Sawaumi, 2015).

#### ***3.2 Participants***

The study involved 29 Japanese student participants who had enrolled in two different English communication classes taught by the first author. The classes were held during the spring 2015 semester in a university within the urban area of Tokyo. All participants were freshman students majoring in economics, who had studied EFL for at least six years during junior and high school. At the beginning of the course, the students were asked to complete a consent form, allowing the video and text data to be used for the purposes of research or education; they were further informed that their participation in the research would not affect their grade, nor would the study require any additional work or time. Two students did not consent to their data being used for research. Their data were thus omitted from the analyses, resulting in a final total of 27 participants.

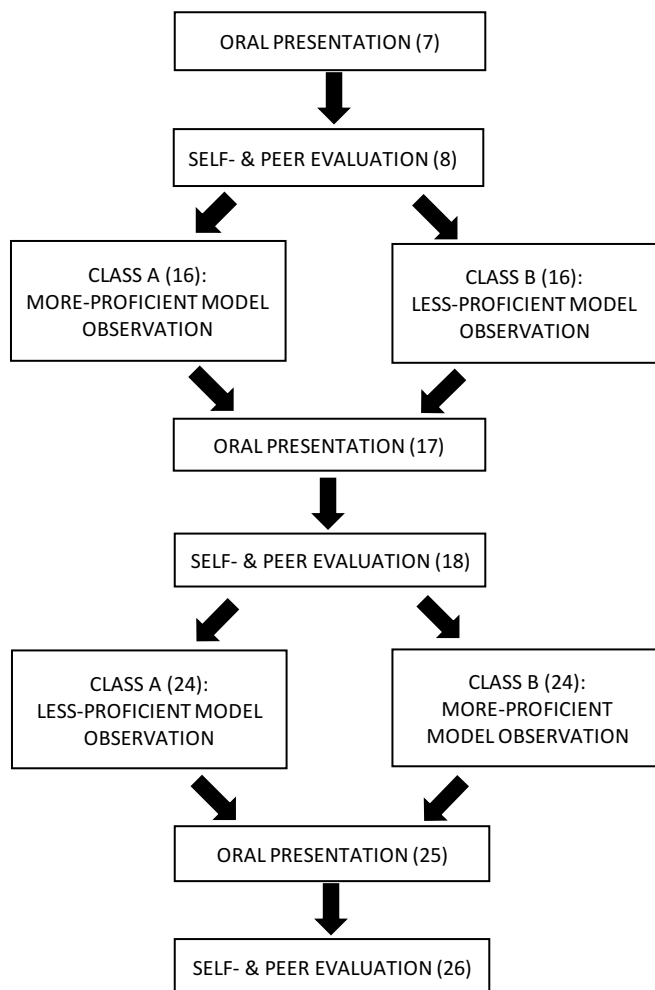
### 3.3 *Settings*

At the beginning of the semester, the participants were assigned to each course based on their scores on the Test of English for International Communication (TOEIC) Bridge test. As the students' choice of a third language (e.g. Chinese, German, or Spanish) was also a consideration for their class placement, the participants were not randomly assigned in the study. The average scores were 130 for Class A ( $n = 12$ ) and 123 for Class B ( $n = 15$ ). There was no statistically significant difference in the test scores between the two classes ( $t(25) = 1.57, n.s., d = 0.62$ ). These classes were, therefore, considered equivalent in terms of English language proficiency. With respect to class size, there was no significant difference between the two classes.

The course was mandatory for graduation and emphasized developing oral communication skills through group and pair work. The 90-minute class met twice a week during the 14-week semester. Both participating classes covered the same content and used the same textbook for language instruction. After the instructor had explained the nature of the study, the students clearly understood that their performance in the classroom would be used in this study (although only 45% of the course workload was used as data: 30% for oral presentations and 15% for draft submissions and self- and peer evaluation; the other 55% consisted of 40% for vocabulary quizzes and 15% for reading assignments).

### 3.4 *Teaching procedure*

Drawing upon the findings from Okada, Sawaumi and Ito (2014), students were encouraged to give three memorized oral presentations.<sup>1</sup> Figure 1 presents the teaching procedure of this study. For each presentation, students were instructed to write a three-paragraph draft consisting of introduction, body, and conclusion within 180 to 200 words. Prior to each presentation, students' drafts were submitted to the instructor at least once in order to receive feedback on their content and language.



Note: The numbers in the parentheses indicate the number of session in which the data were collected

**Fig. 1. Overview of teaching procedure**

### 3.4.1 Oral presentation phase

In Sessions 7, 17, and 25, students' oral performances were recorded using a video camera on a tripod, and graded by the instructor. After the second and third presentations, students were asked to write about how video observation had affected their own performance. In Session 25, the students gave the third oral presentation, after which they completed a performance reflection form.

### 3.4.2 Self- and peer evaluation phase

In Sessions 7, 17, and 25, students' oral performances were recorded using a video camera on a tripod, and graded by the instructor. After the second and third presentations, students were asked to write about how video observation had affected their own performance. In Session 25, the students gave the third oral presentation, after which they completed a performance reflection form.

### **3.4.3 Model video observation phase**

In Sessions 16 and 24, prior to the second and third presentations, respectively, students in both classes watched model video presentations and commented on what they had observed in the videos. In Session 16, Class A watched the more-proficient model videos first, while Class B watched the less-proficient presentations first. Then, in Session 24, each class viewed the videos in the alternative order of quality. After viewing and commenting on each clip, students were divided into groups of three to four to discuss the model videos for approximately 2 minutes, after which they were expected to write down both the strengths and weaknesses of the model speakers.

The model video clips shown to the students, comprising eight more-proficient speaker video clips and eight less-proficient speaker video clips, were selected by the first author from former students' video-recorded performances. More-proficient models were selected from learners in higher-level classes. In choosing more-proficient models, presenters' excellent use of word stress and pausing was considered, in addition to good pronunciation, as non-native speakers find these elements particularly difficult to acquire (Okada, 2011, 2012). The less-proficient model selections comprised videos from lower-level classes and students could explicitly point to inadequate elements in them. Each video clip lasted for approximately 2 to 3 minutes and the presentations were on a slightly different topic from that of the current study.

## **3.5 Research materials**

In this study, the following instruments were used: (a) self- and peer evaluation forms for the three oral presentations; (b) a student performance reflection form; and (c) a video observation reflection form. These forms were explained and elaborated upon by students in Japanese, and the students' comments were translated into English.

### **3.5.1 Self- and peer evaluations**

The evaluation items used for self- and peer evaluation in the study were based on those of previous studies (Okada, Sawaumi, & Ito 2014; Yamashiro & Johnson, 1997). Although Okada, Sawaumi and Ito (2014) used 14 items with a comment column, we referred to former students' suggestions and decided not to include three content items in order to reduce the burden of filling out the form. Thus, a total of 11 questions were rated on a 4-point scale from 1 "weak" to 4 "great." Items 1 to 4 concerned elements of voice control such as projection, pace, intonation, and diction; Items 5 to 8 dealt with elements of body language such as posture, foot and hand positions, eye contact, and facial expression; and Items 9 to 11 covered elements of effectiveness such as topic choice, language use, and vocabulary. A comment column was integrated into the evaluation form (see Appendix 1).<sup>2</sup>

### **3.5.2 Student performance reflection**

After the second and third oral presentations, students were asked to freely write down their thoughts on whether the model video observation had affected their oral presentations. As both classes had watched more-proficient and less-proficient model videos prior to the students' oral presentations, the form was used to explore how the two types of model video had differently affected their oral presentations.

### **3.5.3 Video observation reflection**

This form was completed by students after the third self- and peer evaluations. It included the following two sections: (a) students were asked to rate the perceived difficulty of each topic while preparing a script on a 5-point scale (1 "easy" to 5 "difficult"),<sup>3</sup> and (b) they were asked to write

down comments about their observations of the two model videos.

## 4 Results

### 4.1 Quantitative data analyses

For quantitative analyses, repeated measures' analyses of variance<sup>4</sup> (ANOVA) were performed to examine the effects of class (Class A vs. Class B, i.e. the between-participants factor) and time of the presentation (first vs. second vs. third, i.e. the within-participant factor) and their interaction on self- and peer-rated presentation scores.

To simplify the analyses, 11 items on the evaluation form were aggregated and used as the following subscales: voice control (Items 1 to 4), body language (Items 5 to 8), and effectiveness (Items 9 to 11). The three subscale types were the dependent variables. The Cronbach's alpha coefficients were .83, .90, .94 (first); .92, .90, .98 (second); and .88, .85, .92 (third), respectively, for the self-evaluated subscales. For the peer-evaluated subscales, they were .82, .66, .73 (first); .84, .56, .81 (second); and .82, .59, .89 (third), respectively.<sup>5</sup>The quantitative analyses were performed using IBM SPSS 22.0, and an alpha level of  $p < .05$  was set for the statistical significance in all inferential statistics. In the following analyses, the Greenhouse-Geisser correction was used where the sphericity assumption was not met for the within-participant factor.

#### 4.1.1 Self-evaluation

The mean scores of self-evaluated voice control, body language, and effectiveness are presented in Table 1. First, a two-way ANOVA (class  $\times$  time) on the mean score of self-evaluated voice control did not reveal any significant main effects,  $F(1, 18) = 0.03, p = .86, \eta_p^2 < .01$  for class;  $F(2, 36) = 1.11, p = .34, \eta_p^2 = .06$  for time. The class  $\times$  time interaction effect was not significant either,  $F(2, 36) = 3.14, p = .06, \eta_p^2 = .15$ .

Second, an identical two-way ANOVA was conducted on the mean score of self-evaluated body language but did not reveal any significant main effects,  $F(1, 18) = 0.79, p = .39, \eta_p^2 = .04$  for class;  $F(2, 36) = 1.46, p = .25, \eta_p^2 = .08$  for time. The class  $\times$  time interaction effect was not significant either,  $F(2, 36) = 1.91, p = .16, \eta_p^2 = .10$ .

Finally, an identical two-way ANOVA was conducted on the mean score of self-evaluated effectiveness but did not reveal any significant main effects,  $F(1, 18) = 0.47, p = .50, \eta_p^2 = .03$  for class;  $F(1.52, 27.43) = 0.99, p = .36, \eta_p^2 = .05$  for time. The class  $\times$  time interaction effect was not significant either,  $F(1.52, 27.43) = 2.29, p = .13, \eta_p^2 = .11$ .

**Table 1. Mean scores and SDs of self-evaluation**

	Time 1			Time 2			Time 3		
	n	M	SD	n	M	SD	n	M	SD
Class A									
Voice	12	2.48	0.62	10	2.90	0.57	11	2.66	0.46
Body	12	2.54	0.77	10	2.73	0.66	11	2.84	0.53
Effect	12	2.28	0.83	10	2.97	0.60	11	2.76	0.63
Class B									
Voice	15	2.50	0.65	15	2.68	0.82	11	2.89	0.84
Body	15	2.27	0.84	15	2.77	0.77	11	2.68	0.79
Effect	15	2.84	0.82	15	2.80	1.00	11	3.21	0.85

Note: The data of students who were absent at the time of evaluation were eliminated at each stage of analysis. Voice: Voice control; Body: body language; Effect: effectiveness.

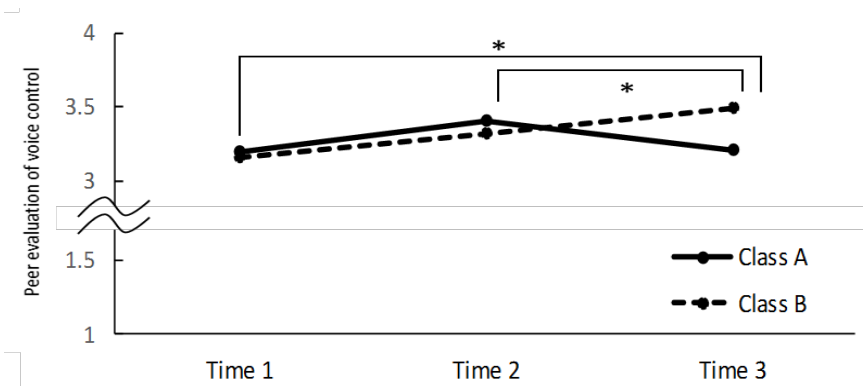
4.1.2 Peer evaluation

Next, the mean scores of peer-evaluated voice control, body language, and effectiveness are presented in Table 2. First, an identical two-way ANOVA was conducted on the mean score of peer-evaluated voice control. The ANOVA revealed a significant main effect of time,  $F(1.49, 29.84) = 3.90, p = .04, \eta_p^2 = .16$ , but the main effect of class was not significant,  $F(1, 20) = 0.22, p = .64, \eta_p^2 = .01$ . The class  $\times$  time interaction effect was significant,  $F(1.49, 29.84) = 5.74, p = .01, \eta_p^2 = .22$ . Post-hoc tests for the significant interaction effect showed that Class B's third peer evaluation of voice control was significantly higher than that of the first and second presentations. This observation was not seen for Class A (see Fig. 2).

Table 2. Mean scores and SDs of peer evaluation

	Time 1			Time 2			Time 3		
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>n</i>	<i>M</i>	<i>SD</i>
Class A									
Voice	12	3.20	0.25	10	3.41	0.25	11	3.21	0.24
Body	12	3.19	0.18	10	3.37	0.24	11	3.24	0.22
Effect	12	3.31	0.21	10	3.45	0.17	11	3.42	0.13
Class B									
Voice	13	3.16	0.27	15	3.32	0.30	15	3.50	0.21
Body	13	3.00	0.20	15	3.19	0.17	15	3.30	0.19
Effect	13	3.45	0.06	15	3.51	0.08	15	3.65	0.08

Note: The data of students who were absent at the evaluation were eliminated at each stage of analysis. Voice: Voice control; Body: body language; Effect: effectiveness.

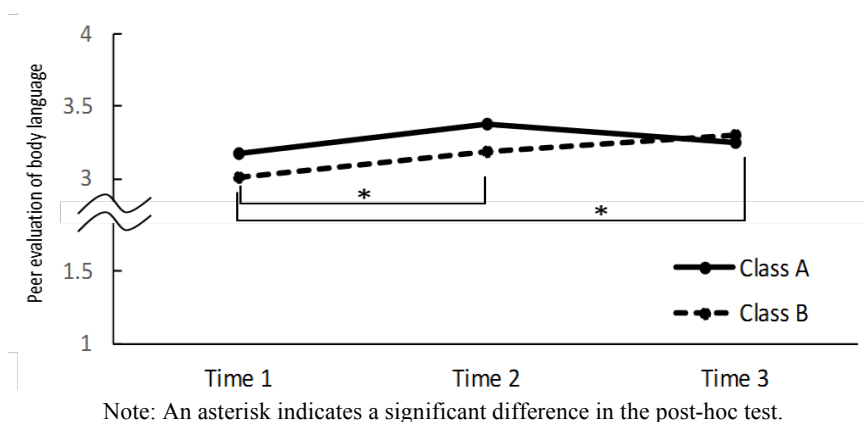


Note: Asterisks indicate significant differences in the post-hoc test.

Fig. 2. Peer evaluation of voice control as a function of time and class

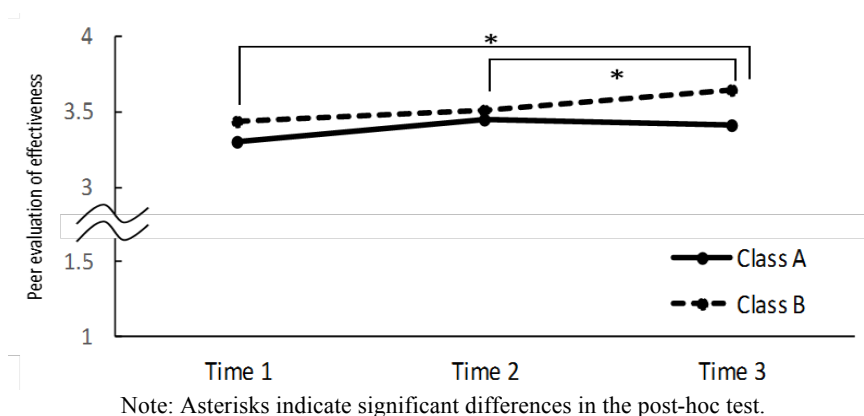
Second, an identical two-way ANOVA was conducted on the mean score of peer-evaluated body language. The ANOVA revealed a significant main effect of time,  $F(2, 40) = 6.86, p < .01, \eta_p^2 = .26$ , but the main effect of class was not significant,  $F(1, 20) = 3.77, p = .07, \eta_p^2 = .16$ . The class  $\times$  time interaction effect was significant,  $F(2, 40) = 4.16, p = .02, \eta_p^2 = .17$ . Post-hoc tests for the significant interaction effect showed that the second and third peer evaluations of body language were significantly higher than that of the first presentation only in Class B (see Fig. 3).





**Fig. 3. Peer evaluation of body language as a function of time and class**

Finally, an identical two-way ANOVA was conducted on the mean score of peer-evaluated effectiveness. The ANOVA revealed significant main effects of class,  $F(1, 20) = 16.83, p < .01, \eta_p^2 = .46$ , and time,  $F(1.51, 30.24) = 5.68, p = .01, \eta_p^2 = .22$ . These effects were qualified by a significant class  $\times$  time interaction effect,  $F(1.51, 30.24) = 5.47, p = .02, \eta_p^2 = .22$ . Post-hoc tests for the significant interaction effect showed that the third peer evaluation of effectiveness was significantly higher than that of the first and second presentations in Class B, which was not the case in Class A (see Fig. 4).



**Fig. 4. Peer evaluation of effectiveness as a function of time and class**

#### 4.2 Text mining and qualitative analyses

To qualitatively examine students' responses to the first and second performance reflections and the video observation reflection, we performed a text mining analysis, using Text Mining Studio 5.1 by NTT DATA Mathematical Systems Inc. As the Japanese language, unlike English, is an agglutinative language, in which words are made up of a sequence of meaningful word elements, called morphemes, it is necessary to analyze morphemes in the process of text mining analysis. Other examples of agglutinative languages include Korean and Turkish. The software used for this study enabled us to first analyze morphemes and then count the morphemes in the text data. In the following section, the results of the word frequency analysis will be shown, followed by the qualitative analysis of the text data.

#### 4.2.1 Analysis of student performance reflection

Table 3 shows the results of the word frequency analysis on students' performance reflection, which was administered immediately after the second and third presentations. The results indicate that students seemed to become conscious of what they had observed in the model videos when giving their oral presentations, as the word *consciousness* (48)<sup>6</sup> was most frequently used in the reflections. Additionally, it is clear that students in both classes referred to both visual and auditory information, for example, *speech* (36), *look* (32), *posture* (20), *speak* (19), and *pronunciation* (17) when reflecting on their own performance. The frequent appearance of *consciousness* indicates that students successfully focused on various aspects of presentation skills by watching the two types of model video.

**Table 3. High frequency words in student performance reflection**

	Word	Meaning	Part of Speech	Total
1	意識	consciousness	noun	48
2	スピーチ	speech	noun	36
3	見る	look	verb	32
4	前回	previous time	noun	23
5	人	person	noun	21
6	ビデオ	video	noun	20
7	姿勢	posture	noun	20
8	話す	speak	verb	19
9	良い	good	adjective	18
10	発音	pronunciation	noun	17

According to students in both classes, watching more-proficient model videos was an effective way to focus on particular aspects of the model speakers in order to imitate them, whereas less-proficient model video observations helped students identify ways to improve the model presentations, regardless of which model video students watched first. For instance, after the second presentation, nine students in Class A stated that they had focused on improving facial expressions, pronunciation, and posture because these aspects seemed important in the more-proficient model videos. In Class B, 12 students indicated that they had worked on areas where the model speakers had failed in the less-proficient speaker videos.

- I learned from the model video to present with a smile. (S76, more-proficient model first)
- I practiced my presentation while looking at myself in the mirror. I gave the presentation as if I had been thinking in Japanese, concentrating on what I really wanted to say. (S7, more-proficient model first)
- From watching the less-proficient model video presentations, I learned that posture and eye contact were also important factors that made the presentations better. Therefore, I practiced my presentation paying attention to these aspects in addition to my volume. (S102, less-proficient model first)

For the third presentation, the contents of students' comments were reversed as the students watched the model videos in reverse order. The students in Class A reported that they had found some negative aspects of the model videos (e.g. eye contact and posture) and that they would not deliver presentations as the model that the speakers had.

- Since some model speakers didn't make eye contact, I attempted to look at the audience in my presentation. Then, I thought my voice would carry well without any effort. (S1, more-proficient model first).

On the other hand, the students in Class B who watched more-proficient models wrote that they practiced their presentations focusing on positive features (e.g. eye contact and pronunciation) in the more-proficient model videos, and that they could successfully change their postures despite having difficulties with pronunciation.

- Because we had watched model speakers who were good at pronouncing English and stressing words in the previous lesson, I wanted to improve my pronunciation in the presentation. In addition, I thought I should make use of making good eye contact as the model speakers did. However, I became nervous when I stood in front of students, and consequently I couldn't pronounce well and felt uneasy. Although I calmed down and had good posture in the third presentation, I became frustrated because I couldn't practice what I had originally planned. (S116, less-proficient model first).

The students' comments in both classes showed that observation of the more-proficient speaker video increased their motivation to rehearse their presentation by imitating the speakers' strengths. Meanwhile, the less-proficient model video observation enabled students to identify negative presentation patterns, so that they could avoid negative effects when giving their presentations.

**Table 4. High frequency words in model video observation reflection**

	Word	Meaning	Part of Speech	Total
1	見る	look	verb	23
2	人	person	noun	21
3	良い	good	adjective	17
4	スピーチ	speech	noun	16
5	自分	self	noun	16
6	参考	reference	noun	12
7	ビデオ	video	noun	10
8	分かる	understand	verb	9
9	発表	presentation	noun	8
10	うまい	good	adjective	6

Note: In this table, different Japanese words indicating "good" appeared twice. The first word, ranked third, was used to explain people's behaviors or personalities that exceed the standard, whereas the second instance, appearing in the 10<sup>th</sup> place, is used to describe excellent technique or skills.

#### 4.2.2 Analysis of video observation reflection

Table 4 shows the results of the word frequency analysis on the video observation reflections, which the students completed after concluding all of the procedures. The results indicated that words relevant to model video observation, such as *speech* (16), *video* (10), and *presentation* (8) frequently appeared in the text. In addition, *good* (23) was also frequently mentioned in comments. Using the high frequency words in the table, it is suggested that *watching a video presentation by a person who is good at delivering a speech would be a good reference*. Observing video models may therefore have been a positive experience for the learners, who were able to learn from both model video types.

Some students in each class agreed that watching model speakers who were unknown to them was effective, as they could observe these videos objectively. In addition, both more-proficient and less-proficient video observation provided students with opportunities to identify ways of improving their own presentations. The following are examples from students' reflections on the model video observation:

- Since there was something I wanted to imitate in the more-proficient model videos, I focused on this when practicing my presentation. The first model video presentations were very effective.... For the second model videos, I could observe what I needed to improve and practice in my presentation. (S9, more-proficient model first)

- It was very good because I was able to find ways to improve my skills through the (less-proficient) model videos. In addition, I found a difference in my own presentations before and after watching the model videos. (S103, less-proficient model first)

Comments on the video observation reflection showed that students were able to analyze the model behaviors and find their own methods of improving both their language skills and presentation techniques.

## 5 Discussion and implications

This study focused on two different model video qualities that affected learners' language and presentation skills in line with observational learning theory (Bandura, 1971, 1977). After discussing the quantitative and qualitative results, the research questions will be revisited.

The results indicated that the self-evaluation scores of students' oral presentations did not change dramatically in either class, while the peer evaluation scores generally improved over the semester. This may be because young Japanese adults tend to evaluate themselves lower than other nationalities (Oshio, Okada, Mogaki, Namikawa, & Wakita, 2014; Prime Minister of Japan & His Cabinet, 2016). Further, Japanese people value social identity more than personal identity (Endo, 1997), which may suggest that participants in our study scored higher in peer evaluation and lower in self-evaluation. However, the improvement was significant only for Class B students, who observed the less-proficient model videos first and more-proficient videos next. This suggests that the sequence of video observation may affect the quality of students' oral performance, such that observing less-proficient model videos prior to more-proficient videos is more effective for improving students' oral performance. Both classes were taught oral communication skills throughout the course, for which reason it is reasonable to expect that their language and presentation skills would improve through regular lessons alone. Hence, it is noteworthy that, according to peer evaluation, Class A students' presentations did not improve greatly despite the regular lessons and video observation, whereas Class B students improved their oral presentations significantly after watching less-proficient model videos followed by more-proficient model videos.

The text mining results suggest that video observation raised students' awareness of audio and visual aspects of their performance. The findings indicated that both more-proficient and less-proficient model videos enabled learners to watch the models carefully and apply their observations to their own performance. Students' comments also support the importance of the video viewing sequence: those who watched less-proficient model videos first were more satisfied with their own performance than those who watched the videos the other way around.

The findings demonstrated that students imitated skills from more-proficient videos but were also able to use the less-proficient models to make meaningful adjustments. In other words, observing different types of model video can enhance learners' awareness of how to imitate the models' strengths and how to improve on any observed weaknesses. In this study, although the two types of model presentation were shown to each class at two different times, it is possible that showing both models at the same time or using native-speaking models helps learners imitate pronunciation and presentation skills in addition to improving their cognitive awareness at the early stage of oral presentation instruction.

The findings suggest that more-proficient model presentations can lead to positive effects on learners' performance. From the social comparison perspective (Festinger, 1954), people evaluate their own abilities by comparing themselves with others, which is what the learners did with the model videos. More specifically, when people compare themselves to others who perform better than them, a unidirectional push upward to gain an advantage over others occurs (Dumas, Hugust, Monteil, Rastoul, & Nezelek, 2005; Takata, 2011; Wheeler, 1966), resulting in upward comparison. In this study, as a result of learners' upward comparison to the more-proficient model videos, they may have achieved an attainable goal resulting in improvement of their language and presentation skills.

Due to the quasi-experimental design, this study did not have a large sample size. As mentioned previously, quasi-experimental design involves several limitations. As the course in the study was a compulsory university course, not all students would always be willing to participate in the study. Moreover, all tasks had to be completed during class time, and therefore, it was not possible to collect data from students who missed the oral presentations. To generalize the results of the study, we need to conduct a replication study using the same methods but different subjects. Future studies may further benefit from viewing both more-proficient and less-proficient models simultaneously (rather than sequentially). Students could watch videos featuring native speaker models, which may affect their oral presentations in a different manner.

Further, this study did not compare the scores from students with those from the instructor because it was considered that these scores were unlikely to correspond. For example, Falchikov and Boud (1989) pointed out a lower correspondence between student self- and instructor evaluations, if the course was for beginning-level learners and compulsory. In our study, the course was compulsory, and students' TOEIC Bridge test scores were not very high. Moreover, Kobayashi (2010) argues that East Asian learners are characterized as having low self-esteem, because they understand their own weaknesses and underestimate their own performances; the study participants would thus likely have fixated on their errors and underestimated their subsequent performances. Further study may be able to focus on comparing scores between students and instructor.

Finally, it is necessary to consider other factors, such as self- and peer evaluation of video-recorded presentations, which potentially affect students' oral presentations, in addition to model video observation. Previous studies have shown that it is effective for students to self-evaluate their own performances while watching video-recordings, in order to improve their language abilities (e.g. Castañeda & Rodríguez-González, 2011; Okada, 2011). Because students in this study had an opportunity to watch their video-recorded presentations after each presentation, it is likely that self-evaluation using these videos impacted their subsequent presentations. Students had also watched other students' live presentations prior to peer-evaluation with the videos. Thus, viewing both live and video-recorded presentations probably affected the peer evaluation scores, as well as students' subsequent performances.

Despite these limitations, both the quantitative and qualitative data collected in this study revealed interesting differences in the use of the two model video types in developing students' language skills and presentation techniques. Not only in the EFL context but also in other foreign languages, the use of non-native speakers' performance videos as models could be applicable to increasing students' language proficiency, while developing their presentation skills in the target language.

## 6 Conclusion

The study findings have shown that both more-proficient and less-proficient model videos can help increase students' awareness of speakers' strengths and weaknesses as well as their ability to observe videos from different perspectives. The study presented empirical evidence of the significant effect of model video observation, suggesting that the pedagogic use of model videos can develop Japanese EFL students' metacognitive skills in the target language. In particular, we examined the effects of two model video types on Japanese EFL learners' oral performance development. Although this study uncovered the importance of observing and comparing video models, it is necessary to focus on developing learners' language skills through observation and comparison between their own performance and the model. Although the study was conducted in Japan, these research findings would also be beneficial for readers outside Japan. While supporting the development of their use of stress, rhythm, and intonation, with which Japanese students have difficulties (Ohata, 2004), future research needs to examine how learners' language proficiency can be enhanced for work in a global society where English presentation and negotiation skills are required.

## Notes

<sup>1</sup> The topics were “My favorite food,” “My memories of high school club activity,” and “My ideal trip.” Regarding the second presentation topic, some students asked the instructor to change the topic because they were not involved with club activities. Therefore, they were allowed to write about memories of their family instead of high school.

<sup>2</sup> Analysis of written comments on Item 12 was not included in the paper because it focuses on performance and video observation reflections.

<sup>3</sup> The results showed that the means were 2.91, 2.96, and 2.91 in chronological order, located around the midpoint (3). On the whole, the three topics were considered equally difficult for students.

<sup>4</sup> Non-parametric tests, such as a Friedman test, were also conducted and the results were compared with those of the ANOVA as indicated in Appendix 2. Because this study necessarily used a multiple comparison analysis test to compare the three presentations, it was not possible to use Wilcoxon Ranking. In psychology, the Likert scale is considered an interval scale, and it is common for such scales to be analyzed using parametric tests, such as ANOVA. In addition, the statistical power of parametric tests such as ANOVA is higher than that of non-parametric tests. The robustness of ANOVA can be justified even though its distribution may not be normal.

<sup>5</sup> Although the Cronbach’s alpha coefficients of body language for the three presentations were relatively low, mean scores across Items 5 to 8 were used in the subsequent analyses in the interest of consistency.

<sup>6</sup> In this case, 48 indicates the total appearance in both classes, and so forth.

<sup>7</sup> S stands for student.

## Acknowledgement

This work was supported by the JSPS Grant-in-Aid for Scientific Research (C) 15K02730 and Research Institute for Language Education at Seisen University, Japan. We would like to thank Editage ([www.editage.jp](http://www.editage.jp)) for English language editing. This is a revised and expanded version of a paper presented at the Globalization and Localization in Computer-Assisted Language Learning Conference at Pai Chai University, Daejeon, Korea, 12–14 November 2015.

## References

- Adams, K. (2004). Modelling success: Enhancing international postgraduate research students’ self-efficacy for research seminar presentations. *Higher Education Research and Development*, 23(2), 115–130. doi:10.1080/0729436042000206618
- Bandura, A. (1971). Analysis of modeling processes. In A. Bandura (Ed.), *Psychological modeling: Conflicting theories* (pp. 1–62). Chicago, IL: Aldine-Atherton Press.
- Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Castañeda, M., & Rodríguez-González, E. (2011). L2 speaking self-ability perceptions through multiple video speech drafts. *Hispania*, 94(3), 483–501. Retrieved from <http://www.jstor.org/stable/23032122>
- De Grez, L., Valcke, M., & Roozen, I. (2009). The impact of an innovative instructional intervention on the acquisition of oral presentation skills in higher education. *Computers & Education*, 53, 112–120. doi:10.1007/BF03174762
- De Grez, L., 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. doi:10.1080/07294360.2013.832155
- Dumas, F., Hugust, P., Monteil, J. M., Rastoul, C., & Nezlek, J. B. (2005). Social comparison in the classroom: Is there a tendency to compare upward in elementary school? *Current Research in Social Psychology*, 10, 166–187.
- Endo, Y. (1997). Relationship-enhancement and relative self-effacement. *The Japanese Journal of Psychology*, 68(5), 387–395.
- Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research*, 59(4), 395–430.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7(2), 117–140. doi:10.1177/001872675400700202
- Hincks, R. (2010). Speaking rate and information content in English lingua franca oral presentations. *English for Specific Purposes*, 29(1), 4–18.

- Ikeda, K., & Erikawa, H. (2016). Global jinzai ikusei wo mezasu eigo kyoiku seisaku no hensen to mondaiten [A critical study of the English language education policies for promotion of global human resource development]. *Bulletin of the Faculty of Education, Wakayama University Humanities*, 66, 37–43. Retrieved from <http://repository.center.wakayama-u.ac.jp/metadata/2739>
- Kibler, A., Salerno, A., & Palacios, N. (2014). “But before I go to my next step”: A longitudinal study of adolescent English language learners' transitions in oral presentations. *TESOL Quarterly*, 48(2), 222–251.
- Kobayashi, Y. (2010). Communication kyoiku to eigo communication kyoiku [Communication Education and English Communication Education]. *Artes Liberales*, 86, 95–105. Retrieved from <http://hdl.handle.net/10140/3778>
- Koike, I., Takada, T., Matsui, J., Terauchi, H., & The Institute for International Business Communication (IIBC) (2010). *Kigyo ga motomeru eigoryoku* [English Skills: What do companies really need?]. Tokyo, Japan: Asahi Press.
- Ministry of Education, Culture, Sports, Science and Technology (MEXT) (2014). *Report on the future improvement and enhancement of English education (outline): Five recommendations on the English Education Reform Plan responding to the rapid globalization*. Retrieved from <http://www.mext.go.jp/en/news/topics/detail/1372625.htm>
- Ohata, K. (2004). Phonological differences between Japanese and English: Several potentially problematic areas of pronunciation for Japanese ESL/EFL learners. *The Asian EFL Journal*, 6(4), 1–19. Retrieved from <http://asian-efl-journal.com/journal-2004/>
- Okada, Y., & Sawaumi, T. (2015). Factors affecting student evaluation of video-recorded performance. *Nihon University College of Economics Research Bulletin Liberal Arts*, 77, 267–279.
- Okada, Y., & Ito, T. (2014). Jiko hyoka peer hyoka kara mita gakushu-sha no video eizo katsuyo no koka [The effects of using video recording in self- and peer assessment]. *Nihon University College of Economics Research Bulletin Liberal Arts*, 76, 47–55.
- Okada, Y. (2011). Reflecting on students' videotaped presentations. *Journal of Teaching English*, 20, 41–54.
- Okada, Y. (2012). Using video feedback comparison with EFL students. *Studies in Language Education Seisen University*, 4, 17–33.
- Okada, Y., Sawaumi, T., & Ito, T. (2014). Different effects of sample performance observation between high and low proficiency English learners. In Aishah M. K., W. M. Chan, S. W. Chi, K. N. Chin, S. Klayklung, S. Kumar Bhatt, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker, (Comps.), *Knowledge, Skills and Competencies in Foreign Language Education. Proceedings of the 6<sup>th</sup> CLS International Conference* (pp. 394–413). Singapore: National University of Singapore, Centre for Language Studies. Retrieved from <http://www.fas.nus.edu.sg/cls/CLaSIC/clasic2014/Proceedings/>
- Oshio, A., Okada, R., Mogaki, M., Namikawa, T., & Wakita, T. (2014). Age and survey-year effects on self-esteem in Japan: A cross-temporal meta-analysis of scores on Rosenberg's self-esteem scale. *Japanese Journal of Educational Psychology*, 62(4), 273–282.
- Prime Minister of Japan and his cabinet. (2016). *Nihon no kodomotachi no jiko koteikan ga hikui genjo ni tsuite* [The present situation of young people's low self-esteem in Japan]. Retrieved from <http://www.kantei.go.jp/jp/singi/kyouikusaisei/chousakai/dai1/siryou4.pdf>
- Shrosbree, M. (2008). Digital video in the language classroom. *The JALT CALL Journal*, 4(1), 75–84. Retrieved from [http://journal.jaltcall.org/articles/4\\_1\\_Shrosbree.pdf](http://journal.jaltcall.org/articles/4_1_Shrosbree.pdf)
- Takata, T. (2011). *Tasha to kuraberu jibun: Shakaiteki hikaku no shinri gaku* [Self who compares others: Psychology in social comparison]. Tokyo: Saiensusha Co., Ltd.
- Wheeler, L. (1966). Motivation as a determinant of upward comparison. *Journal of Experimental Social Psychology*, 1, Suppl. 1, 27–31.
- 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>

## Appendices

### Appendix 1

#### Evaluation Form

		Rating (1: weak, 4: great)				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	Used 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	Feet & Hand Positions	1	2	3	4	Placed feet shoulder-width apart and kept hands together around waist height.
7	Eye Contact	1	2	3	4	Looked at the audience.
8	Facial Expression	1	2	3	4	Had a relaxed facial expression.
9	Topic Choice	1	2	3	4	Selected an interesting topic.
10	Language Use	1	2	3	4	Used simple sentence structures.
11	Vocabulary	1	2	3	4	Used easy vocabulary words.
12	Please comment on the overall performance.					

Note: The evaluation form used by students was written in Japanese.

### Appendix 2

#### Results of Friedman test and ANOVA

		Friedman test		ANOVA	
		Class A	Class B	Class A	Class B
Self	Voice		*		
	Body		*		
	Effect				
Peer	Voice				*
	Body	*	*		*
	Effect		*		*

\*  $p < .05$

Note: For the results of peer evaluated scores of Class B using ANOVA, voice control was significantly higher in the third than in the first and second presentations; body language was significantly higher in the second and third presentations than in the first; effectiveness was significantly higher in the third than in the first and second presentations.