

How Do Speech Model Proficiency and Viewing Order Affect Japanese EFL Learners' Speaking Performances?

Yasuko Okada (okada92025@gmail.com)

Seisen University, Japan

Takafumi Sawaumi (t.sawaumi@gmail.com)

Ryutsu Keizai University, Japan

Takehiko Ito (take@wako.ac.jp)

Wako University, Japan

Abstract

Learners in foreign language classrooms can benefit from viewing model videos featuring other non-native speakers, which helps to develop their public speaking performance. The present study attempts to replicate Okada, Sawaumi, and Ito (2017), who demonstrated how the viewing order of different model speakers affected learners' public speaking performance. The participants were Japanese freshmen ($N = 24$) enrolled in English communication courses. In the classroom-oriented study, one group was shown more- and less-proficient speaker models prior to delivering the second and third presentations, respectively; the other group was shown the same models in the opposite order. To determine whether the models impacted learners in any meaningful way, self- and peer evaluations were obtained from participants following each speaking performance and analyzed in conjunction with responses to reflection papers. Results from two-way ANOVAs indicated that video-based observational learning in the viewing order of less-proficient speakers first and more-proficient ones next resulted in improved peer evaluation in subsequent performances, which was consistent with the findings of Okada et al. (2017). Responses to reflection papers revealed that both more and less-proficient speaker models helped enhance learners' awareness of self/others and improve their own public speaking skills in the EFL classroom.

Keywords: model video, viewing order, public speaking, replication, observational learning

Introduction

An increasing number of Japanese learners of English as a foreign language (EFL) are showing interest in enhancing their English communication skills. According to the national Ministry of Education, Culture, Sports, Science and Technology (MEXT, 2014), Japanese students are

expected to improve their communication skills in English to respond to the rapidly globalizing society. Nevertheless, Japanese people are often considered to be modest and shy compared to those in Western countries; these traits may prevent them from mastering a foreign language and speaking in front of others with confidence. Although many universities in the United States are offering public speaking courses to develop these skills and reduce students' anxiety of public speaking (Hancock, Stone, Brundage, & Ziegler, 2010), few Japanese universities offer such courses. Accordingly, it is important for students in Japan to improve their public speaking skills when studying at a university or college before becoming full-fledged members of society.

To improve public speaking skills in classrooms, video-recording students can be effective and helpful, because this audio-visual information allows them to reflect on their own learning and develop foreign language skills (Castañeda & Rodríguez-González, 2011; Okada, Sawaumi, & Ito, 2014; Shrosbree, 2008). This is increasingly so in recent years, as developments in social media have made students more and more familiar with recording photos or videos and showing them to others. Thus, it may be worthwhile to conduct a study on the integration of video-based observational learning with teaching public speaking skills in an EFL context.

Review of Literature

Overview of Okada, Sawaumi, and Ito (2014, 2017)

A study by Okada et al. (2014) examined the effects of viewing more-proficient speaker model videos on learners' speaking performances in Japanese EFL classrooms. Findings indicated that low-proficiency learners ($n = 17$) were intimidated by viewing these models, whereas model observation helped high-proficiency learners ($n = 12$) enhance their motivation for delivering speeches. This research seems to support the use of videos as practical models, in line with the theory of observational learning (Bandura, 1971, 1977).

Based on the original study (Okada et al., 2014), Okada et al. (2017) conducted a replication study to find out whether learners ($N = 27$) can benefit from viewing not only more-proficient speaker models but also less-proficient ones. Their study reported that both more-proficient and less-proficient models help enhance learners' awareness of their strengths and weaknesses of speaking skills, and that students observed the models for different purposes and for different reasons. Further, analyses showed that the group that observed less-proficient models first and more-proficient ones second scored significantly higher in peer evaluation than the group that

observed the models in the reverse order. It was thus speculated that the viewing order sequence affects student performance in the target language.

Video-Based Observational Learning

According to observational learning theory, new behaviors are acquired by watching and imitating models (Bandura, 1971, 1977). Social learning of this nature has typically involved observing real-life models; however, recent studies (e.g., De Grez, Valcke, & Roozen, 2014; Okada et al., 2014; 2017) indicate that videos can be equally effective as educational models. In the observational learning process, individuals undergo four stages when acquiring new behavioral patterns: *attention*, *retention*, *reproduction* and *motivation* (Bandura, 1977). During the attention phase, one focuses on a model's behavior so as to understand it; then, in retention, one memorizes the patterns of the model in order to facilitate symbolic mediation, which assists in understanding complex behavioral patterns during the preliminary stages of language skill acquisition. Symbolized images are subsequently transformed into actual behaviors, in the reproduction phase; unlike behaviors, language skills cannot be learned through observation alone, and reproduction is therefore necessary once an overview of the model's behaviors has been obtained. During the motivation stage, a distinction appears between acquisition and performance, wherein observed patterns are not necessarily performed unless individuals anticipate that doing so will produce positive results. In this way, observational learning assists students in recognizing behavioral patterns and developing cognitive skills that ultimately affect performance.

De Grez et al. (2014) examined the effect of observational learning on the development of presentation skills among 38 Dutch-speaking Belgian university students. Participants were divided into two groups: one group watched video clips followed by practice activities, while the other group did the reverse. Learners in the former group improved strongly after observing the clips; however, members of the latter group improved just slightly following the practice activities, but significantly after watching the clips.

Effects of Instructional Sequence in Foreign Language Teaching

Researchers have focused on the effects of methodological order sequence to help students develop language skills, especially reading and listening. Shimizu (2007) compared higher- and lower-proficiency learners across two types of question order: inference (local) questions and thematic (global) questions. Her results showed that lower-proficiency learners scored higher in answering local-questions first and global ones next, while higher-proficiency learners scored higher in the reverse order. She suggested on this basis that learner language

proficiency needs to be considered when an instructional treatment is applied, as is proposed in work on aptitude treatment interaction (ATI), where effects are optimal when the instructional treatment and a learners' aptitude are matched (Cronbach & Snow, 1977).

Examining the order effect on second language (L2) reading comprehension, Chen (2012) reported that a group who read an English text first and watched the film version with Chinese captions scored second significantly higher in reading comprehension than did those who watched the film first in English and read the text then; the former group also significantly outperformed those who read the text first and then watched the film in English. The research concluded that language learning may benefit when first language (L1) captions are added to a film, but that L2 captions may not be helpful for novice learners. Overall, these past studies demonstrate that instructional sequence order affects reading comprehension.

Similarly, instructional sequence order may have an impact on student listening and speaking performances in language learning. Examining the effects of viewing order of captioned videos, Winke, Gass, and Sydorenko (2010) found that viewing the video first with captions and then without was more effective for listening performance than the reverse. They concluded that this order helps reduce learners' anxiety and produced a positive effect on subsequent recognition. Winke et al. thus suggested that captions can serve as a useful tool to facilitate second language learning.

Research Questions

The purpose of the second replication study was to examine the effects of showing video clips to EFL learners to develop their English language and public speaking skills. With the preceding literature review of sequence order effects and the scarcity of empirical research on the use of different levels of model video to develop EFL learners' performances, this replication may encourage language teachers to use learners' performance videos to enhance the quality of learners' public speaking performances. It is also hoped that EFL students will benefit from carefully observing video-recorded performances of other learners, as well as their own. To these ends, two research questions were posed:

1. How does speech model video observation affect EFL learners' speaking performance in the classroom?
2. Are learners able to benefit more from viewing speech model videos in less-to-more-proficient speaker sequence order or vice versa?

Method

Participants and Setting

During the Fall 2015 semester, 30 Japanese students (26 males, 4 females; reflecting the gender balance in the economics department) enrolled in two different English communication courses at a university in the Tokyo metropolitan area were asked to participate in this study. These individuals were freshmen economics majors who had studied English for at least six years in junior and senior high school. Prospective participants were informed that partaking in the research would not require any additional work or affect their grades. Permission was sought to use video-recordings of students' presentations as well as textual data (their qualitative responses) for future research or educational purposes. Four students did not consent to having their data collected, and two failed to complete the course. Hence, there were a total of 24 participants (10 and 14 in Groups A and B, respectively).

Students were assigned to the course at the beginning of the academic year, according to their Test of English for International Communication (TOEIC) Bridge scores, and thus were not randomly selected; the average scores for Groups A and B were 130 and 116, respectively.¹ There was a statistically significant difference between test scores of the two groups, $t(18.86) = 2.55, p = .019$; that is, English language proficiency was not equivalent between the groups. However, only 21% of the variance in test scores was explained by variance between groups, so no substantial meaning was attached to this discrepancy. Therefore, we did not further consider English language proficiency as an intervening variable in this study.

The course was mandatory for graduation, and emphasized developing oral communication skills through group/pair work. Classes were 90 minutes long and were held twice weekly throughout a 14-week semester; the content and textbook used in both groups were identical. Students were informed that components of the research project (regardless of whether their data were retained) would constitute 45% of their final grade (30% for presentations and 15% for draft submissions and self-/peer evaluations).

Materials

Model video. Video clips were prepared from 16 clips produced by Japanese EFL students from preceding years with the consent of each student, selected subjectively by the instructor. Clips of eight more-proficient speakers were selected, primarily from among students in higher-level English classes. These videos demonstrated outstanding voice control and body language in their performances, and therefore, were considered the most appropriate to help

enable other students to reach their performance targets. Conversely, eight video clips of less-proficient speakers were taken from lower-level English classes. These speakers still showed deficiencies in areas such as pausing and pronunciation, although their efforts to prepare for their performances seemed apparent despite their low proficiency in their target language.

Self-/peer evaluation. To evaluate student performance, self-/peer evaluation forms were used. These evaluations helped to reflect on students' learning and provide feedback to other students. Based on Yamashiro and Johnson's (1997) study on developing public speaking skills in the EFL context, the current study used 11 variables based on a four-point Likert-type scale from 1 (weak) to 4 (great). Items 1–4 focused on voice control (i.e., projection, pace, intonation, diction), Items 5–8 on body language (i.e., posture, foot/hand position, eye contact, facial expression) and Items 9–11 on effectiveness (i.e., topic choice, language use, vocabulary). In addition, a comment column was included in the form.²

First and second reflections. After the second and third rounds of performance, participants completed this form to comment on how viewing the model videos had affected their own speaking performances.

Final reflection. After the third round of evaluation, a three-section form was administered to identify the effects of viewing both model videos and students' own recordings. In Section One, participants were asked to answer how difficult they had found preparing scripts for each topic on a five-point Likert-type scale from 1 (easy) to 5 (difficult). In Section Two, to measure the effect of viewing models and students' own recordings on their speaking performance, students were asked to rate the overall benefit on a five-point Likert-type scale from 1 (not useful at all) to 5 (very useful). In Section Three, students were asked to write in Japanese their general thoughts on viewing either more- or less-proficient speaker videos, to identify differences in their attitudes by video type and whether different types of model were beneficial for different purposes.

Data Collection Procedure

In order to obtain data, participants were instructed to deliver three memorized speeches. The themes were "My favorite food," "Memories of high school club activities," and "My ideal trip." Students were taught how to compose a three-paragraph presentation, consisting of an introduction, body, and conclusion, between 180 and 200 words. Prior to each speech

performance, students submitted scripts to the instructor and received at least some feedback on content, organization, and grammar.

In Japanese EFL contexts, “speech” generally implies a prepared talk in which students memorize a script on a theme and present it in front of others. While Duong (as cited in Duong & Nguyen, 2006) argues that learners should avoid using memorization because “they then do not interact with the audience and lose naturalness in communicating,” Duong and Nguyen (2006) point out that “memorization seems to be a valid learning strategy, provided that memorization is used appropriately to help students internalize what they have learned to apply in actual communication” (p. 2). Thus, students in this study were required to complete a written piece on the given theme and then use it as a script for their talk.

Three speaking performances were video-recorded and simultaneously evaluated by the instructor. Immediately after the second and third rounds of performances, participants completed a student performance reflection form.

For student evaluation, recordings of speaking performances were shown to students and self-/peer evaluations were conducted using the relevant form. Students evaluated their classmates and returned the completed forms to the respective speakers. The instructor collected the forms shortly thereafter for future analysis. During the third round of evaluation, students were shown 30-second clips from each of their initial performances to prompt them to consider how their performances had evolved; afterward, they completed a video-observation reflection form.

Prior to the second and third rounds of performance, respectively, students were shown model speaker videos and instructed to identify the speakers’ weaknesses and strengths. Before the second round of performance, Group A viewed the more-proficient speakers and Group B the less-proficient speakers; this order was reversed before the third round. After viewing and commenting on two to three clips, students in each session were divided into groups of three or four, wherein they discussed the models for a few minutes. The data collection procedure is shown in Figure 1.

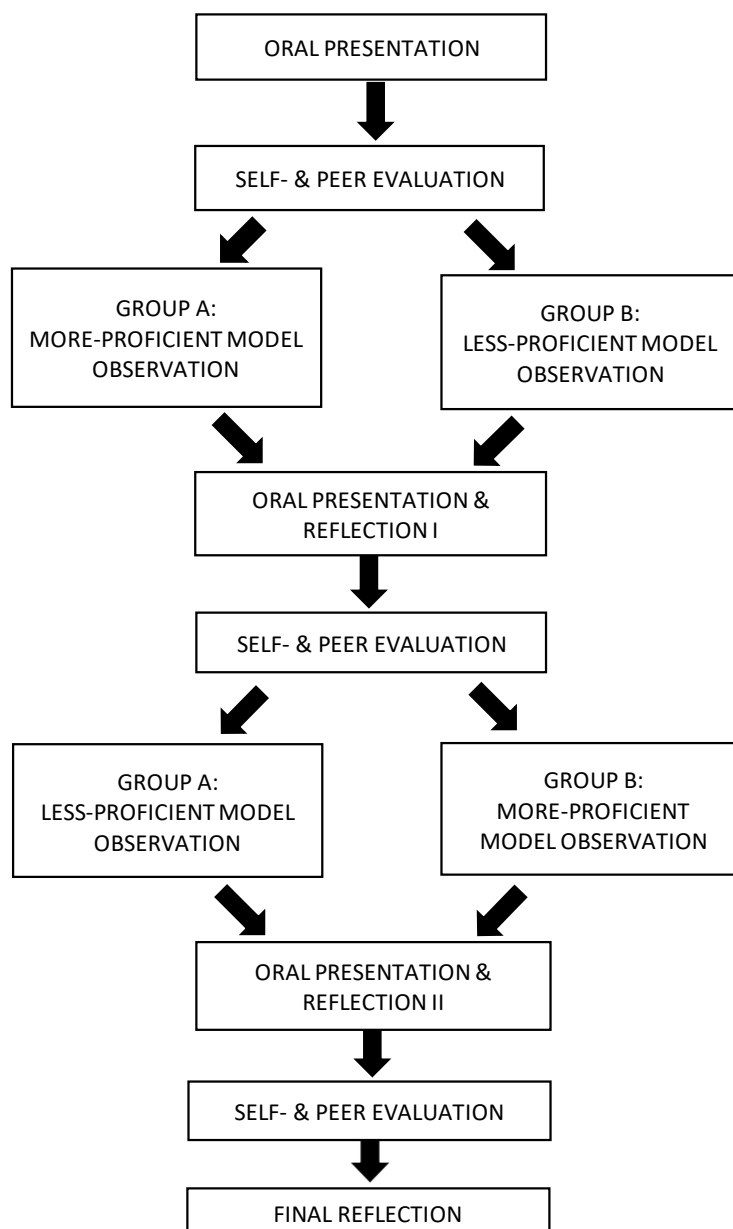


Figure 1. Overview of Teaching Procedure. Quoted from Okada et al. (2017)

Research Design

Since this was a classroom-oriented research study using intact classes, a quasi-experimental revised non-equivalent groups pre-test/post-test design was adopted. Additionally, this study employed a mixed-methods design; more specifically, a convergent design was employed, in which quantitative and qualitative data were collected and analyzed separately to merge the results of these analyses (Creswell, 2015). Quantitative data were collected through self-/peer evaluations, whereas qualitative data were obtained based on text mining analysis of the three reflections.

Results

Quantitative Results

Repeated measures analysis of variance (ANOVA) was used to examine the effect of group (A vs. B [between-participants factor]), performance round (first vs. second vs. third [repeated factor]), and their interaction with self-/peer-evaluated scores. After confirming high internal consistency, ratings for the 11 evaluation items were averaged as follows: voice control (Items 1-4), body language (Items 5-8), and effectiveness (Items 9-11). The Cronbach's alpha coefficients for the subscales were calculated separately for each performance round (.85, .82, .95 [first]; .83, .82, .94 [second]; .89, .88, .98 [third]) and the peer ratings (.89, .66, .82 [first]; .93, .77, .97 [second], .89, .64, .95 [third]). Analysis was conducted using SPSS 22. In the following analyses, the Greenhouse-Geisser correction was used where the sphericity assumption was not met for the repeated factor.

Table 1.

Mean Scores and SDs for the Self-Evaluations

	Round 1			Round 2			Round 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>
Group A									
Voice	10	2.28	0.75	10	2.48	0.69	9	2.64	0.71
Body	10	2.15	0.84	10	2.48	0.70	9	2.75	0.66
Effect	10	2.47	1.04	10	2.57	0.79	9	2.89	0.60
Group B									
Voice	14	2.70	0.65	14	2.84	0.50	14	3.11	0.70
Body	14	2.71	0.70	14	3.04	0.52	14	3.18	0.61
Effect	14	3.33	0.58	14	3.26	0.51	14	3.36	0.61

Note. Data from absent students were eliminated at each analysis stage, but a complete set of data were used to calculate ANOVA.

Voice: voice control; Body: body language; Effect: effectiveness.

The mean scores for self-evaluated voice control, body language, and effectiveness are provided in Table 1. A two-way ANOVA (group \times round) for voice control revealed a

significant main effect for round, $F(2, 42) = 4.69, p = .02, \eta_p^2 = .18$. A Bonferroni post hoc test for the main effect showed that voice control evaluations for the third round were significantly higher than for the first ($p = .04$); nevertheless, no other effects were significant: group, $F(1, 21) = 2.54, p = .13, \eta_p^2 = .11$; group \times round, $F(2, 42) = 0.17, p = .85, \eta_p^2 < .01$.

A two-way ANOVA for body language also revealed a significant main effect for round, $F(2, 42) = 8.67, p < .01, \eta_p^2 = .29$. A Bonferroni post hoc test for the main effect showed that body language evaluations for the third round were significantly higher than for the first ($p < .01$). The main effect for group was marginally significant, $F(1, 21) = 3.72, p = .07, \eta_p^2 = .15$, while the interaction effect for group \times round was not significant, $F(2, 42) = 0.18, p = .84, \eta_p^2 < .01$. A two-way ANOVA for effectiveness showed a significant main effect for group, $F(1, 21) = 7.29, p = .01, \eta_p^2 = .26$; namely, effectiveness was significantly higher among students in Group B. No other effects were significant: round, $F(2, 42) = 1.64, p = .21, \eta_p^2 = .07$; group \times round, $F(2, 42) = 0.87, p = .43, \eta_p^2 = .04$.

Table 2
Mean Scores and SDs for the Peer Evaluations

	Round 1			Round 2			Round 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>
Group A									
Voice	9	3.34	0.31	10	3.28	0.44	10	3.36	0.25
Body	9	3.34	0.19	10	3.37	0.20	10	3.40	0.16
Effect	9	3.52	0.12	10	3.35	0.40	10	3.46	0.09
Group B									
Voice	14	3.05	0.39	14	3.13	0.40	14	3.45	0.22
Body	14	3.07	0.25	14	3.21	0.29	14	3.50	0.13
Effect	14	3.34	0.17	14	3.37	0.16	14	3.61	0.11

Note. Data from absent students were eliminated at each stage, but a complete set of data were used to calculate ANOVA.

Voice: voice control; Body: body language; Effect: effectiveness.

The mean scores for peer-evaluated voice control, body language, and effectiveness are provided in Table 2. A two-way ANOVA (group \times round) for voice control revealed a significant main effect for round, $F(2, 42) = 7.35, p < .01, \eta_p^2 = .26$, but not group, $F(1, 21) = 0.62, p = .44, \eta_p^2 = .03$. However, this main effect for round was qualified by a significant group

× round interaction, $F(2, 42) = 5.31, p < .01, \eta_p^2 = .20$. Post hoc tests showed that Group B's third round of voice-control evaluations was significantly higher than the first or second rounds ($p < .01$ and $p = .01$, respectively); however, this was not apparent in Group A's results (see Figure 2).

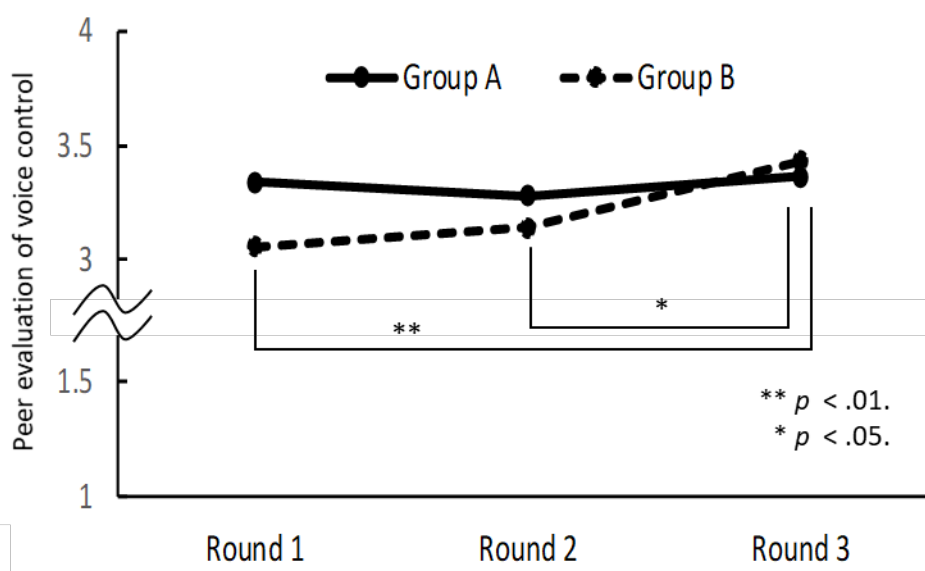


Figure 2. Peer Evaluation of Voice Control as a Function of Round and Group.

A two-way ANOVA for body language also revealed a significant main effect for round, $F(2, 42) = 16.51, p < .01, \eta_p^2 = .44$, but not group, $F(1, 21) = 1.63, p = .22, \eta_p^2 = .07$. The main effect for round was qualified by a significant group × round interaction, $F(2, 42) = 11.34, p < .01, \eta_p^2 = .35$. Post hoc tests showed that Group B's third round of body language evaluations was significantly higher than the first or second round ($p < .01$ for both), and that ratings for the second were also significantly higher than the first ($p = .02$). A similar phenomenon was not apparent in Group A's results (see Figure 3).

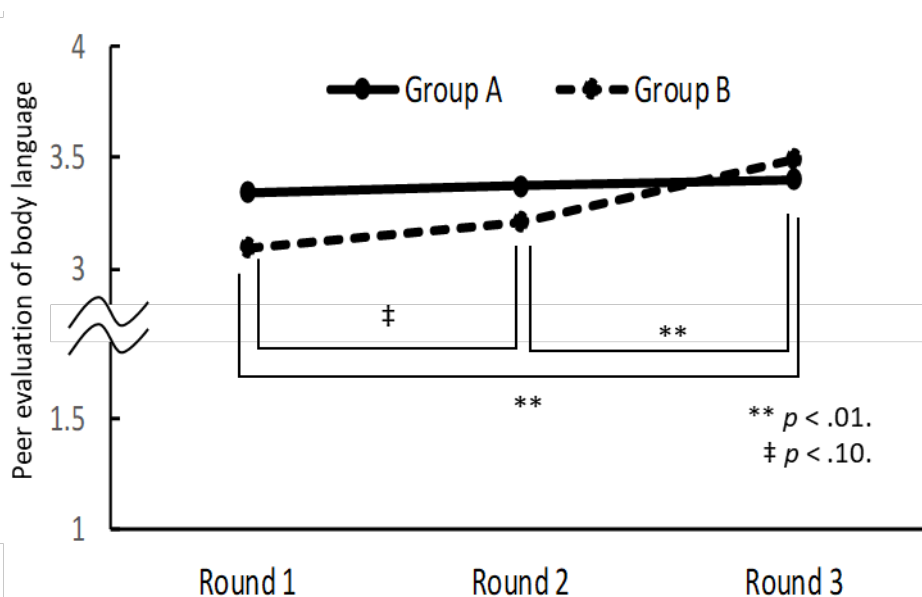


Figure 3. Peer Evaluation of Body Language as a Function of Round and Group.

The two-way ANOVA for effectiveness showed a significant main effect for round, $F(1.32, 27.66) = 7.02, p < .01, \eta_p^2 = .25$, but not group, $F(1, 21) = 0.03, p = .87, \eta_p^2 < .01$. This main effect of round was qualified by a significant group \times round interaction, $F(1.32, 27.66) = 6.16, p = .01, \eta_p^2 = .23$. Post hoc tests showed that Group B's third round of effectiveness evaluations was significantly higher than the first or second rounds ($p < .01$ for both), although such a finding was not apparent in Group A's results (see Figure 4).

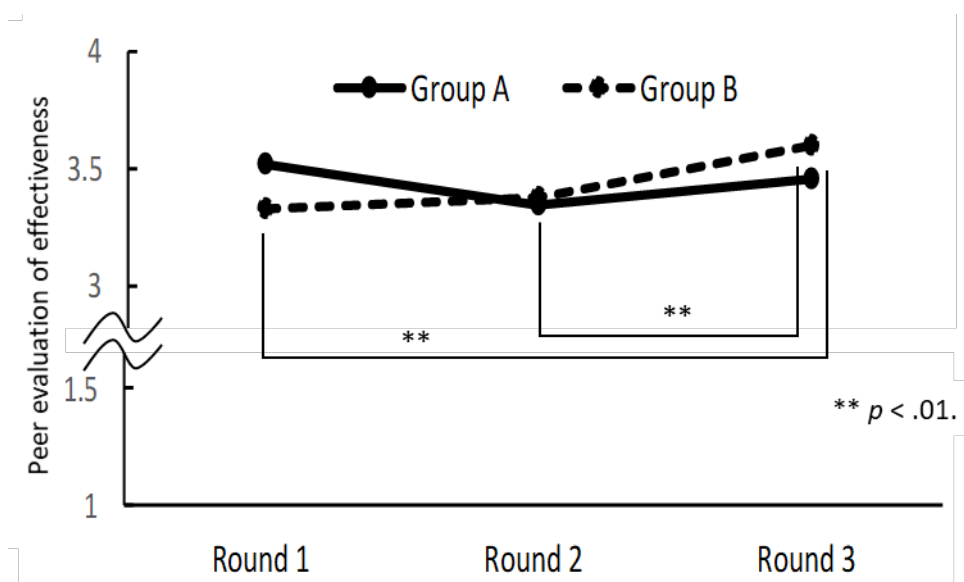


Figure 4. Peer Evaluation of Effectiveness as a Function of Round and Group.

Text Mining and Qualitative Results

Text mining is an analytic process that enables numeric measurement of qualitative data. In this study, featured words were extracted using Text Mining Studio 5.1, by NTT DATA Mathematical Systems Inc., to determine the effects of model speaker videos on students' performances as well as to collect their thoughts and feelings regarding the more- and less-proficient speaker videos. In addition, the textual data were analyzed qualitatively to explain the text mining results.

Table 3

Featured Words in Reflection I

	Word	Meaning	Frequency
1	思う	to think	27
2	意識	consciousness	15
3	発表	presentation	14
4	話す	to speak	13
5	スピーチ	speech	12
6	見る	to watch	10
7	人	person	9
8	声	voice	9
9	良い	good/well	7
10	姿勢	posture	6

Table 4

Featured Words in Reflection II

	Word	Meaning	Frequency
1	思う	to think	15
2	見る	to watch	8
3	ビデオ	video	6
4	人	person	5
5	できる	to be capable	4
6	やる	to do	4
7	話す	to speak	4

Reflections I & II. Tables 3 and 4 respectively show words that appeared in students' reflections following the second and third rounds of performance. In both rounds, the word *to think* appeared the most frequently, due to the nature of the question. In addition, words involving sound and vision (e.g., *to speak, speech, to watch, voice, posture*) were often used to describe specific aspects of the performance. For instance, students in Group A reported that after watching more-proficient speakers, they imitated the models to improve their speaking performance.

- I tried to stand straight and hold my hands during my performance, because I thought that I could easily imitate these behaviors from the video. (S1)
- I learned to imitate the more-proficient speakers because the speakers looked calm. I consequently tried not to move my hands or body during my own performance. (S2)

On the other hand, a student in Group B pointed out the importance of having the audience understand what the speaker is saying.

- While watching the less-proficient speaker models, I realized that it was important to be understood by one's listeners, and so I tried to speak clearly at an appropriate volume. I performed better than in my initial performance, but I am still not satisfied. I wish that I could pronounce English more smoothly to deliver a better performance. I also want to understand all of what others are saying. (S3)

Table 4 indicates the words that appeared with high frequency in the second student performance reflection. Students used words associated with cognitive functions (e.g., *to think, to watch, to be capable, to do* and *to speak*) to describe their own performances, whereas the word *video* was used when referring to the models, whether more- or less-proficient. For example, after watching less-proficient speakers, students in Group A focused on how the models were speaking, because they did not want to perform similarly themselves.

- After watching the videos, I made it a point to try and smile while delivering my performance. I did not want to stop while I was speaking, but I did. In the future, however, I will work harder. (S5)
- I made a mistake during my previous performance...and therefore spoiled the mood. This time, I rehearsed several times so that I would not stumble over the words. I was careful not to move my body or hands, since such behavior seemed to cause the less-proficient speakers to lose their concentration. (S6)

Like students in Group A in the initial reflection, those in Group B after the second reflection emphasized that viewing the more-proficient speakers motivated them to prepare for their performance.

- Since I noticed that more-proficient speakers tended to pronounce English clearly with appropriate intonation, I verified the pronunciation of unfamiliar words, and likewise focused on stressing important words. (S7)
- I thought that I would deliver my performance as before, since the more-proficient speakers seemed too good to imitate. Nevertheless, I wanted to improve my performance. (S8)

Final reflection. In Section One, students responded that the topics were equally difficult overall; M and SD values for each were as follows: “My favorite food” ($M = 3.50$, $SD = 1.18$), “Memories of high school club activities” ($M = 3.54$, $SD = 1.38$) and “My ideal trip” ($M = 3.42$, $SD = 1.28$). There was no significant difference in difficulty found between the three topics, $F(2, 46) = 0.08$, $p = .93$, $\eta_p^2 < .01$. In Section Two, the perceived usefulness of each video type was as follows: model videos ($M = 3.58$, $SD = 0.78$), own videos ($M = 3.88$, $SD = 0.95$), peer videos ($M = 3.63$, $SD = 0.82$). There was no significant difference here either, $F(2, 46) = 1.32$, $p = .28$, $\eta_p^2 = .05$, indicating that the three videos were equally helpful for students in both groups.

Table 5

Featured Words in the Final Reflection

	Word	Meaning	Frequency
1	思う	to think	27
2	レベル	level	16
3	見る	to watch	16
4	スピーチ	speech	13
5	ビデオ	video	13
6	人	person	13
7	平均的レベル	average level ³	8
8	わかる	to understand	7
9	参考	reference	6
10	する	to do	5

As shown in Table 5, *to think* was the most frequent word again in this reflection, consistent with the results of Tables 3 and 4. Other words that frequently appeared, such as *to watch*, *speech*, *video* and *reference*, seemed to indicate that the models assisted students in giving their performances. For example, a student in Group B noted,

- Aside from helping me to evaluate my own performances, I was able to see how well other students delivered their performances. I also became aware of my weaknesses, and [therefore] tried to imitate the more-proficient speakers. Watching my own and other students' videos enabled me to explore ways to possibly improve my performance. (S11)

Students in both groups also reported that more- and less-proficient speakers helped them enhance their speaking performance.

- I clearly understood the weaknesses in each model. I therefore learned how to improve my performance by comparing the more- and less-proficient speakers. (S10, Group A)
- There was a significant difference in performance between the more-proficient and less-proficient speakers. The more-proficient speakers spoke almost perfectly at an appropriate pace, and presented suitable content. I attempted to imitate their eye contact and posture, since imitating their pronunciation seemed too difficult. I compared my performances to those of the less-proficient speakers to identify weaknesses that I could improve upon. (S12, Group B)

A student in Group A commented on the negative impact of viewing more-proficient speakers first on his or her speaking performance.

- It was beneficial for me to view the less-proficient speaker models, because my performance was not so good. I then felt compelled to watch the more-proficient speakers. I believe that it is better to watch the less-proficient speakers first. (S9)

Discussion

Quantitative analyses found significant performance round and group interaction effects in peer-evaluated scores for Group B, who watched less-proficient models first and more-proficient models next. These findings were consistent with those of Okada et al. (2017). It appears that this sequence order of model viewing had a positive impact on student performance.

It should be emphasized that students in both groups were taught similar content by the same instructor; therefore, it was expected that their language- and speech-related skills would

develop solely through regular lessons. A performance gap between the model speakers and students would have been natural; if the gap was too large at the instruction, students would not be motivated for delivering a speech; if, however, it was not, they would be motivated for practicing further. We would then assume that the viewing of less- and more-proficient speaker models assisted them during each instruction for Group B. On the other hand, for students who watched more-proficient first, it might be too difficult to model at the instruction, and therefore this video viewing order would probably not have helped students develop their language and public speaking skills. In summary, our findings from quantitative analyses support the idea that students' development over the course should be carefully considered and such speaker models should be shown to students.

Given that Group A's peer evaluations for the three aforementioned factors were already high following the first round of performance, this group might not have had as much room for improvement, resulting in a ceiling effect. Accordingly, peer-evaluated scores in this group may not have improved as much as those in Group B, regardless of model viewing order. There were two possibilities that may have caused a ceiling effect in Group A. First, the average TOEIC Bridge score for Group A was higher than that of Group B; it is likely that Group A's performances were already satisfactory in the first round of performance. Second, there is a possibility that students in Group A were more lenient than those in Group B, and thus the former group scored higher in peer evaluation. Future research should disentangle these possibilities using a more simplified teaching procedure to examine model viewing effects on EFL learners' performance.

On the other hand, qualitative analysis found that video observation positively affected students' speaking performances, supporting the findings of Okada et al. (2017). For example, viewing more-proficient speakers helped participants enhance their English pronunciation and public speaking skills and, seemingly, develop greater awareness of other students as well as greater self-awareness, thereby improving their speech performance. The less-proficient speaker models also helped participants improve their performance while being aware of the speakers' weaknesses. There is thus a possibility that either more-proficient models or less-proficient ones enable students to raise the quality of their speech presentations for different purposes and for different reasons.

Due to the small quantities of textual data from the three reflections, there was little variety in the words shown in the tables. It might also have been possible to clarify differences in the reflections between the two groups, if more textual data had been collected in the study. Nevertheless, these featured words were informative enough to encourage us to further examine textual data in response to our research questions.

Pedagogical Implications

Live and video-recorded models may both be useful, but for different purposes and for different reasons. That is, it is not a contest between them; rather, they complement each other. Viewing video-recorded models can help learners enhance their linguistic awareness and boost public speaking skills, as can watching recordings of their own performances, while live models can help them understand how speakers deliver speeches in front of an audience that is actually looking into the speakers' eyes and hearing their voice. As mentioned previously, Bandura's (1971, 1977) observational learning theory is comprised of four stages: attention, retention, reproduction, and motivation. His theory was originally used to describe the acquisition of behavior among children; however, language learners who observe a video presentation may undergo a comparable progression of stages. For instance, the participants in this study viewed both more- and less-proficient speakers (involving attention) while focusing on their positive and negative aspects (retention). Students then gave performances and concentrated on imitating the models (reproduction). Finally, learners viewed recordings of their own performances, and noted improvements and aspects requiring further refinement (motivation). It seems, then, that video-based observational learning contributes to enhancing the quality of learners' public speaking skills in the foreign language classroom.

Theoretical Implications

This replication study aimed to use non-native speaker model videos to improve broad-based language skills, including English pronunciation and public speaking skills. While students were taught English pronunciation in their regular lessons, it was not included as an evaluation item; however, we would expect these students to benefit from practicing pronunciation in the context of an exercise like this, especially since most of them had not had instruction in English pronunciation in junior or senior high school.

As Murphy (2014) pointed out, characteristics in the paralinguistic (e.g., facial expressions), linguistic, and rhetorical (e.g., topic) realms play an important role in communication by helping ensure that the speaker is intelligible to listeners. The present study indicates that non-native speakers' video clips affect students' learning performance outcomes by making them more aware of these paralinguistic and linguistic aspects. The primary reason to use non-native speaker models, not native speaker ones, is that they are able to encourage learners and inspire confidence in them (Adams, 2004). It may be therefore reasonable to consider that teaching public speaking skills using different types of models can be beneficial for EFL learners who deliver speech performances, as the learners are able to manifest what they observed in and learned from the videos regarding linguistic, paralinguistic, and rhetorical skills.

Limitations

This study contained some of the same limitations as the first replication study. First, as discussed in Okada et al. (2017), this was a quasi-experimental study, in which there was not a large sample of participants to generalize the findings of the study. Second, although the study's purpose was to examine the effects of model video observation and its sequence order on students' speaking performances, several other factors might have affected students' evaluation of their speaking performances. Additionally, to evaluate students' speaking performances, self- and peer evaluation scores were used for analysis; future research should employ teacher evaluation in addition to student evaluation, so that it can avoid potential problems with the internal validity of the study.

Conclusion

In this replication of Okada et al. (2017), we found very similar results: (1) learners who watched less-proficient speaker models first and more-proficient ones second showed a significantly higher improvement in peer evaluation scores; (2) more-proficient speaker models were used for learners to imitate the speakers' strengths, and less-proficient speaker models were used to raise learners' awareness of self/others to improve their own speaking performance. While the study failed to quantitatively show that the viewing order of model speaker videos solely affect learners' speech performance due to the several limitations of the study, the qualitative results could have important implications for language teachers and researchers. The first implication is that it may be necessary for students not only in Japan but also in other Asian countries to improve public speaking skills as part of curriculum. As many American universities do (Hancock et al., 2010), Japanese institutions could offer courses that focus on improving students' public speaking skills. Second, having an opportunity to develop public speaking skills seems to be essential for students in a globalizing society. For those who are afraid of public speaking, it may be helpful and practical to have their own public speaking performance video-recorded so that they can become aware of their strengths and weaknesses to improve their speaking skills. As a multi-media assisted learning tool, video-based observational learning as well as reflective learning brings about a possibility that learners will be able to develop public speaking skills in the EFL classroom without having a native speaker model.

Acknowledgements

This work was supported by the Japanese Society for the Promotion of Science (JSPS Kakenhi Grant Number 15K02730) and the Research Institute for Language Education at Seisen University, Tokyo. We would like to thank Editage (www.editage.jp) for English language editing.

Footnotes

- ¹. Although Group B comprised 14 students, only 13 scores were computed, since one student did not complete the test.
- ². The comments in the column were excluded from the analysis because this paper focuses specifically on student performance and video-observation reflections.
- ³. The term “average level” was used with the participating students as an alternative to “less-proficient,” to avoid negative impressions of the video clips.

References

- Adams, K (2004). Modeling success: Enhancing international postgraduate research students' self-efficacy for research seminar presentations. *Higher Education Research and Development, 23*(2), 115–130. <http://dx.doi.org/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.
- Chen, M. L. (2012). Effects of the order of reading text or viewing a film and L1/L2 captions on reading comprehension. *Perceptual and Motor Skills, 115*(1), 18–26. doi 10.2466/23.PMS.115.4.18-26
- Creswell, J. W. (2015). *A Concise Introduction to Mixed Methods Research*. Los Angeles, CA: Sage Publications, Inc.
- Cronbach, L., & Snow, R. (1977). *Aptitudes and instructional methods: A handbook for research on interactions*. New York, NY: Irvington.
- 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*(2), 256–271. doi.org/10.1080/07294360.2013.832155

- Duong, T. H. O., & Nguyen, T. H. (2006). Memorization and EFL students' strategies at university level in Vietnam. *The Electronic Journal for English as a Second Language*, 10(2), 1–21.
- Hancock, A. B., Stone, M. D., Brundage, S. B., and Zeigler, M. T. (2010). Public speaking attitudes: Does curriculum make a difference? *Journal of Voice*, 24(3), 302–307.
- 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*. Tokyo: MEXT.
- Murphy, J. M. (2014). Intelligible, comprehensible, non-native models in ESL/EFL pronunciation teaching. *System*, 42, 258–269. doi.org/10.1016/j.system.2013.12.007
- 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 6th CLS International Conference* (pp. 394– 413). Singapore: National University of Singapore, Centre for Language Studies.
- Okada, Y, Sawaumi, T, & Ito, T. (2017). Effects of observing model video presentations on Japanese EFL learners' oral performance. *Electronic Journal of Foreign Language Teaching*, 14(2) 129–144.
- Shimizu, M. (2007). Effects of inference question sequence on reading comprehension. *Annual Review of English Language Education in Japan*, 18, 51–60.
- Shrosbree, M. (2008). Digital video in the language classroom. *The JALT CALL Journal*, 4(1), 75–84.
- Winke, P., Gass, S., & Sydorenko, T. (2010). The effects of captioning videos used for foreign language listening activities. *Language Learning & Technology*, 14(1), 65–86.
- Yamashiro, D. A., & Johnson, J. (1997). Public speaking in EFL: Elements for course design. *The Language Teacher*, 21(4), 13–17.