<div style="border:1px solid">Note</div>

Factors Affecting Student Evaluation of Video-Recorded Performance

OKADA, Yasuko

SAWAUMI, Takafumi

I　Introduction

Peer evaluation and self-evaluation have gained much attention since educators' concentrated studies involving learning assessment in the 1950s.[1] These two types of student-based evaluation can develop learners' responsibility for their own work or their peers' work and can be an effective educational tool to facilitate students' reflection on their own learning process. While self-evaluation is defined as learners' involvement in assessing their own learning and the products of their learning,[2] peer-evaluation is a learning process in which they assess their peers' work. Self- and peer evaluation can be implemented separately, but it is meaningful to combine the two and use them to assess students' learning outcomes because "students are assessing peers but the self is also included as a member of the group and must be assessed."[3] In the foreign language classroom, learners rarely assess their peers' speaking or oral presentation skills, partly because they have difficulty evaluating the verbal and nonverbal aspects of the performances, such as pronunciation, pausing, stressing, and intonation. This may be because they are nervous knowing they will soon have their own performance evaluated by their peers. It may therefore be crucial for learners to evaluate their peers' performance and have their performance evaluated by peers simultaneously.

Video recording can solve these problems because it offers learners an opportunity to assess their performance by reviewing the recording. They can watch themselves as many times as they wish, thereby reducing their burden as assessors. As previous studies suggest, video technology can

---

[1]　Falchikov, *Improving assessment through student involvement.* RoutledgeFalmer: New York, 2005, p. 81.

[2]　Boud & Falchikov, "Quantitative studies of self-assessment in higher education: A critical analysis of findings." *Higher Education,* 18, 1989, p. 529.

[3]　Dochy, Segers, & Sluijsmans, "The use of self-, peer and co-assessment in higher education: A review." *Studies in Higher Education,* 24（3）, 1999, p. 340.

"provide a welcome break from the rigors of more traditional study."[4] For this paper, therefore, the researcher explored the differences in peer evaluation and self-evaluation using video recordings in an English as a Foreign Language (EFL) classroom. After a brief overview of the background of the previous studies on video recording and learner evaluation in language classrooms, the current study investigates how Japanese EFL learners assess their own and their peers' oral presentations using videos. It also examines the ways learners with different levels of English proficiency evaluate their own performances by watching the video recordings.

Ⅱ   Literature Review

1   Learner Involvement in Assessment

Along with the recent focus on teaching communicative skills, various studies have involved students in assessment, examining whether self- and peer assessment are effective in teaching speaking skills in the target language. For example, a study that compared self-assessment with teacher assessment in an EFL classroom at a university in Taiwan,[5] showed that with sufficient training and support, learners began to assess their performance as well as their teacher did. In addition, their attitudes toward self-assessment seemed favorable because of sufficient training and evaluative feedback.

With regard to peer assessment of oral presentation skills in EFL classrooms, Shimura examined how three groups (Advanced, Upper Intermediate, and Lower Intermediate) of Japanese learners with different levels of English proficiency assessed their peers' presentations and evaluated how their assessment differed from that of their instructors.[6] Her study reported the underlying supposition that with higher English proficiency, the students' ratings could be closer to those of their instructor, was not supported. It was revealed, however, that ratings by the Upper Intermediate students were close to that of their instructor, that the Advance students tended to give good grades to every student in a narrow range, and that Lower Intermediate students had difficulty with peer evaluation and tried to distinguish rating in a wider range.[7] Moreover, the Lower and Upper Intermediate groups depended on more visual information to assess their peers' performance than

---

4)   Shrosbree, "Digital video in the language classroom." *The JALT CALL Journal,* 4 (1), 2008, p. 76.

5)   Chen, "Learning to self-assess oral performance in English: A longitudinal case study." *Language Teaching Research,* 12 (2), 2008, pp. 235-262.

6)   Shimura, "Peer- and instructor assessment of oral presentations in Japanese university EFL classrooms: A pilot study." *Waseda Global Forum,* 3, 2006, pp. 99-107.

7)   *Ibid.,* p. 105.

the Advance groups did.[8] Her findings suggest that learners assess their peers differently according to their own proficiency levels.

White's study explored learners' perception of peer assessment and its impact on their learning experience in an EFL context.[9] Students in a university public speaking course, each made two presentations that were videotaped, then assessed their peers and were assessed by the teacher. Afterwards, the researcher administered a student survey to elicit the students' views about the use of peer assessment in the course. The findings showed that the learners' clear understanding of the assessment criteria may have helped them to implement peer assessment effectively, and the peer assessment procedure "did indeed help support and promote student learning about constructing, delivering and judging effective presentations."[10]

As mentioned above, there has been growing interest in the use of peer assessment and self-assessment to teach oral communication skills in language learning, and yet little research has explored whether it is effective for learners to develop their language proficiencies, more specifically their oral presentation skills, through class activities in language classrooms.

2　Using Videos in Language Learning

Video recording has been widely used not only in real-life settings to capture family gatherings and other important events, but also to enjoy watching and sharing videos with others. Even in various educational settings, videos have been accepted to develop teaching skills and learners' performances in order to facilitate the learning process. Recently, technology and education have been closely intertwined, and teachers are expected to use educational ICT tools such as PCs, interactive blackboards, and digital video cameras effectively in order to create interactive classes and make lessons more enjoyable. For example, both teachers and learners can benefit from digital video technology in foreign language learning because it enables them to observe the learners' language development objectively.[11] Previous studies of assessment-purposed videos investigated how Japanese EFL learners evaluated the verbal and nonverbal aspects of their class presentations. In one of the studies, where learners watched video-recordings of their class presentations and evaluated their own performance, they were later able to control their volume and speed of speaking, whereas the pausing and stressing of words seemed problematic for them.[12] The study suggested

---

8)　Shimura, *loc. cit.*

9)　White, "Student perspectives of peer assessment for learning in a public speaking course." *Asian EFL Journal,* 33, 2009, pp. 1-36.

10)　*Ibid.*, p. 30.

11)　Koberl, "Presentable presentations." *ESP-Spectrum*（Slovakia）13, 1997, pp. 12-17.

12)　Okada, "Reflecting on students' videotaped presentation." *Journal of Teaching English,* 20, 2011, pp. 48-49.

that self-reflection using video recordings may enhance learners' awareness of their behavior, both linguistic and nonlinguistic. In her subsequent study, the researcher compared video recordings of a practice session and the final presentation and concluded that learners' linguistic skills and nonverbal behavior improved in the final presentation; however, the learners had difficulties with pausing and stressing words in the target language.[13] These two studies of video use suggest the importance of teaching paralinguistic aspects of the target language.

Ⅲ　Method

1　Research Aims

Based on the above discussion, in the present study, we assumed that Japanese EFL learners would benefit from using the data of video recordings of their classroom performances, as it would allow them to critically analyze their own performances and improve them for the future. Therefore, the first purpose of this study is to investigate how Japanese EFL learners evaluate their peers' and their own classroom performances when reviewing recorded video of the rehearsals and presentations. The second purpose is to examine how learners with different levels of English proficiency evaluate themselves and their peers using video-taped performances in the EFL context.

2　Participants

The participants were 19 college students in two required English communication courses at a university in the Tokyo metropolitan area: seven in one class and twelve in the other. It should be noted that although the study started with 27 students in the two courses, eight students had insufficient data. Therefore, the data from these students were discarded in the subsequent analyses. The participants were all freshmen majoring in economics. Before entering the university, they took the Test of English for International Communication（TOEIC）and were placed in one of the two courses according to their score, which ranged from 185 to 600. The average score was 441 for one group and 291 for the other. It should be noted that there was a significant difference in the test scores between the two classes, $t(9.43) = 3.32$, $p < .01$. However, the students were grouped into high- and low-proficiency levels according to their score because there was a student with a lower score in the high-proficiency class and vice versa.

---

13)　Okada, "Using video feedback comparison with EFL students." *Studies in Language Education Seisen University,* 4, 2012, pp. 17-33.

3　Setting

In fall 2012, the Japanese EFL students were enrolled in English communication courses. The courses emphasized the development of English oral communication skills through various activities with different topics, as well as presentation skills. The class met three hours per week over a 15-week period. The first author was the instructor for both courses.

4　Procedure

As part of the course requirement, the students had to give two oral presentations. The topic of the first presentation was what they had learned from a past experience. The topic of the second was their dream for the future. The first presentation accounted for 15% of their grade, and the second presentation, presentation draft, and video reviewing activity together accounted for 40%, while the remaining 45% consisted of quizzes and homework. Although the importance of training as evaluators was pointed out in the literature review section, the students did not receive sufficient training for this study because of time constraints.

Figure 1 shows the overview of the two-stage process of rehearsal and presentation. The rehearsal for the first presentation and the actual presentations took place in the 12th and 13th class sessions, respectively; for the second presentation, they were in the 26th and 27th meetings, respectively. Before rehearsing their speeches, the students were given instruction intended to improve their language and presentation skills. For example, they were taught how to maintain good posture and eye contact, as well as English pronunciation, rhythm, and intonation throughout the course.

5　Rehearsal & Presentation

For each rehearsal, the students were asked to memorize half their speech because there was not enough time for them to memorize the whole speech. After the rehearsals were recorded, the students evaluated their peers' and their own performance while watching the video recordings during the same class period. It is important to note that the students remained engaged by evaluating all the other students' rehearsals.

In the following session, the students gave their real presentations, which were also recorded. Then, in the next session, they reviewed the video recordings. At this time, they evaluated only their own presentation, because it was considered that the results of peer evaluations would not reflect on their performance in class.

6　Instruments

To evaluate the students' performance, an evaluation form was used（see the Appendix）. The

Class

First Speech: "My Past Experience"

12th

Rehearse the presentation (half the speech),
Review the video-recorded performances, **Self- and peer evaluation**

13th

Presentation

14th

Review the video-recorded performances, **Self-evaluation**

Second Speech: "My Dream"

15th-25th

Write a draft through the peer reviewing process
Receive content and grammar feedback from the teacher

26th

Rehearse the presentation (half the speech)
Review the video-recorded performances, **Self- and peer evaluation**

27th

Presentation

28th

Review the video-recorded performances, **Self-evaluation**
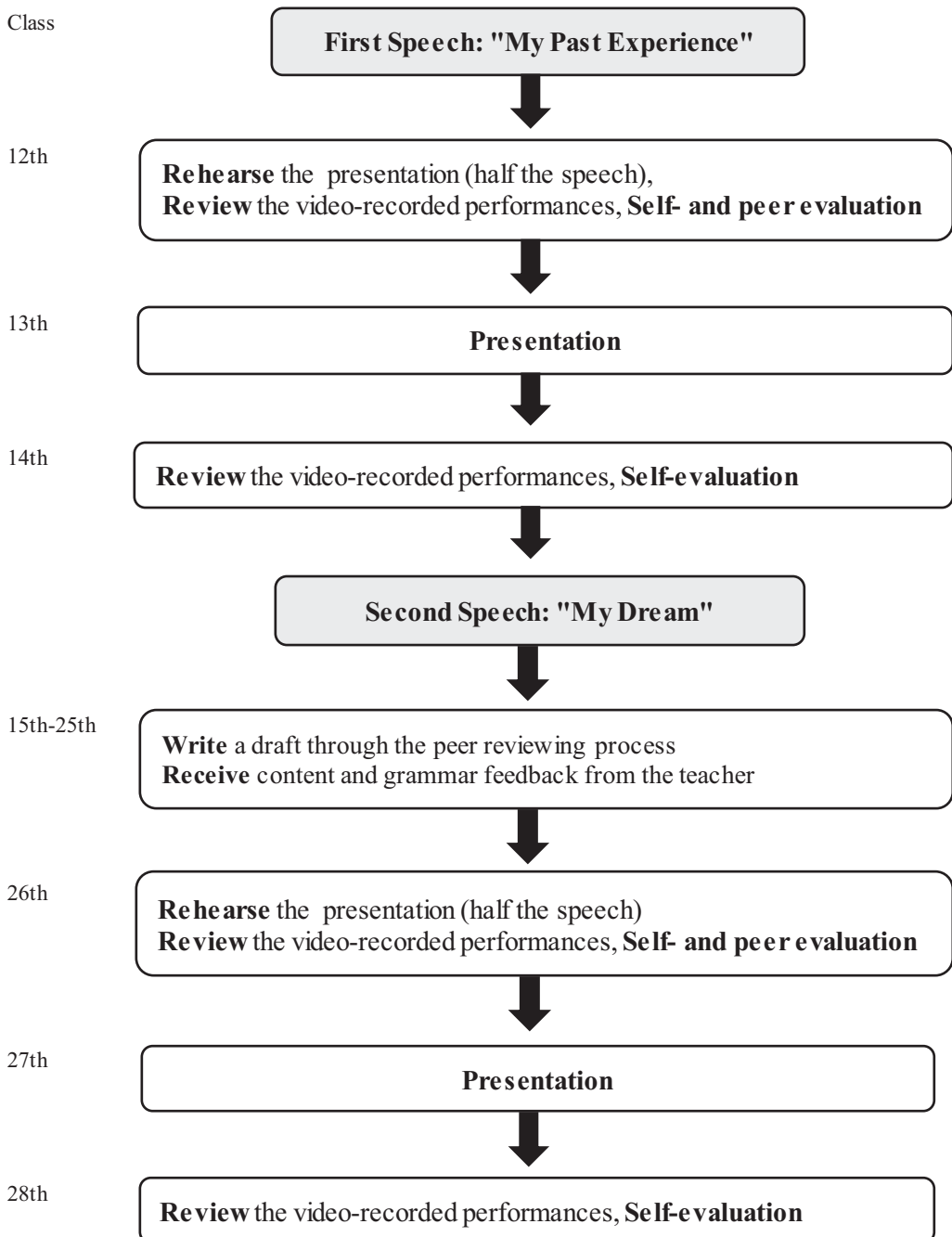
Figure 1 Overview of the Study

students were asked to indicate how they felt about their own and their peers' performance, using a
5-point Likert-type scale that ranged from 1 "strongly disagree" to 5 "strongly agree." Before creating

the evaluation form, the first author reviewed a previous study[14] and modified its oral presentation rubric to fit the context of this study. Ten items were used for the peer and self-evaluation. Among them, four items were relevant to nonverbal aspects of the performance:（1）hand position,（2）foot position,（3）posture, and（4）eye contact. Six items were related to verbal aspects of the performance:（5）speech speed,（6）speech volume,（7）pausing,（8）intonation,（9）rhythm, and（10）articulation. The students had already received a series of instructional inputs on these presentation skills, so they should have understood how to assess them when evaluating their peers' and their own performance. Additionally, the students were asked to comment in Japanese on anything they noticed about their own performance at the bottom of the form.

7　Data Analysis

This study examined the self- and peer evaluation scores for learners' performance using video recordings. For the first research aim, in order to explore the rater difference between the first and second rehearsals and presentations, a paired-sample t-test was performed. For the second research aim, 19 learners in the two groups were first divided into three groups according to their TOEIC score, and then the one-third who had the highest and lowest scores were considered the high-proficiency group（$n = 6$）and low-proficiency group（$n = 7$）, respectively, to enable a comparison of peer and self-assessment by the learners' English proficiency levels. An independent-samples t-test was conducted to examine the learners' peer assessment scores in the first and second rehearsals and self-assessment scores in the first and second rehearsals and presentations. SPSS（Version 20）was used for the quantitative data analysis. The alpha level was set as $p < .05$. In addition, the high- and low-proficiency learners' comments on their first and second presentations were analyzed as qualitative data.

Ⅳ　Results

1　Results of Quantitative Analysis

The results of a paired samples t-test showed that there were significant differences in all items in the first and second rehearsals for self- and peer evaluation; see Tables 1 and 2.

---

14)　Yamashiro & Johnson, "Public speaking in EFL: Elements for course design." *The Language Teacher,* 21（4）, 1997, pp. 13-17.

Table 1　Average Scores Assigned for the First Rehearsal（$N = 19$）

Item	Self		Peer		
	M	SD	M	SD	$t(18)$
1	3.11	1.29	4.34	0.47	4.31 ***
2	3.32	1.16	4.38	0.27	3.76 **
3	2.95	1.18	4.32	0.37	4.81 **
4	2.79	1.32	3.85	0.29	3.78 **
5	3.16	1.17	4.30	0.25	4.03 **
6	3.26	1.19	4.20	0.50	3.68 **
7	2.68	1.11	4.13	0.29	5.50 ***
8	2.74	1.28	3.89	0.33	3.73 **
9	2.42	1.22	3.89	0.29	5.07 ***
10	2.63	1.30	4.03	0.28	4.27 ***

**$p < .01$; ***$p < .001$

Table 2　Average Scores Assigned for the Second Rehearsal（$N = 19$）

Item	Self		Peer		
	M	SD	M	SD	$t(18)$
1	4.16	0.83	4.62	0.37	2.56 *
2	4.16	0.83	4.63	0.24	2.52 *
3	3.42	0.84	4.53	0.27	5.38 ***
4	2.53	1.07	3.74	0.35	4.83 ***
5	3.47	0.84	4.31	0.25	4.00 **
6	3.00	1.25	4.10	0.55	3.96 **
7	3.05	0.91	4.09	0.45	4.70 ***
8	2.89	0.99	3.82	0.37	3.29 **
9	3.11	1.05	3.83	0.33	2.79 *
10	2.95	0.97	3.95	0.38	4.21 **

*$p < .05$; **$p < .01$; ***$p < .001$

Table 1 shows that for all items, the evaluations were significantly higher in the peer evaluation than in the self- evaluation in the first rehearsal. This indicates that the students gave higher scores to their peers' rehearsals than to their own. In the second rehearsal, the mean scores of peer evaluation were also significantly higher than those of self-evaluation for all items, as shown in Table 2.

Table 3 shows the mean self-evaluation scores for the first and second presentations. The results of a paired samples t-test revealed that there was a significant difference in Item 8 (intonation)

between the first and second self-evaluation, $t(18) = 2.92$, $p < .01$. This indicates that intonation was self-evaluated higher in the second presentation. For the other items, there was no significant difference in self-evaluation between the first and second presentations.

Table 3　Average Self-Evaluation Scores Assigned for Presentations（$N = 19$）

Item	First		Second		$t(18)$
	M	SD	M	SD	
1	3.89	1.15	4.05	0.71	-0.55
2	3.84	1.07	4.00	0.82	-0.62
3	3.42	1.22	3.79	0.71	-1.38
4	2.89	1.20	2.74	0.99	0.46
5	3.00	1.33	2.89	0.99	0.33
6	2.84	1.38	3.47	0.90	-1.84
7	2.74	1.24	2.53	0.70	0.89
8	2.16	1.12	3.00	0.82	-2.92 *
9	2.53	1.26	2.74	0.81	-0.89
10	2.63	1.26	3.00	1.05	-1.20

*$p < .05$

Tables 4 and 5 present the means of the high- and low-proficiency learners in the first and second presentations, respectively. As shown in Table 4, an independent-samples t-test revealed that there were significant differences in Items 1 and 2 between the two groups for the first presentation. This indicates that the low-proficiency group self-evaluated higher than the high proficiency group with regard to hand and foot positions for their first presentation.

Table 5 shows a significant difference in Item 7 between the high- and low-proficiency learners for the second self-evaluation, which indicates that the high-proficiency learners self-evaluated their pausing higher than the low-proficiency learners for the second presentation. There was no significant difference in peer evaluation between the two proficiency groups.

Table 4　First Self-Evaluation by Language Proficiency Group

Item	High (n = 6) M	SD	Low (n = 7) M	SD	t(11)
1	2.83	1.47	4.43	0.53	2.68 *
2	2.83	1.17	4.57	0.53	3.54 **
3	3.00	1.41	4.14	0.90	1.77
4	3.17	1.17	3.14	1.57	-0.03
5	2.83	1.33	3.00	1.63	0.20
6	2.50	1.38	3.71	1.25	1.66
7	2.33	0.82	2.86	1.57	0.73
8	2.00	0.89	2.00	1.41	0.00
9	2.17	0.98	2.86	1.68	0.88
10	2.17	0.98	3.29	1.50	1.56

$*p < .05; **p < .01$

Table 5　Second Self-Evaluation by Language Proficiency Group

Item	High (n = 6) M	SD	Low (n = 7) M	SD	t(11)
1	4.00	0.63	4.14	0.38	0.50
2	3.67	1.03	4.14	0.38	1.14
3	3.50	0.84	4.00	0.58	1.27
4	2.83	1.17	2.57	1.13	-0.41
5	2.17	0.75	3.14	1.07	1.87
6	3.33	0.82	3.57	1.13	0.43
7	2.83	0.41	2.14	0.69	-2.14
8	3.50	1.05	2.71	0.49	-1.78
9	2.83	0.98	2.57	0.98	-0.48
10	3.17	1.33	2.71	1.11	-0.67

2　Results of Qualitative Analysis

The analysis of student comments revealed that almost all the learners wrote about weaknesses in their presentations. The high-proficiency learners made more negative comments than the low-proficiency learners about their own second presentations. According to one high-proficiency and four low-proficiency learners, they became nervous or forgot their speech topic when they stood in front of the class to give their first presentations.

Despite the concerns about their performance, two high-proficiency and four low-proficiency

learners commented on their improvement in their second performance, such as better pronunciation, speech volume, and posture.

V Discussion

1 Research Aim 1

The first aim of this study was to examine how differently Japanese EFL learners evaluated their peers' and their own performances when they reviewed recorded video of the rehearsals and presentations. The comparison of self- and peer evaluations for both rehearsals clearly showed that the students scored their peers' performances higher than their own, which seems to be consistent with previous studies on "friendship effects" in peer evaluation.[15] The students in this study had taken other courses together previously, and therefore, they seemed to know each other and be friendly already. It is likely that they wanted to avoid lowering their peers' evaluations even though they had been told that the scores would not affect their grade. Additionally, the standard deviations of peer evaluation in both Tables 1 and 2 are much larger than those of self-evaluation. This indicates that students self-evaluate their performances within a wide range, while they evaluate their peers with a much narrower range, as shown in an earlier study.[16] It is assumed that the students rated their peers' performances higher without distinguishing among their peers.

2 Research Aim 2

The second research aim was to find any differences in the self- and peer evaluations between the high- and low-proficiency learners. The results showed that the low-proficiency learners self-evaluated their hand and foot positions higher than the high-proficiency learners did for the first presentation. In addition, the low-proficiency learners tended to self-evaluate their performances more leniently than the high-proficiency learners. This suggests that learners who are not confident in their English proficiency tend to overestimate aspects of performance based on visual information, such as hand and foot positions.

The results of qualitative analysis indicated that the high-proficiency learners evaluated their own performances strictly, since they provided more negative feedback on their own performances than the low-proficiency learners did. Moreover, while both the low- and high-proficiency learners

---

15) Falchikov, *op. cit.* p.159; Falchikov, "Peer feedback marking: Developing peer assessment." *Innovations in Education and Training International,* 32（2）, 1995, p. 184; Peng, *Peer assessment of oral presentation in an EFL context.*（Doctoral dissertation, Indiana University, 2009）.（UMI No. 3380148）, 2009, p. 145.

16) Shimura, *loc. cit.*, p. 104.

reported that they felt nervous during their first presentation, there were no such comments for the second presentation, which implies that the students' first presentation and rehearsals helped them reduce the stress they felt speaking in front of the class; eventually they became used to giving oral presentations and felt less anxious about them.

3　Factors That Affected the Evaluations

Additionally, it is important to point out that the students had a rehearsal before each presentation in this study. However, they may not have always evaluated their rehearsals and presentations similarly because the following factors might have affected their attitude toward self- and peer evaluation. The first factor is that the students were "learners." As explained earlier, they memorized only half the speech for the rehearsals, but they memorized the whole speech for the presentation. Moreover, their rehearsal scores were not included in their grades, while their presentation scores were. Accordingly, they probably did not practice as much for their rehearsals as they did for their presentations. The second factor is the students' role as "evaluators." They knew that the rehearsals were just for a practice and that the teacher would not include the peer assessment score in their grades. In the study, peer evaluation was carried out only in the rehearsals; therefore, it is not clear whether the learners evaluated their peers more strictly in the presentation than the rehearsal. As these factors might have influenced the students' self- and peer evaluations, it is reasonable to conclude that future research would benefit from studying whether rehearsals are effective at improving learners' performances.

4　Limitations

The limitations of this study should be pointed out. First, the number of participants was not sufficient for the results to be generalized. Therefore, it is necessary to have more learners in order to find out if student evaluation is effective in language learning. Second, while video recording was used to assess learners' performance in this study, its effect on evaluation has in fact not been discussed here. Future research should focus on examining how video recording can contribute to develop learners' language skills as well as oral presentation skills. Third, although effects of training were demonstrated in the previous study, this study did not train students as evaluators due to the time constraint. It will be worthwhile, however, to examine self-evaluation after students have received sufficient training for it. Finally, this study did not implement peer evaluation for the presentations due to time constraints. It would surely be meaningful for students to evaluate their peers' final performances and for researchers to analyze how the performances change over time.

Ⅵ Conclusions

In this study, we attempted to show how Japanese EFL learners evaluated their own and their peers' performances using video recordings, and how high- and low-proficiency learners differed in their evaluations of video-recorded performances. The results confirmed that students take their social relationship into account when evaluating their peers' performances, as asserted by previous researchers, and that high- and low-proficiency learners evaluate their own performances differently from each other. Although the paper did not fully discuss the effects of using video recordings in self- and peer evaluation, the use of a video camera in the classroom has been shown to be an effective way of facilitating learners' reflection on their improvement in the target language, as it allows them to compare their previous and current performances. It is suggested that a teacher can use a video camera to record learners' performance in the target language not only once but several times. Learners can review a series of video recordings made throughout the course to observe how their target language and presentation skills have improved.

Finally, various factors were involved in improving learners' performance in this study. Watching their own rehearsal on video surely encouraged them to prepare for the final presentation and work to avoid lowering their academic grade for the course. No matter what other factors were present, it is clear that with the combination of peer evaluation and self-evaluation, video recordings can be a tool for learners to engage in reflection and thereby make their foreign language learning successful in the classroom.

Note

This is a revised and expanded version of the papers presented by the first author at the following two conferences: The Japan Association for Language Teaching（JALT）Annual Conference at Kobe Convention Center, Portopia, Kobe, October, 25-28, 2013, under the title "Video Recording as a Learner Assessment Tool," and the Globalization and Localization in Computer-Assisted Language Learning（GLoCALL）Conference at the University of Danang, Vietnam, November 28-30, 2013, under the title "Peer Assessment through Video Recording."

Appendix

		1	2	3	4	5
		Strongly disagree				Strongly agree
1	Hand position was appropriate.					
2	Foot position was appropriate.					
3	Posture was appropriate.					
4	I used eye contact.					
5	Speaking speed was appropriate.					
6	Speaking volume was appropriate.					
7	Pausing was appropriate.					
8	I used intonation.					
9	I used word stress.					
10	I paid attention to my pronunciation.					