The European Language Portfolio, Moodle, and Peer Evaluation: A Simplified Approach

Sergio Mazzarelli

[Abstract]

This paper reports the results obtained by taking advantage of the European Language Portfolio, Moodle, and peer evaluation activities in a listening course with an important presentation component. Previous work by the same writer had shown positive results from the use of these tools in two listening and speaking courses. However, those courses met twice a week, which is unusual in Japan, so the amount of time that could be dedicated to such activities was larger than most teachers have at their disposal. On the other hand, the present article details the steps taken to make it possible to integrate Portfolio, Moodle, and peer evaluation in a course that met only once a week. This was achieved by giving students clear directions for Portfolio use and by developing a simplified evaluation form for peer evaluation. The paper shows that the directions worked and that the simplified form was a viable option. An end-of-term survey demonstrated considerable satisfaction about the course among the students.

Introduction

This writer previously reported on his experiences in using Moodle’s Workshop module and a localized version of the European Language Portfolio in two presentation-based English courses he taught at a Japanese women’s university.¹ In particular, he described how Moodle can be employed to let students peer evaluate their in-class presentations (Mazzarelli, 2015). On those occasions, the time available for Portfolio-related activities was considerable because the course met twice a week and out-of-class conferencing with the students was also possible. Moreover, in the peer evaluation of the presentations, numerous aspects were evaluated one by one, so that to complete this
kind of activity the students needed to review video recordings of the presentations. In the end, the outcomes of the courses appeared to repay all the efforts, but there was a need to confirm such results, explore alternative evaluation strategies for peer evaluation, and investigate how to use the Portfolio, Moodle, and peer evaluation in courses that meet only once a week, as most do at Japanese universities. The present article provides an account of experimentation that was undertaken for these very purposes at the same university in the second term of the 2015 academic year.

Because the pedagogical reasons that recommend the use of the European Language Portfolio, Moodle, and peer evaluation were already explained in another article (Mazzarelli, 2015, pp. 106-107), they will not be revisited here. Suffice it to say that a growing number of teachers throughout the world are finding the European Language Portfolio and peer evaluation to be useful to motivate students and help them to become self-regulating learners. As for Moodle, its use for self- and peer evaluation simplifies procedures that can be time consuming and impractical if carried out on paper.

The Context

The present study was carried out in a second-year listening course which had a strong presentation component but in which there was limited time to dedicate to Portfolio activities and peer evaluation. There were nine students, all of whom had used the Portfolio before. Eight of them had previously used the Moodle Workshop module to self- and peer evaluate presentations through video recordings (Mazzarelli, 2015), while the remaining student had used Moodle but not for self- or peer evaluation. In the previous year, all the students were considered to have achieved CEFR B1 level in spoken production. Their listening ability, roughly estimated from their TOEIC score, ranged from CEFR B1 to B2, with three having achieved the higher level.

As is generally the case in Japan, the course met just once a week for fifteen weeks. During the majority of the lessons, the students listened to recordings of academic-style mini-lectures, took notes, answered comprehension questions, and discussed the topics of the lectures. Key vocabulary was assigned as homework. In-class listening comprehension quizzes were administered from time to time. The textbook, Contemporary Topics 1 (Solórzano & Frazier, 2009) paid
particular attention to techniques and expressions that are generally used to organize lectures.

Moodle was regularly used in the course to communicate with the students. In addition, the students were assigned to watch online lectures from the TED website (https://www.ted.com) and submit summaries of them along with personal comments. The topics of the lectures were chosen so as to be related to those of the textbook. Moodle was used to provide links to the lectures, let the students type their summaries and comments, and provide prompt feedback about those summaries and comments.

Since the textbook had twelve units and each unit took a week to cover, only three classes were available for presentations and peer evaluation activities. Moreover, some time was necessary to introduce learning targets and other Portfolio-related activities, although the students’ familiarity with the Portfolio certainly helped on this front. Finally, since students at Japanese universities take many courses, care had to be exercised not to overburden them with homework.

In the first week, the students were shown the following targets for them to achieve through the course:

1. I can follow the essentials of lectures and take notes that I can use later.
2. I can give a logically organized presentation on a topic derived from the textbook. Like the lecturers in the textbook, I can highlight significant points and provide relevant supporting detail.

These derived from long-term Common European Framework B2 goals listed in the Portfolio Appendix, which were also brought to the students’ attention. For example, one of the goals for listening is “I can follow the essentials of lectures, talks, and reports and other forms of academic or professional presentation in my field,” while one of the spoken production goals is “I can give a clear, systematically developed presentation on a topic in my field, with highlighting of significant points and relevant supporting detail.”

Since the Portfolio did not include note-taking goals, the students were shown the following one, derived from the Common European Framework by changing the third-person descriptor found there into a first-person descriptor: “I can understand a clearly structured lecture on
a familiar subject, and can take notes on points which strike me as important, even though I tend to concentrate on the words themselves and therefore to miss some information” (Council of Europe, 2001, p. 96). By looking at the long-term goals, the students were reminded that the course targets were not arbitrary. On the contrary, they were designed to help them achieve long-term goals based on validated proficiency descriptors.

The students entered the course targets in Portfolio target forms so as to maintain constant awareness of the purpose of all activities within the course. At the end of the course, they were supposed to look again at the form with target (1) above and decide whether or not they had achieved their target, and why. The form with target (2) above would later be complemented by two more target forms—one for each presentation—that students would use to record the outcome of their self-reflection.

The target forms ask students to indicate how they intend to achieve their target. For the first target, the following statement was suggested to them: “I will take notes when I listen to the lectures. I will use my notes about the lectures played in class to answer questions in the textbook. I will use my notes about the online lectures assigned as homework to prepare summaries.” For the second target, they were directed to write something like, “I will study the strategies described in the ‘Focus Your Attention’ section in each chapter of the textbook. I will plan my presentation carefully. Then I will write a draft and revise it.”

Ideally, in keeping with the spirit of autonomy the Portfolio was designed to foster, students should determine each target, write the way to achieve it, and work out peer assessment criteria all by themselves, but this would have required an amount of class time that was simply unavailable. Nevertheless, this writer believes that clarifying the purpose of all class activities allows students to reflect on their learning and monitor their progress and, therefore, is a valuable achievement if time does not allow teachers to implement more full-fledged forms of autonomy.

In the seventh and twelfth week of the course, respectively, the students were required to deliver two presentations on topics stemming from the textbook: “Sleep Problems” for the first presentation and “Robots” for the second. The required length for both presentations was three minutes. For these presentations, they were asked to conduct research and employ the knowledge of organizational techniques they
had acquired from studying the mini-lectures covered in class. They were not supposed to use any slides. If they wished to show images, these were to be printed and attached to the blackboard with magnets. Such a presentation format was chosen to let students concentrate as much as possible on the spoken word. It should also be noted that the lectures in the textbook featured no slides.

As mentioned above, for each presentation, the students filled a Portfolio target form a few weeks in advance of the due date. In it, they indicated what they were trying to achieve, how they intended to achieve it, what materials they would need, how much time they were going to budget for preparation, and how they would know if they achieved their target.

The target for the first presentation was “I can give a clear, systematically developed presentation on a sleep problem, with highlighting of significant points and relevant supporting detail.” Exactly the same wording was used for the second presentation, except that the words “sleep problem” were replaced by “robot.” As for the way the students would know if they achieved their target, they were all directed to write something like “Through self-evaluation, peer evaluation, and teacher evaluation.”

Devising Evaluation Procedures

In the course of past experimentation conducted in different courses, this writer had let his students evaluate their presentations by using video recordings uploaded to Moodle. This approach was thought to be preferable for students who had never evaluated their peers’ work because of the considerable concentration needed to assess live presentations and because the number of discrete criteria employed was somewhat large. In two instances, seven evaluation criteria had been used, one rated on a 5-point scale and all the others on a 10-point scale. In a third instance, there had been fourteen criteria, one rated with a four-level rubric and all the others with a three-level rubric.

This time, however, it was decided to experiment with a simplified evaluation form, which would allow the students to grade the presentations on the spot. One reason for this was that it was thought desirable to let second-year students begin performing live evaluations, which would have been too challenging for many first-year students who took part in previous experiments. The most important reason, though,
was the desire to speed up the evaluation process. Since the class time available for the evaluation was limited, if videos had been used the students would have had to watch them from home via Moodle. Thus, the evaluation could not have begun until all the videos had been uploaded to Moodle. Moreover, as stated above, the students were expected to be very busy, so the chances that they would be late in finishing their evaluations were high. Because grades could not be calculated until every participant had submitted her evaluation, the timely completion of the activity could have been prevented by the failure of a single student to complete all her evaluations.

Instead, thanks to the simplified form, the evaluations were done in class. A three-minute interval was inserted between presentations to allow everyone to finish writing. Later, the class moved to a computer lab where the evaluations were typed into Moodle. This took no longer than thirty minutes, so even if a lab had not been available, it would not have been much of a burden for the students to enter the evaluations into Moodle in their own time.

Alternatively, it would have been possible to let the students type the assessment directly into Moodle if the class had been held in a computer room or if recourse to mobile phones had been made. However, computer rooms may not be the right environment for presentations because the monitors and other equipment often make it more difficult for the speaker to maintain eye-contact and more generally establish a rapport with the audience. As for the use of the students’ mobile phones for typing comments into Moodle, this was ruled out because it was feared that typing on small devices may be cumbersome and might adversely affect the quality of the comments. On the other hand, it was believed that having to type what they had previously handwritten would give the students a chance to check their spelling, fix their syntax, and make other adjustments if needed. In fact, they were told that it was also acceptable to write simple notes on paper and then expand them into sentences when typing into Moodle, although no one did that.

It must be mentioned that the presentations were actually recorded anyway. Although the footage was not uploaded to Moodle or otherwise viewed by students during the course, it could have been used by the teacher in case a student’s comment raised a point the teacher did not recall. Also, if no camera had been available and one or more students had been absent on the day of the presentations, those students would
not have been unable to evaluate presentations they had not seen. Finally, the teacher intended to give the students the option to receive videos of their own presentations after the course ended so that they could add them to their Portfolios.

The evaluation criteria were obtained by grouping characteristics that had been previously identified as constitutive of a satisfactory presentation (Mazzarelli, 2015, pp. 110, 119). The evaluation form looked like this:

A. Visual impression: Eye-contact, smile, posture, and gestures.

   0 1 2 3 4 5 6

B. Aural impression: volume, pronunciation, intonation, and fluency.

   0 1 2 3 4 5 6

C. Content: interest, depth, clarity, and organization.

   0 1 2 3 4 5 6

Total points: __________

Under “organization,” the students were to pay particular attention to the “highlighting of significant points and relevant supporting detail” featured in the target description of both presentations and derived from the Common European Framework (Council of Europe, 2001, p. 60). Four lines of space were provided for comments and suggestions at the bottom of the form. The students were also reminded to evaluate their own presentations.

**Evaluation of the Presentations**

All the presentations showed solid organization, with effective use of sentence connectors and support provided for all points. This was due to the study of the lectures in the textbook, but the students must also have benefited from instructions they received in writing classes. Such a positive result was even more remarkable in that this writer did not preview and give feedback on any drafts of the presentations.

Where the presenters differed was mainly in the degree of eye-contact and fluency, which were related to their ability to speak without looking at notes, in volume, which was related to confidence, and in
depth of thought, which was related to effort but perhaps also familiarity with speech and writing that appeal to logic rather than emotions.

In the second presentation, all presenters visibly improved their performance as regards the aspects that had been identified as in need of improvement in the peer evaluation activity, although some were more successful than others.

The comments the students attached to their evaluations were brief, but always perceptive and helpful. By collecting the paper evaluation forms and comparing them with what had been entered into Moodle, it was found that often the students had reworded their comments to make them clearer or more constructive. Therefore, it was confirmed that the pause for reflection occasioned by the need to enter the evaluation into Moodle had been beneficial.

The grades automatically calculated by Moodle were satisfactory enough to be used without modification in the case of the first presentation, and with a single adjustment in the case of the second, where Moodle was manually set to ignore a particularly low self-evaluation. In other words, the correlation between the ranking of presentations resulting from the average grades calculated by Moodle and the ranking of the same presentations resulting from the teachers’ grades was very good. Such grades appeared to be as usable as those obtained in a previous experiment that had involved a group of thirteen students plus their teacher evaluating presentations according to a larger set of discrete criteria and after viewing video recordings (Mazzarelli, 2015, pp. 110, 119).5

The most serious limit of student evaluations was that a few students gave themselves much lower grades than they gave to classmates of similar ability. This is a common problem in Japan, where people find it uncomfortable to express favorable self-evaluations, and its cultural roots make it difficult to solve. In fact, in this class it was not as much of an issue as in other classes this writer has taught.

After the grades for each presentation were ready, the students looked at them as well as the comments their peers and their teacher had provided, and then filled the last field of the target form, which asked them to decide whether they had achieved their target, but in which they had also been instructed to write how they could further improve their proficiency. This was a very important moment because it would show whether the Portfolio was effective in promoting self-reflection and enhancing motivation.
The Portfolio is the property of the student. Nevertheless, this writer arranged to review the presentation target forms for research purposes, while making clear to the students that no grading whatsoever would be done on the Portfolio.

The tone of the self-evaluations in the Portfolio was always positive and forward-looking. The students showed determination to improve and appreciation for the comments they had received. One explicitly said that she was “motivated” by the comments to work on specific areas. Occasionally, the comments also helped a student realize that an issue she had worried about was not a problem at all. When taken as a whole, peer comments managed to identify almost all the problems each presentation had, except for points of grammar and depth of thought, which were raised only by the teacher. Thus, thanks to the peer evaluation, supplemented by teacher evaluation, the Portfolio was confirmed to have fulfilled its role as a tool for self-reflection as well as a motivating agent.

In the case of the second presentation, systematic comparison of student self-evaluations on Moodle with those in the Portfolio showed that the latter were always superior to the former, mostly because they took into account all or nearly all the issues identified by classmates. In the case of the first presentation, the difference between the two kinds of self-evaluation was less pronounced. This was most likely due to the fact that the second time all the students had the comments of their peers and teacher in front of them as they wrote, while the first time many had based their reflections purely on their memory of such comments. The students had been asked to write in the Portfolio outside class, but most ended up writing in it while in class. On the second occasion, they were reminded that they could display the comments on their mobile phones and took advantage of this possibility. This underlines the importance of accommodating the needs of students with busy schedules in order to help them make the most of the Portfolio.

**Student Feedback about the Course**

An anonymous questionnaire was administered to the students in the penultimate class in order to learn their opinion about the course. Below are the first five multiple-answer questions, along with the number of students who chose each answer:
A. How much progress do you think you have made in English LISTENING thanks to this class?
   1. No progress 0
   2. A small progress 0
   3. A moderate progress 3
   4. A considerable progress 4
   5. A great progress 5

B. How much progress do you think you have made in English NOTE-TAKING thanks to this class?
   1. No progress 0
   2. A small progress 0
   3. A moderate progress 1
   4. A considerable progress 5
   5. A great progress 3

C. How much progress do you think you have made in English PRESENTATION skills thanks to this class?
   1. No progress 0
   2. A small progress 0
   3. A moderate progress 3
   4. A considerable progress 5
   5. A great progress 1

D. Have you achieved your first Portfolio target for this class? (“I can follow the essentials of lectures and take notes that I can use later.”)
   1. Not at all 0
   2. To a small degree 1
   3. To a moderate degree 2
   4. To a considerable degree 2
   5. To a great degree 4

E. Have you achieved your second Portfolio target for this class? (“I can give a logically organized presentation on a topic derived from the textbook. Like the lecturers in the textbook, I can highlight significant points and provide relevant supporting detail.”)
   1. Not at all 0
   2. To a small degree 1
   3. To a moderate degree 4
   4. To a considerable degree 3
   5. To a great degree 1
The above results show awareness of the undoubtable progress that the students made in English proficiency, coupled with the recognition that, in some cases, they had not yet fully achieved their targets, especially in presenting. That is correct if we look at the targets as including all the criteria used in the evaluation rather than simply the organizational aspects highlighted in the target statement. As stated before, the presentations were all satisfactory from an organizational point of view. Yet it could not be denied that a satisfactory presentation does not depend solely on organization.

Another group of five multiple-answer questions asked students to assess the usefulness of each component of the course towards the achievement of its goals. The results are listed below:

F. Was the textbook useful?
   1. Not at all 0
   2. To a small degree 0
   3. To a moderate degree 1
   4. To a considerable degree 5
   5. To a great degree 3

G. Were the Moodle listening assignments useful?
   1. Not at all 0
   2. To a small degree 0
   3. To a moderate degree 0
   4. To a considerable degree 4
   5. To a great degree 5

H. Was giving presentations useful?
   1. Not at all 0
   2. To a small degree 1
   3. To a moderate degree 0
   4. To a considerable degree 4
   5. To a great degree 4

I. Was receiving evaluations of your own presentations useful?
   1. Not at all 0
   2. To a small degree 0
   3. To a moderate degree 1
   4. To a considerable degree 2
   5. To a great degree 6
J. Was evaluating your classmates’ presentations useful?
   1. Not at all  0
   2. To a small degree  0
   3. To a moderate degree  2
   4. To a considerable degree  1
   5. To a great degree  6

K. Was evaluating your own presentations useful?
   1. Not at all  0
   2. To a small degree  1
   3. To a moderate degree  0
   4. To a considerable degree  5
   5. To a great degree  3

Overall, there seems to be satisfaction with all components of the course, but the peer evaluation (Questions I and J) saw the highest number of students choosing “To a great degree” as an answer. Therefore, the success of the peer evaluation cannot be doubted. A single student thought presentations and self-evaluation only useful to a small degree. Unfortunately, although there was a space for comments, she did not provide any that could clarify her answers, but it is worth bearing in mind that she still believed that receiving evaluations of her own presentations had been useful to a considerable degree.

The questionnaire then asked students whether they would prefer evaluating presentations on the spot or by reviewing video recordings, and why. Here opinions were split, with five students favoring video recordings and four choosing evaluation on the spot. Of course, the fact that all but one of the students involved in this study had been accustomed during the previous year to evaluating presentations with the help of video may have had a bearing on the result.

Among the reasons provided by those who preferred recordings were that the limited time made it difficult to write complete evaluations, that without videos one cannot see one’s own performance, and that they wanted to watch the presentations again. Those who preferred the live evaluation believed that the presenter’s live voice is clearer than a recording, that a live event is different from what is captured on camera, or that it is simply easier to evaluate a live presentation.

The final comments appended by some students to the questionnaire were all positive about the course in general and the presentations in particular. Let a few examples suffice. One said, in part,
“I like this class because it is useful to improve my listening and speaking skills.” Another stated, “I thought presentation in the class was very good to get confident to speak in front of people.” A third remarked, “I felt my improvement as I do my presentation again and again.”

**Conclusion**

It is hoped that this paper has confirmed how the use of the Portfolio helps students in language classes and that its combination with peer evaluation and Moodle is effective in improving students’ presentation skills.

Moreover, although more rigorous evaluation criteria may be desirable in some contexts, it seems that a simple evaluation form like the one described above can yield satisfactory results while allowing CEFR B1 and B2 learners to perform live evaluations of presentations and minimizing the amount of time that teacher and students need to complete peer evaluation activities. Although five students stated that, given the choice, they would prefer to evaluate from video, this writer did not detect any substantial difference between the quality of the evaluations assisted by video described in a previous paper (Mazzarelli, 2015, pp. 112, 114-115, 119-120) and that of the live evaluations described in the present paper. However, this could be because all the students had some experience in evaluating presentations. It remains to be seen whether students who had no such experience would obtain equally good results with and without the help of video. Further research would be necessary to ascertain that. Finally, video would be very useful if not indispensable were any students to be absent on presentation days, especially when total course enrolment is small.

One more issue also remains to be investigated. Given the superiority of the Portfolio self-evaluation, conducted by students after reading comments by teachers and peers, it is worth considering whether the initial self-evaluation, composed just after the presentation, could be dispensed with. The written part does no harm, but the grades the students give themselves tend to be, as stated above, rather low. Especially in small classes, very low self-evaluation grades may significantly lower the final grades Moodle calculates even when they do not alter the ranking of presentations because everyone is modest. The teacher is then forced to adjust Moodle’s calculations.

The purpose of the initial self-evaluation was to allow students to
participate in the grading of their own presentations, but this could also be accomplished by asking them to grade themselves after they have completed the self-evaluation in the Portfolio. The average grade calculated by Moodle out of teacher and peer evaluations could be compounded with the grade the student would later give herself. Hopefully, the higher quality of the reflection carried out in the Portfolio would result in more accurate grades.

Humility is widely recognized as a traditional cultural trait of the Japanese people. It could be that having to give themselves grades is perceived by many, though not all, Japanese students as a context in which it is particularly difficult express a favorable self-evaluation. This difficulty may then be removed once other participants’ opinions of one’s performance are known. Of course, it is also possible that societal norms prescribing humility are so strong that some students may give themselves low grades even after they received good ones from their teacher and peers. The form of discursive self-evaluation required by the Portfolio, which—it must be remembered—is written only for oneself, may be the only context in which all students feel free to express more favorable opinions about their own performance. If so, perhaps asking students to participate in the actual grading of their own presentation may be counterproductive for cultural reasons. Since the self-evaluation in the Portfolio is effective in building their motivation and helping them progress, teachers should be satisfied with that and should not require their students to grade themselves—no matter how empowering for the students this sounds.

This writer hopes to be able to clarify these issues in future studies.

Notes

1 The Portfolio used was a slightly adapted version of the Language Portfolio for Japanese University (Framework and Language Portfolio SIG, 2009).
2 Tablet computers and a wireless LAN were not available.
3 The criterion of length, used the previous year, was replaced by depth. It was felt that a systematically developed presentation based on research should be judged by depth of content. A shorter presentation would most likely lose points under depth.
4 The acceptability of the ranking was evident without the need for any statistical analysis. Nevertheless, the grades calculated by Moodle and those given by the teacher were tabulated and the Spearman coefficient calculated. For the first presentation, it was .92 ($p < .001$). In the case of the second presentation, the Spearman coefficient was .81 ($p > .001$) before the adjustment and .83 ($p < .001$) afterwards. As in a previous paper (Mazzarelli, 2015, p. 7), statistics are here provided in notes for the reader’s convenience. This writer does not wish to give undue significance to calculations
performed on such small samples or claim any expertise in statistics.

On that occasion, with fourteen raters, comparing Moodle average grades with the teacher’s grades yielded a Spearman coefficient of .93 ($p < .001$) for a presentation evaluated with seven discrete criteria and .90 ($p < .001$) for one evaluated with fourteen discrete criteria. However, it must also be noted that the same criteria that had yielded a .93 ($p < .001$) coefficient with fourteen raters produced a much lower .71 ($p = .006$) when another presentation was evaluated by just seven raters (Mazzarelli, 2015, p. 7).

References


Received February 1, 2016.