



Implementing Peer Assessment for Optimal Effects: Learners' Voice

Jeongsoon Joh
Konkuk University

The present study attempted to figure out the optimal conditions for peer assessment to be successful from learners' perspectives. Participants were twenty Korean college students majoring in TEFL, who practiced peer assessment in a content course, completed a pre-, and a post-survey about their perception of peer assessment, and responded to an interview after the semester. Several meaningful themes emerged from their voice in relation to the fundamental elements to be considered in designing and implementing peer assessment. First, working in group to perform the task to be peer-assessed proved to be highly effective in that it could offer much more opportunities to learn from each other first within the group members and then across different groups. Second, face-to-face peer assessment in class was generally preferred over on-line out-of-class, with the immediacy of feedback and other benefits interwoven with it. Third, the more experience of peer assessment in a curriculum and more frequent peer assessing activities in a course was often associated with the learners' perceived learning achievement. Fourth, the teacher's intervention could be a necessary condition to make the benefits of peer assessment greater. Implications for teachers and researchers are stated with limitations of the study.

Keywords: peer assessment, group work, mode of feedback, frequency, teacher intervention

Introduction

Peer assessment, defined as “an arrangement for classmates to consider the level, value, or worth of the products or outcomes of learning of their equal-status peers (Topping, 2013, p. 395)”, has attracted many researchers and educators for the last few decades as a promising form of alternative assessment. Conceived of fairly formative and learner-centered in nature, peer assessment of various forms has frequently been employed in diverse educational settings, particularly often in L2 writing classrooms (e.g., Chong, 2017; Hu & Lam, 2010; Lee, 2015; Nawas, 2020; Park, 2018) but also in many other subject matter courses including teacher education (e.g., Kissau & King, 2015; Patton & Marty-Snyder, 2014; Sahin-Taskin, 2018). Related previous studies have reported its beneficial impact for enhanced learning via deeper understanding of the learning materials, and highlighted its long-term potential, such as boosting learner autonomy, cultivating critical thinking and communication skills, which are highly valuable both in and out of school (Azarnoosh, 2013; Falchikov, 2007; Topping, 2018).

At the same time, questions have been raised about its reliability and/or validity, largely associated with students' deficiency in the ability to assess their peers either from lack of sufficient training in assessment or from their limited knowledge in the relevant area. Such concerns combined with many other practical or contextual reasons (e.g., extra time and efforts required on the part of teachers, students' reluctance to assessing and being assessed by peers for various reasons including concerns about subjectivity of



assessment) seem to have led some educators and researchers rather reserved about implementing peer assessment in classrooms, not very sure of its success (cf. Adachi, Tai, & Dawson, 2018a).

In exploring peer assessment (frequently shortened as PA henceforth, for convenience' sake), some of the previous studies attempted to compare the teacher-, and the peer assessment (e.g., grading or scoring by the teacher vs. that by students) as a means to verify PA's reliability and/or validity (e.g., Babaii & Adeh, 2019; Ganji, 2009; Ruegg, 2015), while others were more interested in figuring out substantial effects of peer feedback on learning (e.g., Berggren, 2015; Birjandi & Tamjid, 2012; Paulus, 1999), or focused on examining students' behaviors while engaging in giving or receiving peer feedback (e.g., Zheng, 2012; Zhu, 2001). Culture-specific or affective variable issues associated with PA were occasionally discussed as well (e.g., Nelson & Carson, 2006; Yu & Lee, 2015; Zhu & Mitchell, 2012).

Although research in peer assessment has steadily been conducted with various foci, revealing both positive and negative aspects about employing PA as a major classroom procedure, there remains much to be further investigated in pursuit of the precise nature of PA, i.e., exactly how PA works in many possible different instructional settings. Given the variety of contexts where PA has been employed and researched (e.g., the nature of courses, the type, manner, or frequency of PA activities, the characteristics of students, the venue of research, etc.), more detailed and refined investigation situated at diverse learning environments seems required. That is, more rigorous investigations would be necessary that aim at defining specific instructional conditions or designs for PA to bring out maximum benefits for the students situated at different learning contexts. Accumulated findings from such empirical studies should provide teachers more pragmatic and useful information about how to successfully design and put into practice their own PA procedures that best fit their learners and their own teaching environments, with optimal benefits for both parties.

Despite considerable body of PA research over the last three decades, few approached this issue directly and empirically. Adachi et al. (2018a, 2018b) exceptionally attended to this issue by interviewing in depth 13 teachers who had experience of practicing PA in classrooms, but the voice of learners themselves is still rarely heard. Although there have been studies that reported the learners' perception about PA, they usually focused more on the students' perception about the effect of PA or their appraisal of peers' ability as assessors, thus rather limited in elucidating how learners think of and evaluate the specific aspects of the design or procedure by which PA is arranged and executed in their own classroom. Listening to students' voice would be undoubtedly critical for successful implementation of PA when they are the very agents of peer assessment who assess and are assessed by peers, even though they do not design PA for themselves in most cases.

Given this, the current study attempts to elicit learners' perspectives about under what conditions (i.e., with what implementational procedures by which PA is practiced) they think PA most benefits their learning, or under what circumstances they perceive of it less or not very beneficial, by involving a group of college students who have engaged in peer assessment in content courses. The learners' vivid voice from their own experience would provide substantial clues to how to effectively design and utilize PA both for the teachers' and learners' satisfaction. What is special about the current study is that the participants are the teacher-trainees. For them, being engaged in peer assessment is not simply a learning experience, but it could simultaneously work as a training opportunity in that they are quite likely to design and orchestrate PA procedures on their own in near future, now as a teacher. In what follows is a summary of literature review.

Review of Literature

Peer Assessment: Benefits and Concerns

A great number of previous studies (both theoretical and empirical, classroom-based studies) have mentioned beneficial effects of PA on learning. Interactive in nature, it can make the classroom more

learner-centered (Azarnoosh, 2013), and can even be more beneficial than teacher feedback in light of the quantity and immediateness, although likely to be less accurate than teachers' (Topping, 2013). Moreover, PA has been viewed to boost reflective, self-directed learning (e.g., Falchikov, 2007; Reinholz, 2016; Roscoe & Chi, 2007; Saito, 2008). Critical thinking ability, communication and problem-solving skills, and evaluative judgement skills can also be developed while learners give feedback to their peers, evaluate them by giving them grade or score, and react to peers' feedback comments (e.g., Adachi et al., 2018a; Boud & Falchikov, 2006; Yang, Badger, & Yu, 2006).

Several studies compared PA with teacher assessment in highlighting the effectiveness of PA in learning gain. According to Ganji (2009) peer feedback was more beneficial in writing improvement than either self-, or teacher feedback. Similarly, Ruegg (2015) found that teacher feedback worked better than peer feedback only in the area of grammar in an EFL writing classroom. Sippel and Jackson (2015) also found peer feedback was more effective in the acquisition of German grammatical features, although the teacher's corrective feedback was also beneficial. Murillo-Zamorano (2018) compared the effect of teacher-, and peer feedback on the improvement of oral presentations in business-majoring college students and found peer feedback was more effective than teacher feedback. More recently, Babaii and Adeh (2019) reported that the greatest performance of EFL writing was found in the paired peer feedback condition over either teacher feedback-, or group peer feedback condition.

Meanwhile, concerns about incorporating PA into classroom procedures have also been expressed in previous studies. For example, Chang, Tseng, Chou and Chen (2011) was skeptical about reliability and validity of PA, for they found a significant difference between peer assessment and the final course grade. Questions about the effectiveness of PA were also raised in some studies, where ESL learners were found to considerably prefer teacher feedback over peer feedback (e.g., Cheng & Warren, 2005; Tsuil & Ng, 2000). In addition, peers' subjectivity in assessing other peers (e.g., friendship bias) was mentioned as a possible factor that might threaten the validity of PA (Alfallay, 2004; Azarnoosh, 2013; Buchanan, 2004; Morahan-Martin; 1996). More recently, Adachi, Tai, and Dawson (2018a) observed that personal power relationships within the peer group and the students' perception of their peers' ability to assess could make difficult successful implementation of PA.

Learners' Perception of PA

Inquiries of learners' perception of PA have revealed mixed results. Zhang (1995) found that his secondary school English learners in Hong Kong strongly preferred teacher feedback over peer feedback. By contrast, Lee (2015) reported that Chinese secondary school English learners showed very positive perception of peer feedback. Tahir (2012) also found that ESL college students were contented with peers' feedback comments both in terms of quality and affective aspects. Meanwhile, Zou, Schunn, Wang, and Zhang (2017) observed a rather complicated relationship between the learners' attitudes about PA and the degree of their participation in it, reporting that their doubts about the procedural rationality of peer assessment was negatively associated with their participation while generally positive perception about the PA's benefit did not necessarily result in higher participation. Park (2018) found that Korean EFL college students favored and used more of teacher feedback than peer feedback in their writing, although they recognized the unique strengths of peer feedback. This variation in findings from seemingly comparable research contexts imply that multiple factors could contribute to the shaping of the learners' perception of PA.

Related to this, several studies tried to figure out the possible factors that could affect the learners' approaches to, and perception of, PA. Roskams (1999) attended to the role of cultural context where learners have been exposed in the shaping of their perception of PA, reasoning that the quantity of peer interaction among the participants and their reserved attitudes towards PA seem to have rooted in their Asian cultural norms. Yu and Hu (2017) similarly pointed to socio-educational context as well as personal factors (e.g., beliefs and experiences) as possible factors that could affect one's approach to PA. Meanwhile, Allen and Katayama (2016) showed how peer assessment practices (e.g., type and quantity of

feedback) could be influenced by learners' perception of their peers' ability as well as of their own. Recently Joh (2019) reported that competitiveness among the peers and their desire to keep face played a key role in Korean college learners' active participation in PA activities, which contributed to their enhanced learning and very positive perception about the beneficial impact of PA.

Frameworks for Designing and Implementing PA

Topping's (1998) typology of PA is regarded as the first systematic attempt to define and classify the essential elements to be considered in researching and implementing PA. Gielen, Dochy, and Onghena (2011) expanded Topping's (1998) frame by introducing more elements and re-organized them into 5 fundamental categories: (1) decisions concerning the use of peer assessment (e.g., subject area, timing, assessment type); (2) link between PA and other elements (e.g., relationship of PA to other assessment elements); (3) interaction between peers (e.g., anonymity of feedback giver, type or mode of feedback); (4) composition of assessment group (e.g., assessment by group or individual, manner of peer matching); and (5) management of assessment procedure (e.g., moderation of feedback, training prior to assessing).

Similarly, Topping (2013) offered a kind of checklist which would make easier the comparisons across PA studies conducted under different contexts, also reminding us of the multi-dimensional aspects involved in PA practices. The checklist consists of 55 entries subsumed under 6 categories: (1) the focus of the peer assessment (e.g., students' metacognitive gains, teacher's time saving); (2) the nature of the sample (e.g., previous experience, pre-existing skills); (3) the matching and contact procedure (e.g., selective or random matching); (4) criteria and training (e.g., clear assessment criteria); (5) variations in feedback (e.g., qualitative or quantitative feedback); and (6) final issues (e.g., rewards, evaluation of peer assessment).

More recently Topping (2018) presented 43 dimensions, a slight revision of Topping (2013), against each of which different options are available, e.g., either formative or summative (or both), either with or without rubrics, either anonymous or non-anonymous, either individuals, pairs, or in groups, etc. Both Gielen et al. (2011) and Topping (2013, 2018) clearly demonstrate how much variation is possible in designing and implementing PA, requesting careful consideration in designing, implementing, or researching PA. Azarnoosh (2013), too, observed that the efficacy of PA might depend on a variety of factors even though its impact on learning is seemingly considerable. Related to this, Adachi et al. (2018a, b) suggested that the two most crucial elements for successful implementation of PA should be motivating students to willingly participate in PA and helping them develop skills in giving feedback and emphasized the need for further inquiry aimed at developing guidelines and frameworks for practicing PA for its successful implementation.

Given the multiple dimensions to consider when designing PA as well as the variety of contexts where PA could be introduced, more investigation seems necessary that involves teachers and learners who have experience of practicing PA. Reflecting their voices into the designing, implementing, or researching PA will greatly contribute to the success of this alternative approach to assessment of great potential. With this research orientation in mind, and as an attempt to investigate the optimal conditions for PA to be successful, the present study focuses on learners' perspectives on this issue, which has rarely been examined in previous studies. More specifically, the present study intends to answer the following questions, i.e., what aspects or dimensions of PA practices influence the learners' evaluation of its impact on their learning, and exactly how they do.

Method

The Context and Participants

The current study was conducted in two independent intact classes for TEFL majoring students at a

university located in Seoul, South Korea, during a regular semester. The two courses were *TEFL Methods* (a course for sophomores) and *English Language Assessment* (a course for juniors), where peer assessment was employed as a major learning activity. Many of the students in the junior course had already practiced peer assessment in one or two courses they previously took as TEFL majors, while most students in the sophomore course were relatively less familiar with such peer assessment activity as implemented in the current study, although some of them had experience of assessing peers in other classes, mostly liberal arts courses.

All the students enrolled in the two courses (42 students in total) responded to a survey designed to examine their perception of PA's impact on their learning at the beginning of the semester, performed PA activities during the course as instructed by the teacher, and responded again to the same survey at the end of the coursework. After the semester was over, all the students were invited to a retrospective interview about their experience of PA, and 20 students accepted the invitation. These 20 students (13 juniors and 7 sophomores) were finally counted as the participants of the current study, and their responses to the interview questions comprised the main data sources to be analyzed and discussed.

Peer Assessment in the Current Study

In both classes peer assessment was composed of giving feedback comments to the peers' performance or product and grading it using the evaluation criteria provided by the instructor. The learning output to be peer-assessed was, however, different in the two courses according to the nature of respective course: *TEFL Methods* required the students to perform a micro-teaching presentation in the class, while *English Language Assessment* required constructing items or tasks to assess English language ability.

In both courses, the students worked in group to prepare and present their product. Each group (of 4 to 5) collaborated to prepare the given task and to present its output in the class, but students worked individually in giving feedback comments and evaluating other groups' products. The presenting group was instructed to provide the rest of the class with a copy of materials they created, such as a lesson plan and worksheets, or test items. The peer assessment session was composed of two phases: peer discussion phase where the observing students (i.e., those students other than the presenters) gave comments or asked questions about the presentation, and the presenters responded to, or explained their rationales about, those comments or questions; and peer evaluation phase where the observers filled out a peer evaluation form based on the discussion.

The peer discussion phase fulfilled multiple functions such as offering the students the opportunities to better understand the presenters' rationale for implementing the micro-teaching or in designing specific assessment tasks, working as a window for sharing useful learning ideas among the peers, and forming the base of evaluating the peers' performance. The participants were strongly encouraged to ask questions and give feedback comments before filling out the evaluation form. The instructor completed the peer assessment session by providing her own feedback comments both on the students' presentations and on the feedback comments given by the observing students. Sometimes the instructor gave complementing remarks regarding insightful peer feedback comments, which could probably have provided the students with opportunities to learn, or train in, how to assess.¹

In what follows is given a brief description of the structure of the two courses along with how PA was incorporated into instructional procedures in each course. In the course *TEFL Methods*, most of the coursework was devoted to the lectures on the well-known language teaching methods (e.g., Audio-Lingual Method, Silent Way, Task-based Language Teaching), usually accompanied by video-clips demonstrating a specific method where available or the instructor's demonstration otherwise. Then the last 3 weeks before the final exam were allocated for the micro-teaching presentation. Before the micro-teaching session began, the students were trained for constructing a formal lesson plan on their own.

In the course *English language assessment*, several semester-initial weeks were devoted to learning the

¹ This reasoning is based on the analysis of the interview protocols.

important terms and concepts, and fundamental principles and procedures in language assessment, and then the later weeks were set for constructing assessment tasks largely by skills (e.g., listening, speaking, etc.). After a book chapter on a skill was covered at the class, the students created in group the assessment items/tasks for the skill covered in the class, presented their output to the whole class, and peer assessment session followed according to the procedure already described in a previous paragraph.

This course had several aspects different from the course *TEFL Methods* regarding the elements of peer assessment activities. First, the students had to present their product more frequently (as many as five times (i.e., for each of the 4 skills and grammar/vocabulary, according to the chapters of the main textbook used for this course), compared with just once in *TEFL Methods*, which means students in *English Language Assessment* were assessed by peers more often than those in *TEFL Methods* and assessed peers far more frequently than that. Second, about half of the peer assessment sessions occurred out of class, on-line, using the E-campus system of the university they belonged to. For on-line peer assessment, the students uploaded on E-campus the assessment items/tasks they constructed, gave feedback comments to other groups' products, and responded to the feedback comments given to them, anytime within the time limit designated by the instructor. By contrast, in the course *TEFL Methods* all the peer assessment procedures were face-to-face, in-class, immediately after the presentation. This combination of on-, and off-line peer assessment in the course *English Language Assessment* was largely because time was not enough to cover all the class procedures and peer assessment sessions inside the classroom within the given time frame. In short, the two courses were a little different in terms of the frequency, and the mode (or channel) of peer assessment.

Materials²

Surveys

The pre-survey consisted of two sections: (1) Questions about personal information such as gender, grade year, personality type (i.e., introverted vs. extroverted), cognitive style, prior experience of peer assessment; (2) Questions about how the students perceive of the impact of peer assessment on their learning. In the second section, the participants were asked to give their perspectives by choosing one of the 5-point Likert-scale (from “*Strongly agree*,” to “*Never agree*”) about the following 4 statements: (1) “Peer assessment activities help better understand the learning materials.”; (2) “Peer assessment activities help enhance my motivation to learn.”; (3) “Peer assessment activities help prepare for, and actively participate in, the class.”; (4) “Peer assessment activities help improve learning achievement.”

For the purpose of examining whether the participants' perception changed over the semester, a post-survey was designed, which was basically the same as the pre-survey. The post-survey did not include the questions for demographic information, which information was already collected from the pre-survey, and the 4 statements about perception were re-phrased in the past tense, e.g., “Peer assessment activities *helped* (me) better understand the course materials,” for the participants would respond retrospectively at the time of the post-survey.

Peer evaluation form

This form was composed of two sections: peer feedback comments or questions to ask the assessed, and the 5 criteria of evaluation which were written in question forms. For the course *TEFL Methods*, the criteria included teaching objectives (“Were the teaching objectives clearly stated?”, “Were the objectives adequately accomplished?”), quality of the teaching materials (“Were the teaching materials well prepared?”, “Were the materials appropriately incorporated into the teaching procedures?”),

² The current study used a similar set of materials to those in an earlier study with different research purposes by the same author.

correspondence between teaching-learning activities and the learner characteristics (“Did the employed activities/materials well match the target learners’ characteristics, e.g., proficiency level, age?”), and creativity and usefulness of format and contents (“Were the contents and format used for teaching creative and useful?”).

For the course *English Language Assessment*, the criteria covered the fundamental principles of assessment such as practicality, reliability, validity, and washback, stated in the form of questions such as “Were the items or tasks appropriate for eliciting the knowledge or ability they attempted to assess?”; “Were the items possible and easy to implement in real classrooms?”; “Did the items or tasks contribute to accomplishing teaching objectives?”; “Did the items or tasks match the level of the target students?” For each criterion, the participants marked on one of the 5 Likert scales, from ‘5’ for ‘Excellent’ to ‘1’ for ‘Very poor.’

Interview questions

Questions for the interview were slightly different across interviewees, as the individual interview was largely based on the participants’ responses to the pre-, and the post-survey. The main purpose of the interview was to elicit as much as possible the participants’ thoughts that the surveys could not have sufficiently visualized, and the focus was on the reasons for the interviewee to reach a specific response. Example interview questions include the following: “You responded more favorably at the end of the semester to the questionnaire item about the benefits of PA in enhancing learning achievement than at the beginning. Is there any special reason for such change in perception?”; “You showed consistently positive attitudes about the impact of PA on your motivation to learn. What makes you so sure that PA boosted your learning motivation?”

Procedure

The students completed a pre-survey at the first class meeting. During the coursework, peer assessment activities were conducted as described in the previous section. At the end of the semester, a post-survey was administered as a means to verify whether their perception about peer assessment changed, and how, if ever. After the semester was over, all the students were contacted for in-depth individual interviews, and 20 students accepted the invitation, as described earlier in this chapter.

Before the interview, each interviewee’s responses to the surveys were meticulously examined to specify their perception about PA and its transition over time, if any. Then, questions were asked of each interviewee about the reasons for their specific perception of PA at each time, and those for any change in case their perception was different across the times of being surveyed. Depending on the individual interviewee’s responses, additional probing questions were asked to elicit more meaningful information about the factors underlying their perception. It took about 35 to 40 minutes for interviewing an individual, and all the interviews were audio-recorded using a high-quality recording equipment for later transcription.

Data Analysis

As the number of the participants was not large enough for running inferential statistics, descriptive and qualitative analysis of the collected data was carried out. First, the participants’ responses to the surveys were summarized for each interviewee to see whether and how his/her perception changed over time. Then the recorded interview protocols were transcribed and combined with the summarized survey data for each interviewee, which made it more easily visible how the students perceived of the various aspects of peer assessment they practiced in the class, and what seemingly affected their perception and its change over time if nay.

Next, the data were more closely scrutinized in search of the interviewees’ perspectives that should be

taken into account in designing and implementing PA for optimal benefits for learners. By repeatedly examining the interview protocols with that goal in mind, several major points which were often shared among the interviewees were finalized to be discussed in relation to the elements in designing and implementing PA in the classroom.

Results and Discussion

The participants generally showed more favorable perception about the benefits of PA on their learning at the end of the semester compared to the beginning, with a little variation in extent and across survey items. Repeated and meticulous examinations of interview protocols mapped on the survey responses revealed important clues to the factors underlying their perception of PA and its transition over time. Several meaningful themes or elements were drawn from this procedure about things to consider for successful implementation of PA, seen from the learners' perspectives. Most salient ones, among others, were as follows: (a) group work; (b) channel of feedback (i.e., in class, face-to-face vs. outside class, online non-face-to-face); (c) frequency of PA activity; and (d) teacher's intervention. In what follows is presented the discussions on each of them.

Benefit of Working in Group

Composition of assessment groups, i.e., whether assessment occurs individually or by group, is one of the important elements in PA design or research (e.g., Adachi et al., 2018b; Gielen et al., 2011; Topping, 2013, 2018). In the current study, the students worked in group to perform the task to be peer-assessed. Although it was primarily for the sake of efficiency of classroom management under time pressure, this group work format proved to be fairly promising in multiplying the beneficial effect of PA on learning. Almost all participants reported that they benefitted from working together within a group in preparing and presenting their task output, for a variety of reasons. The following is some of the excerpts mentioning the benefits of group work, most of which were similar in nature.

Working together is less burdensome than working alone both in preparing the presentation, and in coping with the [unfavorable] feedback from peers.

Giving and receiving peer feedback within the group greatly helped and usually made the final product of the group better.

My group members often saw what I didn't see, and their different viewpoints on the same issue made my perspectives wider and richer.

I was lucky to have wonderful colleagues in my group. They helped a lot for me to deal with huge amount of course materials.

Overall, there seems to have been two main benefits of group work, i.e., emotional security (affective benefit) and synergy effect or what is called 'collective intelligence' effect (practical benefit). In other words, the students felt psychologically secure by facing the peers' evaluation about their work together with other members of the same group, especially when they thought peers' evaluation of their own work was not that favorable. After all, working in group in the context of peer assessment is likely to bring psychological comfort in dealing with the peers' assessment which might not be always favorable, as well as could practically lessen the learners' workload by sharing responsibility.

Although there were a few concerns or complaints about group work format (e.g., "having to allow additional time for gatherings outside classroom," "feeling bad when receiving unfavorable peer feedback

as a result of excluding my opinion in the final group product”), even those participants who talked about it agreed that the benefits far exceeded such minor concerns. In short, group work format seems to have potential for increasing the beneficial effect of peer assessment by providing an additional channel for sharing peer feedback within the same group, in more cooperative and friendly atmosphere before they are assessed by peers in public which could possibly make them feel a little anxious and rather competitive against other groups.

It seems to deserve paying attention that almost all interviewees pointed to ‘diverse perspectives’ as the major virtue of peer feedback. The ‘diverse perspectives’ they mentioned could have been associated with the unique format of PA procedure implemented in the current study. As already mentioned, the learners in the current study worked in group to complete the task and present it to the class and were assessed by the peers except the members in the same group. Within the equal or comparable ability group, learners are likely to confront similar difficulties in performing the given task, which was often mentioned by the interviewees. Given this, the group work mode adopted in the current study would have made it possible for them to get a variety of feedback comments, both from the peers inside the same group and from the peers in other groups, which could have provided them diverse perspectives on how to approach the same task. Meanwhile, the product of group work was rarely peer-assessed in many of previous PA studies, where pair-work prevailed as the manner of peer assessment, which format is not likely to allow the feedback comments to be shared among all the students in the same classroom. Even when a group work format was employed (e.g., Babaii & Adeh, 2019), the manner was different from in the current study, i.e., only the group leader was given the opportunity to evaluate the members in the group.

In this respect, findings described in the above provide meaningful implications for implementing peer assessment in classrooms: other than individual or pair in the manner of configuring assessment groups, and beyond writing samples in the object to be assessed. Although rather different in research context, Pocock, Sanders, and Bundy (2010) reported that group work in peer assessment provided practical experience of teamwork by engaging the students in negotiating with their colleagues. Recalling that one of the acclaimed long-term benefits of PA is its potential to help nurture communicating skills, group work format seems to be worth considering in implementing PA when it naturally provides the learners with opportunities to communicate and negotiate with the group members in completing their common project before being assessed by peers, as well as the practical benefits of collaboration in making their product better.

Mode of Feedback: In-Class, Face-to-Face vs. Out-of-Class, Online Non-Face-to-Face

Previous studies that presented a typology of PA elements (Adachi et al., 2018b; Gienlen et al., 2011; Topping, 2013, 2018) all included as one of its entries the type of feedback, how it is delivered, etc., although there is variation in terms across the typologies. The possible impact of the type or channel of feedback on the overall effect of PA is considered one of the areas that need further investigation. In the current study, those participants who took the course *English language assessment* experienced two different modes of giving/receiving peer feedback: (1) in-class, face-to-face, immediately after their product was presented; (2) out-of-class, via E-campus system, usually sometime after their product was uploaded, as convenient as their personal schedules permitted, within the timeline designated by the instructor.

Most of these participants had enough prior experience of in-class, face-to-face peer feedback in their previous courses, which would certainly have made it possible for them to empirically compare the two different modes of PA. Therefore, it seemed plausible and reasonable to elicit valuable information from the interviewees who practiced PA by both modes about whether and how the different modes of feedback influenced their learning in different ways.

The results were mixed, revealing both positive and negative aspects of each, but with generally more preference for off-line feedback, the reasons for which seems to deserve attention. The two most frequently mentioned strengths of online peer feedback were associated with time and affective pressure.

Different from in-class face-to-face feedback which occurred immediately after each presentation, online feedback via E-campus system, the participants said, allowed them more time both to closely examine the peers' products and to think about what feedback comments to give them. Online feedback also made it possible for more students to engage in giving feedback, which was difficult in in-class feedback session due to the limited time for all students to say something for every presentation. In addition, online feedback comments were provided in written, and were available any time for the presenters to review, whereas the presenters could occasionally miss offline feedback comments which were given in spoken in class, even though they took notes while listening to the peers' feedback comments.

Another important benefit of online feedback was the lowered affective filter. About half of the participants, largely rather introverted ones, reported that the non-face-to-face online feedback was less anxiety-provoking, both when giving and receiving. They said that it was usually difficult for them to speak something in public face-to-face about the peers' performance, but that they could say what they wanted to say in the non-face-to-face online feedback system. Also, reading online rather unfavorable (negative) feedback comments was, according to them, less hurting than listening to those in face-to-face feedback in class.

On the other hand, negative aspects of online non-face-to-face feedback were often associated with learning outcomes, i.e., almost all participants reported that they learned less from online feedback than from offline face-to-face feedback. This finding is quite interesting, for the participants usually felt less anxious in online feedback sessions but they said they learned less from those sessions. It sounds somewhat paradoxical: more anxiety and offended feeling, but learned more; and less anxiety, more time allowed, but less learned. The following excerpts illustrate some of the reasons for [their feeling of] 'less learning' from online feedback.

The quality of the online feedback was generally lower than that of offline feedback, probably because less efforts were exerted [than for offline, face-to-face feedback] when giving feedback.

Differently from in-class face-to-face feedback, I often checked only those feedback comments given to my own group's product, without glancing at feedbacks on other groups' products, which turned out to be not very helpful for my learning.

I took feedback received online rather light compared to the similar feedback received offline, just as much as unfavorable (negative, thorny) feedback received online was less hurting, and it didn't help my learning after all.

I paid less attention to feedback comments given online.

Similar pattern was observed in giving feedback. Most participants reported that they tended to be less rigorous or less serious when giving feedback online than offline. They said they often provided minimum, rather simple and loose comments only, usually regarding the points about which other students already gave feedback. Interestingly enough, such phenomenon was somewhat related to the presence or absence of the instructor at the time of giving/receiving feedback. The following excerpts show how this happened.

The thought that the professor (= the instructor) is not watching us here and now made me exert less efforts and spend less time on feedback giving.

I wish the instructor to intervene and give us feedback immediately after the presentation like in-class feedback sessions.

I think I usually work less hard when there is no supervisor organizing and supporting the class procedures.

As the online feedback does not occur publicly, and there is no teacher looking at us, either, as in the offline feedback session, I feel less anxious and concentrate less on giving feedback to my peers.

In addition, some students pointed out as one of the shortcomings of online feedback the lack of immediate interaction, which, they thought, makes the class more dynamic, saying as follows.

In online feedback sessions the presenters' reaction to the peers' feedback was less faithful than in offline ones.

It was hard to me that simultaneous reciprocal exchange of ideas on my group's product was not usually possible in online system, and it seemed to decrease the usefulness of feedback giving and receiving.

The strengths of face-to-face, in-class feedback were at the other side of the weaknesses of online, non-face-to-face, outside class feedback, just like a mirror image. The following excerpts indicate how the students felt about off-line feedback practices.

Offline feedback is usually more anxiety-provoking, but the anxiety makes me more concentrated, which helps me better understood, give more appropriate feedback to the peers, and also receive more useful feedback from peers.

The fear of getting somewhat negative feedback comments in public is quite burdensome, but it makes me prepare with more efforts [to avoid unfavorable feedback comments], which naturally results in higher learning achievement.

On-site, face-to-face feedback in class is surely of higher quality. Also, I learn a lot from viewing other groups' products.

The strength of offline feedback is the immediacy. That is, the presenters can pick up the peers' feedback immediately after presenting their product, unlike the online one where I should wait until the peers upload their feedback comments on the system.

I think the feedback in the true sense of meaning is given in the in-class offline, although there was some degree of emotional pressure about presenting in public.

I like offline feedback session better because the instructor gives immediate feedback on site about our product. (mentioned by 3 participants)

In the class, I watch other groups' presentations very attentively and give them focused feedback. I don't think I concentrate as much in online feedback as I do in the classroom.

I learned quite a lot from the peers' feedback given to other groups' products, as much as I learned from the feedback given to my own group's product. (mentioned by 3 participants)

As can be expected from these excerpts, the most salient negative feature of offline feedback was anxiety about the possibility of receiving negative feedback from peers and about giving rather unfavorable feedback comments to peers in public. The following excerpts provide some hints about how the students felt.

The feeling of anxiety was the worst thing about offline feedback.

I was often reluctant to give my peers feedback comments face-to-face, as I am an introverted person.

As I am introverted, it was difficult for me to give feedback comments openly in the class, often saying only to myself what I'd like to say to the presenters. I was not sure if my feedback comments were appropriate.

It was always burdensome to me to make a presentation in the presence of others. Also, I get easily hurt if I receive unfavorable feedback immediately following my presentation.

It is not usually easy for me to speak frankly when I have to give feedback openly in class.

Much fewer interviewees mentioned negative aspects of offline feedback than those who talked about its benefits. Three of them talked from the viewpoints of feedback givers, and two explicitly associated their introverted personality with the negative aspects of offline mode of feedback. Overall, however, more students preferred offline only, as reported already. This means not all introverted students preferred non-face-to-face online feedback over offline, although several introverted students favored non-face-to-face online mode of feedback. Even the introverted students, if not all, may have thought the benefits of offline feedback (with its immediate and synchronic interaction among peers) deserved the anxiety they had to confront. Several participants reported they preferred combining on-, and off-line format, with a little bit more of offline opportunities.

As can be seen in these findings, both online and offline have strengths and weaknesses of each as a medium of feedback giving and receiving, and more of the participants preferred face-to-face offline in-class feedback because they thought it was more beneficial in enhancing their learning, although it sometimes accompanied certain degree of anxiety which largely came from their fear of losing face in the presence of peers. The participants particularly appreciated the immediacy of feedback on their performance both from peers and the instructor, which was available only in offline face-to-face mode. In addition, they also highlighted the learning benefits they got from feedback given to the other groups, not simply the feedback given to themselves. Such learning benefits was far more easily available in in-class offline face-to-face feedback session even without intentional efforts on the part of the students, whereas strong will/motivation and additional time would have been required for them to browse online all the feedback comments given to the other groups' products for more learning benefits.

In addition, it seems that the decision on the mode of feedback should depend on the structure of a specific course and the nature of a product to be peer-assessed. The participants in the present study generally agreed that giving feedback online would be more appropriate in the course *English language assessment* (as had been tried in the current study) than in the course *TEFL methods*. As outlined earlier, assessment tasks, usually in the forms of test items, if not entirely, were the object of peer assessment in the former, while in the latter course micro-teaching sessions were assessed by peers. As there were more products for the students to present in the former course, it was sometimes impossible to give feedback comments to all the presented products within the limited time of the given class meeting. Such circumstances somewhat inevitably led to the adoption of online peer feedback mode for some part of the class. Meanwhile, the students particularly expressed concerns about the amount of time it would take them to video-record their micro-teaching for uploading it on the web, which they thought would be a bigger burden for them than simply performing just once in class.

In summary, the teachers are encouraged to try combining optimally the two different modes of feedback, taking into account both the nature of the products to be assessed and the personal characters of the students. As can be seen in the interview excerpts presented in the above, some students could be affectively very vulnerable to negative feedback provided in public, even when the feedback comments in spoken may not sound very harsh. Periodic formal and informal conferences with students would surely

help figure out the optimal solution by modifying the modes of PA if necessary.

Frequency of PA Activities

Another element of peer assessment in terms of design and implementation is ‘timing’, e.g., when, how much time, or how often (Adachi et al., 2018b). Liu and Carless (2006) emphasized the importance of explicitly specifying how much time is required for PA, and Adachi et al. (2018b) pointed out the deficiency of previous studies on this issue, saying that such deficiency could make comparison of different PA models difficult. A possible question relevant to this issue, among others, would be “Does the learning benefits from peer assessment co-vary with the amount (or frequency) of practicing it?” In the current study, most participants responded more positively in the post-survey than in the pre-survey about the learning benefits from peer assessment activities, and it seemed meaningful to figure out what contributed to such change in perception.

From the analysis of interview protocols, two key phrases that emerged in relation to this element were ‘*the more, the better*’ and ‘*trust in the peers’ ability*.’ In other words, more frequent exchange of peer feedback was associated with more learning benefits students perceived, along with increased trust in the peers’ ability to assess, although negative side-effects were also mentioned, such as physical fatigue or loose concentration when giving feedback.

The participants’ first practicing of peer assessment as TEFL majors was in the course for sophomores ‘*TEFL methods*’, and naturally they were not much accustomed to such activity at the beginning. Even though some of them had previous relevant experience in other courses, the manner of assessment and the nature of the product assessed were not identical with those in ‘*TEFL methods*.’ The participants reported that at first they were not very sure of the benefits of peer assessment in their learning mainly because they did not trust the peers’ knowledge and ability to assess themselves, accepting only the instructor’s feedback as correct and trustworthy. As they went through peer assessment practices over the continuing semesters, however, they came to trust peers to a considerable degree, seeing that some of the feedback comments from peers were quite instructive and sometimes even comparable in quality to the instructor’s feedback. The following excerpt from junior students illustrate how their perception developed as they went through more experience of peer assessment.

I was not familiar with peer assessment at the time I took the course TEFL methods. My peers did not give me critical feedback comments. I could not trust peers as assessors as I did not think their feedback was correct [or useful]. Of course, I myself could not give them critical comments, either. I trusted only the instructor’s feedback as it was always correct and trustworthy. However, I gradually grew to trust peers as assessors as I found their feedback helped a lot, and some peer feedback was wonderful, almost as good as the instructor’s. With accumulated experience and knowledge, my peers’ and my perspectives expanded, and we could give each other in-depth feedback. That’s why I came to be more positive about the beneficial effect of peer assessment in the later survey. Before experiencing, I was doubtful about peer assessment, thinking “How can my peers assess me?” But such thought almost disappeared while continually engaging in such activity. It helped [my learning] greatly and motivated me to study harder and prepare for the class more eagerly, therefore I wish more active introduction of peer assessment into classroom procedures in other courses, too.

I put forth more efforts in order to receive higher evaluation from my peers as I found my peers and my own knowledge and ability had enhanced, and it resulted in my improved capacity.

In the course TEFL methods as a sophomore, I was kind of at a loss not knowing exactly how to give feedback to peers, but in junior courses we could give each other better quality feedback, now that we have better idea about what to give feedback on, with more experience of practicing PA. Naturally it was of great help to me.

Repeated training in peer assessment over three consecutive semesters improved my peers' and my ability to appraise, and such improved ability in turn contributed to my learning, cultivating my motivation, and of course to enhancing my learning achievement.

These self-reports imply the importance of learners' belief in their peers' ability in their perception of, and approaches to, the peer assessment. This finding is in line with earlier studies that highlighted learners' perception of peers as an important factor in peer assessment (cf. Babaii & Adeb, 2019; Ballantyne, Hughes, & Mylonas, 2002; Brown, Irving, Peterson, & Hirschfeld, 2009; Kaufman & Schunn, 2011; Le Hebel et al., 2018; Strijbos, Narciss, & Dünnebier, 2010). More importantly, the findings from the current study indicate how experience and training can contribute to forming the students' beliefs towards mutual trust and consequent learning benefits. That is, diachronically (i.e., over the semesters) accumulated training in PA, by repeatedly engaging in relevant activities in the classes, contributed to the learners' increased knowledge and ability to assess, which enhanced their trust in the peers as assessors, and ultimately benefitted their learning.

Another point of interest was whether the synchronic frequency (that is, within a semester time frame) of peer assessing activity has any relevance to learners' perceived effect of PA on their learning. As already described, the participants much more frequently assessed peers and were assessed in the course 'English language assessment' than in the course 'TEFL Methods'. Some of the participants showed highest satisfaction in terms of learning gain after taking the course 'English Language Assessment', and when asked why, they answered multiple presentations of their products (i.e., being assessed by peers many times) helped greater than just being assessed once as in earlier courses they took. The following excerpts illustrate this.

The course English Language Assessment helped enhance my learning achievement more than the two earlier courses I took, and I think it is because I had more frequent opportunities of giving and receiving feedback among peers.

With more things to do for presenting our products, I had to put more time and efforts in preparing. Getting feedbacks from peers again and again, it became easier for me to make my product better, as I could predict to a degree about what my peers would give me feedback comments. It also helped me find with more speed what feedback I should give to my peers' products. Anyway, I'm sure that receiving feedback many times is better [than once].

Making presentations many times, I felt my knowledge and ability had grown a lot. It was good to have an opportunity to compensate my deficiency at the next presentation. The more feedback receiving and giving, the better for my learning achievement, I think.

Although these students affirmed the benefits of frequent feedback giving and receiving, there was a negative aspect of it, too: fatigue, both mental and physical. Having to continuously work on their products to present, and to give feedback comments to and evaluate even more of others' products seemed to have made several students quite tired, as shown in the following excerpt.

It was hard and tiring for me to give feedback to so many different products of my peers, which made me lose interest in this activity.

These results imply that engaging in PA activities over longer period of time, across different courses in a curriculum, could considerably enhance the students' ability to assess and mutual trust in one another as assessors, which often results in the students' satisfaction with PA, feeling that they have learned something meaningful via peer assessment. In addition, having more frequent PA activities in a course (especially being peer-assessed many times than just once) could also increase the students' perceived

learning achievement, by being offered opportunities to improve their products in the following performances and confirm themselves about their growth. As frequent PA activities, however, might demotivate some students, who think the workload is too big, the instructor needs to pay close attention to how the individual students are doing and carefully adjust the frequency of PA activities according to the specific context of the given classroom.

Teacher's role in PA

Some of the facets or elements in designing and implementing PA is related with the role of teacher. Although literature does not explicitly name it 'teacher's role in PA', several descriptors in the lists or typologies point to the manner how the teacher intervenes in the process of PA. For example, Topping's (2013) checklist of factors in PA research includes whether PA activities are monitored by supervisory staff, aligned with traditional assessment, or supported by some scaffolding devices (p. 409), all of which belong to the teacher's domain. Adachi et al. (2018b) also included in their design elements for PA 'relationship to other assessments' (p. 457), with 'giving peer assessment prior to staff assessment' as an example of this element.

In the current study, much of the participants' perspectives on this element were naturally collected from their responses to the questions about why they became more positive about the beneficial effect of PA at the end of the semester compared to the beginning. While recalling the benefits of PA, they often mentioned the unique strengths of peer feedback and those of teacher feedback, and their views about the desirable role of the teacher in the whole PA process were spontaneously expressed. Overall, the participants wanted feedback both from peers and the teacher, clearly recognizing the respective strengths of each. While usually appreciating various beneficial effects of peer feedback, the students still highly valued and depended on the authority and professionalism of teacher feedback. The majority of them wished the teacher to consolidate the peer feedback session by providing a kind of summary and additional feedback comments not mentioned by peers. Several students, however, strongly preferred teacher feedback only, mostly from lack of trust in the quality of peer feedback.

It deserves attention that the students showed a kind of dual attitudes towards peer feedback. Some students said they regarded peer feedback rather light whereas they took teacher feedback quite seriously, which was often associated with their thought that the peers are comparable with themselves in knowledge and ability, as illustrated in the following excerpt by a junior student.

Different from teacher feedback, it was often predictable where peer feedback would be targeted, and peers' feedback contributed less to my growth than the teacher's. That's why teacher feedback is indispensable. On the other hand, a good thing about peer feedback is its positive impact on affective aspects: Peers' feedback was easier to accept because feedback givers and the receivers are at similar level [in knowledge and capacity].

As exemplified in the above, the students were occasionally somewhat reserved about accepting the peers' feedback, since they thought that they were equal in ability, but at the same they said it was easier to accept for the very reason, which seemingly sounds contradictory. The following excerpts demonstrate what the students thought of the respective value of peer-, and teacher feedback, how diverse their ways of thinking could be, probably interwoven with their personal characters and prior relevant experience, and how they wanted these two sources of feedback to be arranged for their optimal learning growth.

Teacher feedback is good as it is to-the-point and looks into what the peers cannot see, whereas I was not sure about the beneficial effect of peer feedback as I thought it tended to be superficial. However, I came to find in the junior course that I could learn a lot from peer feedback, too. Now I think it great to have teacher feedback after receiving peer feedback first, since such arrangement gives me diverse perspectives.

I usually take the teacher feedback more seriously due to its consistency and authority (professionalism), whereas I think peer feedback is rather subjective and sometimes difficult to agree with. I do not exert much effort to improve in the aspects about which peers pointed out, as I do not trust in their professionalism: They might be wrong. By contrast, I think highly of teacher feedback. I often utilize the teacher's feedback to other groups' products when performing my next learning task. When there is discrepancy between the peer feedback and the teacher's, I definitely take the latter because I firmly trust teacher feedback.

I paid much more attention to the teacher's feedback than to peers.' Although peers occasionally provide useful feedback comments, most of them is limited in scope or perspectives, often affected by surface features such as the manners when presenting or fanciness of the materials presented.

I'm very positive about the usefulness of peer feedback now, although I was negative about it before experiencing PA, thinking, 'How could (dare!) my peers evaluate me?' PA especially helps boost learning motivation and active participation in the class. Regarding the time ratio of teacher- vs. peer feedback, I prefer more time allotted for peer feedback as we have done in this course, since peer feedback is more effective for my motivation and class participation. The teacher feedback, I think, had better to center on what is not mentioned in peer feedback, along with a kind of summary.

Peer feedback is easy to accept and understand as we (peers and me) usually go through similar difficulties or problems while involved in the same task. Peers offer a variety of perspectives about how they coped with those difficulties or problems, which were very useful to me in that I could see both how they thought similar to me (which makes me feel relieved) and how differently they thought from me (which makes me learn from them). Therefore, both teacher-, and peer feedback are necessary, with more time allotted for peer feedback and leaving teacher feedback for overall summation and for filling in the niches not covered by peer feedbacks.

Even the students who were rather reserved about peer feedback (not very affirmative about its impact) confirmed that it would be better to go with it than without it, as long as teacher feedback is given also, which is well illustrated in the following excerpt.

Peers tend to evaluate unnecessarily in detail and fragmentarily unlike the teacher who sees the whole picture. Thus, I accept teacher feedback without doubt about its quality and trust-worthiness, but I sometimes wonder about peer assessment if it's correct, thinking I AM correct. Still, peers' diverse perspectives sometimes help and some of their feedback comments are good. Also, it facilitates my class participation and concentration because I need to understand the peers' products in order to give them feedback, which stimulates my motivation to learn. So, I think it not bad to have peer feedback only if the teacher feedback is also provided.

As indicated in these self-reports, peer feedback followed by teacher's summary of the peer feedback session along with additional remarks of the teacher's own was conceived ideal by almost all students, which is further supported in the following excerpts.

*I think it best to listen to peers' feedback first and then to the teacher's wrap up, and actually it was very good to do so in this class. Without the teacher's wrap up, it doesn't seem to be cleared up anyway. So, I prefer to having both teacher-, and peer feedback.
The teacher's and the peers' feedback have its own strength, so both of them are necessary. Teacher feedback has academic depth, while peer feedback offers different perspectives at the similar level to mine. If only one of them is provided, it would not be sufficient, I suppose.*

In summary, the participants agreed on the crucial role of the teacher in peer assessment, regardless of the extent to which they valued the quality of peers' evaluation of their own performance. As for the manner of teacher intervention, they favored a consolidation by the teacher at the end of the peer feedback session of the day where the teacher provides an evaluative summary of the feedback comments from the peers, and additional remarks not made by the peers. It seems apparent that this final stage of PA moderated by the teacher also offered opportunities for the students to train in peer assessment, by giving them a model of how to give feedback, and a kind of check list as well against which they could assure themselves of learning gain.

The teacher's intervention in the PA process could especially matter when peer assessment occurs on-line, in non-face-to-face mode. The participants in the current study who experienced PA both face-to-face in class and non-face-to-face online recalled that one of the most serious drawbacks of online feedback outside the class was the absence of the teacher's immediate and direct feedback to their performance as in the classroom, which helped their learning a lot. Therefore, the teacher needs to be very sensitive about when and how much time the learners want teacher feedback to be provided and take this into account when designing and implementing PA activities, as well as the nature of the task to be peer-assessed since "PA is subject-specific" as Adachi et al. (2018b, p. 456) pointed out.

Conclusion

Summary of the Study

As indicated well in Gielen et al. (2011), Topping (2013, 2018), and Adachi et al. (2018b), there are a variety of elements to consider in designing and implementing PA. Specific circumstances or contextual constraints in individual classrooms (e.g., nature of the subject matter area, institutional restrictions, students' composition and their characteristics and needs, etc.), however, can make it difficult to incorporate some of those designing elements. Partly for this reason, existing PA studies are quite varied in these elements, making it difficult to compare and discuss the findings from different PA studies conducted in various educational settings and under different frameworks.

Therefore, it seems necessary to search for more effective implementation of PA by empirically investigating the concrete and specific conditions where PA is practiced and the manners how it is practiced. Accumulated research findings from the studies across diverse educational settings would greatly contribute to our understanding of how to obtain maximal benefits from PA. They also will provide instructive information and affective support for the teachers and learners who are reluctant, largely due to the lack of trust in its strengths, to engage in this highly learner-centered and promising form of assessment and learning.

Given this, the current study attempted to figure out under what conditions the beneficial effects of PA would be maximized, based on the voice of learners who have experienced PA in their content courses as education majoring students and on the researcher's own observation of real classrooms that employed PA as a major learning procedure. The following is an overall picture of the current study.

The context of the current study was two content courses for pre-service teachers where peer assessment was utilized as a core learning procedure, with multiple purposes such as training them as assessors, encouraging them to cooperate with others towards the same goal, and providing them opportunities to learn from each other and then improve themselves. The product to be peer assessed was group oral presentation (i.e., a micro-teaching) for the sophomores, and assessment item construction for the juniors. Both quantitative rating (from 1~5 on each criterion) and qualitative feedback information (i.e., commentary feedback) were exercised. Feedback comments were given in oral for in-class face-to-face feedback session, and in written for outside-class online session, depending on the course environment.

As for the composition of assessment group, the students performed and were assessed in group, but

worked individually when assessing (giving feedback to, and grading) other groups. The students formed a group of 4 or 5 as they would like to, without teacher intervention in this process. For more reliability and validity of assessment procedure, the students were provided with assessing criteria created by the teacher, about which the teacher explained before they used it. In the present study, the teacher wrapped up the peer assessment session of the day by giving feedback to the presenters and to their peer-assessors. The students favored a lot the teacher's summary given after the peer assessment session.

At the beginning and the end of the semester, the students completed a survey about their perception of PA's impact on their learning. After all the coursework was over, they were contacted for an individualized interview, and about half of them accepted the invitation. Questions were asked of each interviewee based on his/her responses to the pre, and post-survey in search of clues to their perception and its change over time. Several themes emerged from these interview protocols, which provided precious insights about how to design and implement PA for better results.

Findings and Implications

Some of the major findings from the current study are summarized in the following, with implications for teachers or researchers that can be drawn from those findings. First, working in group to perform the task to be peer-assessed proved to be quite fruitful in that such format benefitted the learners both affectively and academically. By having much more opportunities to share their learning ideas and strategies and thus learning from each other, first within the group while preparing for the group presentation, and then exchanging feedback with other groups, the learners could have the feeling of achievement or growth. Such group format could also be efficient for involving as many students as possible in PA activity, compared with pair or individual assessment, within the limited time frame of the given class, while maintaining beneficial impact of PA on learning. Therefore, it seems to be recommendable that teachers try to exercise group work format in their classroom when employing PA as a learning activity.

Second, the participants largely tended to prefer in-class, face-to-face feedback over on-line feedback, recalling that the former was more helpful for their learning with its immediacy and directness (i.e., given right after their presentation) combined with teacher feedback at the same site. They reported in-class face-to-face feedback helped them more concentrated than online, although it could sometimes be anxiety-provoking. However, they also recognized, especially those who are more introverted, the benefit of non-face-to-face online feedback, e.g., feeling more secure emotionally without having to front some uneasiness when they gave or received not very favorable feedback comments in public. After all, most participants thought it best to combine *appropriately* these two modes, although sounding a little vague, with face-to-face in-class feedback sessions more often. Given this, it seems reasonable for the teacher to be flexible about the mode of feedback exchange by carefully considering the nature of the task (e.g., writing samples, micro-teaching presentation, or test-item construction) and the personal characteristics of the students. Having occasional small conferences with the students would help to figure out which mode they really prefer, putting all things together.

Third, the participants usually recognized the benefit of being frequently exposed to PA activities in that it offered more training opportunities, thus better chances of improving their ability to assess, which also increased their trust in the peers as eligible assessors as well as their own self-efficacy. This does not mean, however, that the principle of 'the more often, the better' will always work, as illustrated in some self-reports that mentioned its possible negative side-effect. The teachers, therefore, are advised to be sensitive to the students' needs and affective state by closely observing how they do and what they think. By having periodical formal and informal conferences with students about the processes of PA, the optimal 'timing' (cf. Adachi et al., 2018b) of PA activities could be determined that could both enhance its beneficial impact on learning and keep the learners motivated to engage in it. In addition, the nature of the subject matter course and the demands of the task to be performed need also to be taken into account in adjusting the frequency or amount of time allocated to PA activities in the given classroom.

Fourth, the participants perceived indispensable of the teacher's intervention in the PA procedure irrespective of the extent to which they valued the benefits of PA in their learning. Apart from their generally increasing trust in the peers as qualified assessors with their repeated practice, they would like the teacher to intervene in PA procedure, preferably at its final stage, by giving evaluative feedbacks both to each performance group and to the whole class especially on those aspects not covered by the students. This implies the participants in the current study recognized unique strengths of peer assessment, while still depending on the teacher as the source of absolute authority in assessment, without which they felt something was missing in the class, their learning process not being completed. Therefore, the teacher is supposed to observe very attentively the students' feedback exchange and help them feel assured of their own learning outcome by having a kind of consolidation phase in which the teacher summarizes the day's peer feedback activities, provides comprehensive feedback comments to the students' performance as well as to the peer feedback comments made by the students. By so doing the teacher can encourage the students to more spontaneously be engaged in PA activities, affirming themselves of the learning gains from those activities and of what they further should do to improve themselves.

Some limitations of the current study should be admitted. Most of all, the context of the current study is quite specific and narrow, situated in a higher education site, which could be inevitable in small scale studies like the current one. Small number of participants made it impractical and difficult to run inferential statistics, which might reduce the interpretability of research findings. Without having a control group or experimental design, generalizability of the findings would be limited. Only very limited number of elements were discussed among the many possible elements or factors in designing and implementing PA.

Still, the current study observed how experience could affect the learners' perception, and elicited learners' vivid voice based on their own experience that can shed meaningful insights about how to design PA towards optimal satisfaction of learners. Accumulated findings from similar empirical studies conducted in diverse educational contexts that attend to learners' inner voice would contribute to build up a more robust and sustainable framework for efficient designing and implementing PA. Such a theoretical framework will then make it easier to compare different studies on PA and more useful information could be drawn to be applied to classroom implementation of PA.

The Author

Jeongsoon Joh is a professor of English Education at Konkuk University in Seoul, Korea. She is particularly interested in L2 reading comprehension from both theoretical and pedagogical perspectives. Her recent research interests include peer assessment as a learning experience for teacher-trainees.

Department of English Education
College of Education, Konkuk University
Seoul, Korea, 05029
Mobile: + 82 1089724373
Email: johjs@konkuk.ac.kr

Acknowledgements

This paper was supported by Konkuk University in 2019.

References

- Adachi, C., Tai, J. H-M., & Dawson, P. (2018a). Academics' perceptions of the benefits and challenges of self and peer assessment in higher education. *Assessment & Evaluation in Higher Education*, 43, 294-306. <https://doi.org/10.1080/02602938.2017.1339775>
- Adachi, C., Tai, J. H-M., & Dawson, P. (2018b). A framework for designing, implementing, communicating and researching peer assessment. *Higher Education Research & Development*, 37, 453-467. <https://doi.org/10.1080/07294360.2017.1405913>
- Alfallay, I. (2004). The role of some selected psychological and personality traits of the rater in the accuracy of self-and peer-assessment. *System*, 32, 407-425. <https://doi.org/10.1016/j.system.2004.04.006>
- Allen, D., & Katayama, A. (2016). Relative second language proficiency and the giving and receiving of written peer feedback. *System*, 56, 96-106. <https://doi.org/10.1016/j.system.2015.12.002>
- Azarnoosh, M. (2013). Peer assessment in an EFL context: Attitudes and friendship bias. *Language Testing in Asia*, 3, 3-11. <https://doi.org/10.1186/2229-0443-3-11>
- Babaii, E., & Adeh, A. (2019). One, two, ..., many: The outcomes of paired peer assessment, group peer assessment, and teacher assessment in EFL writing. *The Journal of Asia TEFL*, 16, 53-66. <http://dx.doi.org/10.18823/asiatefl.2019.16.1.4.53>
- Ballantyne, R., Hughes, K., & Mylonas, A. (2002). Developing procedures for implementing peer assessment in large classes using an action research process. *Assessment & Evaluation in Higher Education*, 27, 427-441. <https://doi.org/10.1080/0260293022000009302>
- Berggren, J. (2015). Learning from giving feedback: A study of secondary-level students. *ELT Journal*, 69, 58-70. <https://doi.org/10.1093/elt/ccu036>
- Birjandi, P., & Tamjid, N. H. (2012). The role of self-, peer and teacher assessment in promoting Iranian EFL learners' writing performance. *Assessment & Evaluation in Higher Education*, 37, 513-533. <https://doi.org/10.1080/02602938.2010.549204>
- Boud, D., & Falchikov, N. (2006). Aligning assessment with lone-term learning. *Assessment & Evaluation in Higher Education*, 31, 399-413. <https://doi.org/10.1080/02602930600679050>
- Brown, G. T. L., Irving, S. E., Peterson, E. R., & Hirschfeld, G. H. F. (2009). Use of interactive-informal assessment practices: New Zealand secondary students' conceptions of assessment. *Learning and Instruction*, 19, 97-111. <https://doi.org/10.1016/j.learninstruc.2008.02.003>
- Buchanan, E. A. (2004). Online assessment in higher education: Strategies to systematically evaluate student learning. In C. Howard, K. Schenk, & R. Discenza (Eds.), *Distance learning and university effectiveness: Changing educational paradigms for online learning* (pp. 163-176). Hershey, PA: Information Science Publishing.
- Chang, C. C., Tseng, K. H., Chou, P. N., & Chen, Y. H. (2011). Reliability and validity of web-based portfolio peer assessment: A case study for a senior high school's students taking computer course. *Computers & Education*, 57, 1306-1316. <https://doi.org/10.1016/j.compedu.2011.01.014>
- Cheng, W., & Warren, M. (2005). Peer assessment of language proficiency. *Language Testing*, 22, 93-121. <https://doi.org/10.1191/0265532205lt298oa>
- Chong, I. (2017). How students' ability levels influence the relevance and accuracy of their feedback to peers: A case study. *Assessing Writing*, 31, 13-23. <https://doi.org/10.1016/j.asw.2016.07.002>
- Falchikov, N. (2007). The place of peers in learning and assessment. In D. Boud & N. Falchikov (Eds.), *Rethinking assessment in higher education: Learning for the long term* (pp. 128-143). New York, NY: Routledge.
- Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70, 287-322. <https://doi.org/10.3102/00346543070003287>
- Ganji, M. (2009). Teacher-correction, peer-correction and self-correction: Their impacts on Iranian students' IELTS essay writing performance. *The Journal of Asia TEFL*, 6, 117-139.

- Gielen, S., Dochy, F., & Onghena, P. (2011). An inventory of peer assessment diversity. *Assessment & Evaluation in Higher Education*, 36(2), 137-155. <https://doi.org/10.1080/02602930903221444>
- Hu, G., & Lam, S. T. E. (2010). Issues of cultural appropriateness and pedagogical efficacy: Exploring peer review in a second language writing class. *Instructional Science*, 38(4), 371-394. <https://doi.org/10.1007/s11251-008-9086-1>
- Joh, J. (2019). Impact of peer feedback on learning: A case of EFL teacher trainees. *The Journal of Asia TEFL*, 16, 1103-1118 <http://dx.doi.org/10.18823/asiatefl.2019.16.4.3.1103>
- Hu, G., & Lam, S. T. E. (2010). Issues of cultural appropriateness and pedagogical efficacy: Exploring peer review in a second language writing class. *Instructional Science*, 38(4), 371-394. <https://doi.org/10.1007/s11251-008-9086-1>
- Kaufman, J. H., & Schunn, C. D. (2011). Students' perceptions about peer assessment for writing: Their origin and impact on revision work. *Instructional Science*, 39, 387-406. <https://doi.org/10.1007/s11251-010-9133-6>
- Kissau, S. P., & King, E. T. (2015). Peer mentoring second language teachers: A mutually beneficial experience? *Foreign Language Annals*, 48, 143-160. <https://doi.org/10.1111/flan.12121>
- Le Hebel, F., Constantinou, C. P., Hospesova, A., Grob, R., Holmeier, M., Montpied, P., Žlábková, I. (2018). Students' perspectives on peer assessment. In J. Dolin & R. Evans (Eds.), *Transforming assessment: Through an interplay between practice, research and policy* (pp. 141-173). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-63248-3_6
- Lee, M-K. (2015). Peer feedback in second language writing: Investigating junior secondary students' perspectives on inter-feedback and intra-feedback. *System*, 55, 1-10. <https://doi.org/10.1016/j.system.2015.08.003>
- Liu, N-F., & Carless, D. (2006). Peer feedback: The learning element of peer assessment. *Teaching in Higher Education*, 11, 279-290. <https://doi.org/10.1080/13562510600680582>
- Morahan-Martin, J. (1996). Should peers' evaluations be used in class projects?: Questions regarding reliability, leniency, and acceptance. *Psychological Reports*, 78(3), 1243-1250. <https://doi.org/10.2466/pr0.1996.78.3c.1243>
- Murillo-Zamorano, M. (2018). Oral presentations in higher education: a comparison of the impact of peer and teacher feedback. *Assessment & Evaluation in Higher Education*, 43(1), 138-150. <https://doi.org/10.1080/02602938.2017.1303032>
- Nawas, A. (2020). Grading anxiety with self and peer-assessment: A mixed-method study in an Indonesian EFL context. *Issues in Educational Research*, 30, 224-244.
- Nelson, G., & Carson, J. (2006). Cultural issues in peer response: Revisiting "culture." In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 42-59). Cambridge University Press, New York, <http://dx.doi.org/10.1017/cbo9781139524742.005>
- Park, J. (2018). Effectiveness of teacher and peer feedback: Through the lens of Korean Tertiary writing classroom. *The Journal of Asia TEFL*, 15, 429-444. <http://dx.doi.org/10.18823/asiatefl.2018.15.2.11.429>
- Patton, B. J., & Marty-Snyder, M. (2014). Practical applications for using peer assessment in physical education field experiences. *Strategies*, 27(6), 25-31. <https://doi.org/10.1080/08924562.2014.960121>
- Paulus, T. M. (1999). The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing*, 8, 265-289. [https://doi.org/10.1016/S1060-3743\(99\)80117-9](https://doi.org/10.1016/S1060-3743(99)80117-9)
- Pocock, T. M., Sanders, T., Bundy, C. (2010). The impact of teamwork in peer assessment: a qualitative analysis of a group exercise at a UK medical school. *Bioscience Education*, 15, 1-12. <https://doi.org/10.3108/beej.15.3>
- Reinholz, D. (2016). The assessment cycle: A model for learning through peer assessment. *Assessment & Evaluation in Higher Education*, 41, 301-315. <https://doi.org/10.1080/02602938.2015.1008982>
- Roscoe, R. D., & Chi, M. T. H. (2007). Understanding tutor learning: Knowledge-building and knowledge-telling in peer tutors' explanations and questions. *Review of Educational Research*, 77,

- 534-574. <https://doi.org/10.3102/0034654307309920>
- Roskams, T. (1999). Chinese EFL students' attitudes to peer feedback and peer assessment in an extended pair work setting. *RELC Journal*, 30, 79-123. <https://doi.org/10.1177/003368829903000105>
- Ruegg, R. (2015). The relative effects of peer and teacher feedback on improvement in EFL students' writing ability. *Linguistics and Education*, 29, 73-82. <https://doi.org/10.1016/j.linged.2014.12.001>
- Sahin-Taskin, C. (2018). Effects of active learning environments supported with self- and peer assessment on pre-service teachers' pedagogical and self-efficacy beliefs. *Asia-Pacific Journal of Teacher Education*, 46(5), 421-440. <https://doi.org/10.1080/1359866X.2017.1355049>
- Saito, H. (2008). EFL classroom peer assessment: Training effects on rating and commenting. *Language Testing*, 25(4), 553-581. <https://doi.org/10.1177/0265532208094276>
- Sippel, L., & Jackson, C. N. (2015). Teacher vs. peer oral corrective feedback in the German language classroom. *Foreign Language Annals*, 48, 688-705. <https://doi.org/10.1111/flan.12164>
- Strijbos, J-W., Narciss, S., & Dünnebier, K. (2010). Peer feedback content and sender's competence level in academic writing revision tasks: Are they critical for feedback perceptions and efficiency? *Learning and Instruction*, 20, 291-303. <https://doi.org/10.1016/j.learninstruc.2009.08.008>
- Tahir, I. H. (2012). A study on peer evaluation and its influence on college ESL students. *Procedia*, 68, 192-201. <https://doi.org/10.1016/j.sbspro.2012.12.219>
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68(3), 249-276. <https://doi.org/10.3102/00346543068003249>
- Topping, K. J. (2013). Peers as a source of formative and summative assessment. In J. H. McMillan (Ed.), *Handbook of research on classroom assessment* (pp. 395-412). Thousand Oaks, California: SAGE Publications, Inc.
- Topping, K. J. (2018). *Using peer assessment to inspire reflection and learning*. New York, NY: Routledge.
- Tsui, A. B. M., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing*, 9, 147-170. [https://doi.org/10.1016/S1060-3743\(00\)00022-9](https://doi.org/10.1016/S1060-3743(00)00022-9)
- Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, 15, 179-200. <https://doi.org/10.1016/j.jslw.2006.09.004>
- Yu, S., & Hu, G. (2017). Understanding university students' peer feedback practices in EFL writing: Insights from a case study. *Assessing Writing*, 33, 25-35. <https://doi.org/10.1016/j.asw.2017.03.004>
- Yu, S., & Lee, I. (2016). Exploring Chinese students' strategy use in a cooperative peer feedback writing group. *System*, 58, 1-11. <https://doi.org/10.1016/j.system.2016.02.005>
- Zhang, S. (1995). Re-examining the affective advantages of peer feedback in the ESL writing class. *Journal of Second Language writing*, 4, 209-222. [https://doi.org/10.1016/1060-3743\(95\)90010-1](https://doi.org/10.1016/1060-3743(95)90010-1)
- Zheng, C. (2012). Understanding the learning process of peer feedback activity: An ethnographic study of exploratory practice. *Language Teaching Research*, 16, 109-126. <https://doi.org/10.1177/1362168811426248>
- Zou, Y., Schunn, C. D., Wang, Y., & Zhang, F. (2017). Student attitudes that predict participation in peer assessment. *Assessment & Evaluation in Higher Education*, 43, 800-811. <https://doi.org/10.1080/02602938.2017.1409872>
- Zhu, W. (2001). Interaction and feedback in mixed peer response groups. *Journal of Second Language Writing*, 10, 251-276. [https://doi.org/10.1016/S1060-3743\(01\)00043-1](https://doi.org/10.1016/S1060-3743(01)00043-1)
- Zhu, W., & Mitchell, D. (2012). Participation in peer response as activity: An examination of peer response stances from an activity theory perspective. *TESOL Quarterly*, 46, 362-386. <https://doi.org/10.1002/tesq.22>

(Received January 04, 2021; Revised February 28, 2021; Accepted March 10, 2021)