Students' Perceptions of Peer Evaluation and Teachers' Role in Seminar Discussions

Linda Kwok
(eclinda@polyu.edu.hk)
The Hong Kong Polytechnic University, Hong Kong

Abstract

Peer evaluation is considered as an alternative to teacher-based evaluation and its effects on writing assessments have been widely researched. This study aims at gauging students' perceptions of peer evaluation in a seminar discussion by a group of second-language university learners. With the imminent evolution in teachers' role in evaluation, this study also tries to explore teachers' role from the students' perspective in evaluation. The qualitative and quantitative responses suggest that although minimal evaluation training was offered to students, their evaluation and comments were considered fair, useful and sufficient by peers. Nevertheless, the students preferred teachers to peers to evaluate their performance in the discussion and they wanted to get more instructive comments from the teacher. The teacher's possible roles in evaluation are presented at the end of the paper.

1 Introduction

In the traditional practice of education, summative assessment of a student's performance has long been an established practice, in which the students get a grade or score from the teacher on their achievement at the end of a course or a unit of study (Nitko & Brookhart, 2007). Summative assessment, briefly called "assessment of learning" (Ainsworth & Viegut, 2006, p. 8), is generally used by teachers for making decisions on grouping and promoting students (Airasian, 1997). Also, teachers usually use the assessment results to measure the effectiveness of their teaching and to improve their instruction of the assessed areas in the future (Ainsworth & Viegut, 2006). In addition to giving the teachers clear indicators of the students' performance, the grades are considered as "concrete evidence" to parents about the ability or deficiency of their children in an assessment (Airasian, 1997, p. 127).

As teachers take up the major responsibility in administering and prescribing assessment details, it is argued that "summative assessment can be viewed as establishing concensus on what is regarded as valid academic knowledge amongst academics in the disciplinary community" (Havnes & McDowell, 2004, p. 11). It seems that teachers and students are two distinct separate groups at two different levels in the assessment process, with the teachers at the top designing and executing the assessment and the students on the receiving end trying hard to reach the standards set. With such a "vertical relationship between teachers and students" (Havnes & McDowell, 2004, p.6), the assessment provides few opportunities for the students to communicate their expectations and perspectives of assessment criteria, to critically examine the marking standards or participate in the assessment process (Havnes & McDowell, 2004). Hence, it is observed that the "students work mostly for a test grade and not for their enjoyment or understanding of the subject matter" (Airasian, 1997, p. 127; O'Donnell & Topping, 1998). Students when receiving the assessment

results usually pay most of their attention to their scores and are almost negligent to the instructional comments given by the teacher for future improvement (Race & Brown, 1988; Crooks, 1988).

With the increasing attention to learner autonomy, the process of assessment at different stages such as setting marking criteria, giving peer evaluation and grading has made assessment more objective and multi-faceted as different voices from the most concerned parties are heard. A substantial amount of research on peer evaluation or assessment has been done on writing skills (O'Donnell & Topping, 1998) of first language learners. It seems that relatively little has been conducted on evaluating speaking skills in seminar discussions with second language learners. Seminar discussions which involve the ability to present ideas verbally, engage in discussion, motivate group members to speak, respond to comments, defend or clarify one's stance or opinions and fend off interruption, are essential for learning in universities for development of ideas with student contribution (Rendle-Short, 2006; Bligh, 2000). Unlike student writers who can revise a peer's draft for a few times back and forth, students evaluating seminar discussions have to be in real time listening attentively noting speech sounds, spoken grammar, group interaction, body language, idea development, proper degree of formality and appropriate choice of vocabulary.

This study aims to examine tertiary students' qualitative and quantitative feedback from an anonymous questionnaire to see whether they find peer evaluation of seminar skills instructive, and whether they consider themselves competent as peer evaluators with minimal training allowed by a crammed syllabus. Moreover, their perceptions of the role of the teacher in evaluation will also be examined. It is hoped that the study can shed light on the roles of the students and teachers in peer evaluation of a speaking activity common to the undergraduates in higher education.

2   Literature review

2.1   Peer evaluation

Peer evaluation or assessment refers to the process in which a group of students identify and observe the mastery or performance of particular aptitudes or skills desired by the group after training or learning. The students explicitly or implicitly hold themselves mutually responsible for the successful completion of the evaluation exercise (Topping, 2003). The observation involves the students' point of view against the marking criteria which are accepted and perhaps created by the students (Ainsworth & Viegut, 2006). Then, informal feedback about possible future improvements or formal assessment contributing to a student's grade will be given by the students.

A variety of positive feedback of peer evaluation has been documented and it is generally believed that peer evaluation can promote critical thinking (Brown, Bull, & Pendlebury, 1997), enhance learning and critical understanding of evaluation criteria and the knowledge gap, develop "social and communication skills, negotiation and diplomacy, and useful transferable skills like giving and handling criticism, self-justification and assertion" (Topping, 2003, p. 67). Although the reliability of peer evaluation or assessment has been challenged from the perspective of subjectivity and reciprocity effects (Brown, 2004; Topping, 2003), positive correlation of reliability between peer and teacher evaluation has been recorded in studies (Segers & Dochy, 2001; Topping, 2003) with well-designed support such as training, clear description of evaluation criteria and teacher assistance.

In higher education, students find peer evaluation "less anxiety-provoking" (Topping, 2003, p. 72) as students enjoy the sense of ownership and direct involvement in judging the quality of student achievements (Brown, 2004; O'Donnell & Topping, 1998), and substantial benefits to learning have been observed among marketing and pharmacy students (Stefani, 1992; Catterall, 1995). Favourable impacts on language learning are also reported, for instance, greater improvement in writing proficiency in the peer feedback group (Topping, 2003), significantly improved performance in oral presentations using peer review (Mitchell & Bakewell, 1995) and peer feedback achieving similar effectiveness as teacher comments on grammatical accuracy for assessment of essays (Jacobs & Zhang, 1989).

The advantages of peer evaluation have been quite well documented on writing classes for first-language learners. Smyth (2004) points out that assessing or judging one's or peer's work is deemed as a way of training a student's critical thinking and an integral part of learning. Inoue (2005, p. 209) argues succinctly that involving students in evaluating essays, teaching them how to assess themselves and their peers, and articulating "how assessment and writing work in their own practices" can help students "to be more critical, self-conscious, and hopefully better writers". Stanley (1992, p. 217) describes the varied advantages of peer evaluation "including reduced writing anxiety, improved sense of audience, and increased fluency". Moreover, working with and evaluating peers in co-operative learning often result in greater motivation and a higher sense of self-esteem among students, and an increased willingness to take risks and try new tasks, not to mention the improved interpersonal relationships (Strom, Strom, & Moore, 1999; Elliott & Higgins, 2005).

2.2   Training students for peer evaluation

With the inclusion of peer evaluation, students are assumed to play an important role in assessment as the focus is shifted from the teacher to the learner and students share part of the responsibility and authority. Some researchers query what kind of interpersonal interaction among students can facilitate this purpose. Carson and Nelson (1996, pp. 16–17) videotaped students undertaking peer evaluation on ESL composition classes for six consecutive weeks. Eventually they concluded that in comparison, the Chinese students might hold back or scale down comments to their Spanish counterparts on the grounds of avoiding conflict and maintaining group harmony. The Chinese writers also considered themselves as ineffectual evaluators and did not fully specify the problems in their peers' work.

A vast number of studies focus on how to prepare or train students for peer evaluation. Smyth (2004) argues that teaching students about self and peer evaluation is an integral part of learning. By discussing and analyzing assessment practices with their peers, students can gain a better understanding of how formative assessment forms part of the learning process and increase their confidence levels in critical evaluation skills. In addition, the students can focus on making contribution to seminar tasks as team players instead of lone individuals. Elliott and Higgins (2005) argue that it is vital to give students time to develop skills in self and peer evaluation. In their study, the cohort of students openly discussed possible grading options for self and peer evaluation and finally decided to use the positive grading options as the means to moderate a peer's grade for their group projects. In Stanley's (1992) study, students were prepared for peer evaluation in an elaborate training scheme of approximately seven hours, in which the students were sensitized to the written drafts by being urged to judge the writer's claims and assumptions in evaluation analysis sessions, and developed an agreed evaluator tact to communicate their feedback to their peers after a number of role plays. Compared with another group receiving only one hour of evaluation preparation, the trained group demonstrated 'a greater level of student engagement in the task of evaluation, more productive communication about writing, and clearer guidelines for the revision of drafts' (p. 217). Nevertheless, Stanley (1992) ended her studies with an acknowledgment that the productive outcomes were the fruit of a considerable investment of time and effort in getting students ready for peer evaluation.

Similarly, Inoue (2005) considers it important to train or develop students to take control of all writing and assessment practices of the class. Over a period of three semesters, the students were pushed to develop their assessment rubrics and become critical peer and self evaluators and reflective writers, "articulate and theorize their assessment practices" to improve writing (p. 234). Inoue (2005) did not assess, evaluate, or grade his students' writing although they still received course grades at the end of the course. The teacher had private conferences with the students about their portfolio grades, administered and set reflection prompts for the discussion board, hosted in-class discussion and read the written work without giving students comments or replies to their posted questions. Instead, the students were given ample opportunity to critique their peers' drafts and finished work, query the assessment criteria in the rubrics, and eventually give grades to their

classmates on their work.

Although most of his students finished the course with a sense of improvement, Inoue (2005) could not help wondering "if drafts really get better, if writers learn to actually write better" (p. 232). The portfolio of a very responsive, analytical student stirred in him the question: "Would she have done better if I had assessed her writing?" (p. 233). While the students were developing the skills and ability to be capable peer evaluators, the issue arising from the teacher's perspective was: was the quality of a student's work more crucial than the development of his ability as a learner and assessor?

2.3 Peer evaluation and the teacher

With the change of the students' roles in learning and assessment, the teacher's role has changed inevitably. The question about the role played by the teacher and his expertise in assessment has to be answered now as assessments need not be teacher-centered. Some research has been assessing the reliability of peer evaluation as substitutes for instructor evaluation. MacAlpine (1999) finds that a more detailed method of peer assessment results in an enhanced correlation of the average grade for a given student from the class with the lecturer's grade for the same student. Moreover, the new approach of grading four aspects of presentation using a Likert scoring scale is found to give a very significant improvement in consistency and discrimination. MacAlpine (1999, p. 24) hence suggests inclusion of peer assessment of student presentations "covering syllabus material in lieu of the lecturer" as part of syllabus teaching. Similarly, Campbell, Mothersbaugh, Brammer and Taylor (2001) find that both holistic and analytical peer evaluations of oral business presentations are reasonable substitutes for instructor evaluations, when the students have been trained for an hour as raters and the rubric provides detailed descriptions of four aspects of a presentation on a five-point scale. These researchers (2001, p. 37) conclude that peer evaluation can provide "needed, supplemental instruction" for students to improve on future performances and to predict the instructor's evaluation.

Inoue (2005), on the other hand, points out that many students "complain about the intense focus on peer assessment and class-constructed rubrics. They say the teacher should be the centre of knowledge about writing, and the creator of rubrics" (p. 232). Therefore, it seems essential to address the issue of striking a balance between the teacher-based and the students-based evaluation so that the strengths and perspectives of both parties can be best explored to maximize learning. Gopinath (1999, p. 10) observes that the role of teachers has changed from experts or judges to coaches or facilitators, "sharing the assessment process with the students can potentially be attractive as well as efficient, and can release faculty time for more educationally worthwhile pursuits". It is thus of research interest to find out how to make the best use of the surplus time and energy of a teacher possibly released by peer evaluation or assessment.

3 Design of the study

This study aims to first measure the impact of peer evaluation on seminar discussions in higher education. As it is an attention-demanding language task that demands real-time evaluation, the study tries to investigate whether the second language learners consider themselves as competent evaluators without elaborate, explicit training as evaluators. They learn to exhibit discussion skills and at the same time to appraise similar traits in the peers due to a tight syllabus. It is also hoped that the qualitative and quantitative results can help the author to reflect on the teacher's role in relation to assessment procedures.

The research was conceived when the author realized that she did not have enough time to give immediate feedback to five seminar discussion groups in two hours of lesson time, and tried to use peer evaluation to give the groups quick and a fair amount of feedback and the sense of being closely observed and scrutinized. Since there was not another comparable class covering the same syllabus taught by the author and the crammed curriculum made it difficult to use other teachers' classroom time, the pilot test was carried out with a group of five year-one students enrolled on an

English enhancement programme. They were divided into two separate groups and observed one another in turns in group discussions. Overall, they found the evaluation form comprehensive and clear. The time allocated for peer evaluation and giving feedback was considered suitable. Most importantly, their responses were fully echoed or supported by the findings collected in the study.

This research involved 19 Cantonese-speaking year-one undergraduates undertaking a 14-week, 42-hour English for Academic Purposes course. In the first meeting of the course, the students were given detailed information sheets about the course and the seminar discussions, which were timetabled for a double-hour lesson in Week 6 after five weeks of classroom teaching, three hours per week. Seminar skills were taught in about 6 hours along with written skills, and their importance for academic studies and career development was stressed. In the seminar discussion, the students in small groups of 4 to 5 would discuss for 20 to 25 minutes on a topic selected by the members in Week 2 related to their disciplines. As students' performance in the seminar discussion would not count towards the final grade and peer evaluation was mainly for student learning and enhancement, very low stakes were involved. However, it was tied up with an individual written assignment on the same chosen topic, which accounted for 30% of the grade for the course and had to be submitted in Week 8. The seminar discussion was made compulsory in the syllabus and provided a chance for the students to clarify concepts, exchange information and opinions for the written assignment.

In class, the students had the chance to learn the appropriate discussion skills, to critique and discuss good or unsatisfactory seminar skills presented in visual teaching materials supported with teacher comments and checklists. The objective of the lessons was to help students master discussion techniques, and no extra hours or set-apart sessions were dedicated to peer evaluation training. The students were informed that each student would play the evaluator and evaluatee in turns in Week 5, but specific arrangements would be announced on the day the seminars were held. However, the evaluation criteria were spelt out to them early on during class discussions of the video clips and class practice, and each student was given a copy of the evaluation form (Appendix 1) in Week 5 for reference and preparation for the seminar discussion. The evaluation form was designed with the purpose of drawing as much comment as possible from the student evaluators. In addition to assessing the peers' seminar ability on a five-point scale with descriptors ranging from 'very good' to 'very poor' on the following sections, namely content, interactive skills, accuracy, pronunciation, referencing and body language, the evaluators were encouraged to write down comments on accuracy, pronunciation and other aspects.

In Week 6, on the day the seminar was held, arrangements were made to ensure that each student would have the chance to play the two roles of evaluatee and evaluator in the double lesson and the evaluation was carried out on a one-to-one basis. The self-formed groups were divided into evaluator or evaluatee groups, and groups of evaluators were randomly assigned to observe groups of evaluatees. When a seminar discussion took place, the evaluators sat in the outer ring to observe the evaluatees' performance. After one round of seminar discussions, the evaluators were given some minutes to fill out the evaluation forms, and then ten minutes to give or receive feedback in one-to-one consultations. Afterwards, the second round of discussions began and the groups would reciprocate their roles: the evaluators and evaluatees exchanging their positions. An anonymous questionnaire was given to each student before the end of the double lesson and the students were asked to complete the questionnaire at home and return it in the following meeting. The teacher would not give each student a grade or individual remarks, but note down some common mistakes and give the class feedback in the next meeting.

4   Results

The respondents were asked to answer questions first as evaluatees and then to answer the second part of the questionnaire reflecting on their experience as evaluators. With regard to the quantitative questions, respondents indicated their answers on a 6-point scale to clearly indicate their degree of agreement or disagreement with the questions by giving measurable feedback. Generally, after each quantitative question, open-ended qualitative questions probing into reasons

and language skills were asked. Quantitative data analysis was integrated with the results from the open-ended questions to illuminate the findings. A total of 17 questionnaires from a class of 19 were returned for analysis. As for the quantitative questions, numbers as indicators of respondents' opinions were calculated for the means and standard deviations. To achieve reliability for qualitative comments, a colleague of the author with years of experience in teaching and research work was invited to interpret the comments of five randomly selected questionnaires out of the total of seventeen. As the comments were interpreted and recorded in the same way by the author and the inter-rater, and we chose one qualitative question with the most comments to work out the table for it. We agreed upon the headings used for the comments and the grouping of all the comments.

Table 1 summarizes the ranking of the means of students' responses to different quantitative questions. It shows that the students enjoyed the experience of peer evaluation. The evaluation and comments were favourably viewed as the means of related questions ranged from 4.59 to 3.82 and took up the top positions in the table.

(Evaluatee)	I think that the evaluation and comments given are fair.	4.59
(Evaluator)	I enjoyed the seminar as an evaluator	4.24
(Evaluatee)	I think that the comments given are useful for making improvement.	4.12
(Evaluatee)	I enjoyed the seminar as an evaluatee.	3.94
(Evaluatee)	I think that the comments given are sufficient.	3.82
(Evaluatee)	I performed better than I had expected.	3.67
(Evaluator)	I was being lenient.	3.65
(Evaluator)	I was a competent evaluator.	3.35
(Evaluatee)	Who do you prefer to be the evaluator: your teacher or your classmate?	2.65

Table 1: Ranking means of quantitative questions in descending order

4.1   Students as evaluatees

The respondents generally expressed the view that the seminar raised their self-awareness in speaking, enhanced their understanding of teamwork and its importance, and increased their confidence in the open-ended questions. Table 1 and Table 2 show that generally the respondents were positive about their experiences as evaluatees (mean=3.94). Almost all respondents (mean=4.59) considered the evaluation and comments given fair. Moreover, quite a number of the respondents found the comments useful for improvement (mean=4.12) and sufficient (n=11, mean=3.82).

The students enjoyed the seminar as evaluatees as they had the chance to speak and practise (55% of comments) and the relaxed atmosphere was viewed as a favourable factor (27%) as shown in Table 3. This suggests that peer evaluation does not affect the students' perception of the spoken task as primarily an opportunity for formal practice and learning, and it helps to reduce anxiety with its element of collaborative learning.

As is shown in Table 2, two-thirds of the respondents thought they performed slightly better than they had expected (mean=3.67). They reported that they made more contribution in the seminars by speaking more and being relaxed, the seminars went smoothly and the time management was good. On the other hand, the remaining five respondents were self-critical and identified a number of aspects they considered they had performed unsatisfactorily. These self-criticisms were mainly related to speaking ability such as lack of fluency in speaking, inadequate vocabulary, unclear development of ideas, with nervousness coming second.

	Strongly disagree (%) 1	Quite disagree (%) 2	Slightly disagree (%) 3	Slightly agree (%) 4	Quite agree (%) 5	Strongly agree (%) 6
I enjoyed the seminar as an evaluatee. (mean=3.94, SD=0.97)	0	6	29	29	35	0
				(n=11)		
I performed better than I had expected. (mean=3.67, SD=0.72)	0	7	27	60	7	0
				(n=10 out of 15)		
I think that the evaluation and comments given are fair. (mean=4.59, SD=0.87)	0	0	6	47	29	18
				(N=16)		
I think that the comments given are sufficient. (mean=3.82, SD=1.13)	6	0	29	41	18	6
				(n=11)		
I think that the comments given are useful for making improvement. (mean=4.12, SD=1.11)	0	6	24	35	24	12
				(n=12)		

Table 2: Students' perceptions as evaluatees (n=17)

More chance to practise and discuss. (could discuss, speak continuously, could use what I learned)	55%
Relaxed atmosphere (felt less pressure, liked working with classmates, group work, learned not because of exam or test)	27%
The situation was formal.	9%
We had good preparation	9%

Table 3: Reasons for enjoying the seminar as evaluatees

Table 4 shows the types of comments given by the evaluators to their peers. Slightly over one-third of the comments were about pronunciation for the obvious reason that the task was seminar discussions, followed by turn-taking skills and content. Only 7% of the comments given were on grammar and vocabulary. Hence, not surprisingly, when asked about the types of com-

ments desired by evaluatees, the respondents wanted comments on grammar most, followed by pronunciation and vocabulary as analysed in Table 5.

Pronunciation and fluency	38%
Mispronunciation and stress	25%
Sound volume and pace	13%
Turn-taking skills and responsiveness	21%
Content	19%
Content contribution and degree of participation	15%
Overall organization	4%
Body language	23%
Accuracy and vocabulary	7%
Grammar	4%
Use of phrases	3%

Table 4: Types of comments received by evaluatees

Grammar	38%
Pronunciation	19%
Range of vocabulary	19%
Clarity and development of ideas	13%
Fluency and intonation	6%
Presentation skills	6%

Table 5: Types of comments desired by evaluatees

Table 6 reveals that more than two-thirds of the respondents (n=12) chose the teacher to be the evaluator in this low-stakes peer evaluation and the majority stated that the teacher would give more professional suggestions. Others referred to the expert knowledge possessed by the teacher and reported that they could learn more from the mistakes identified by the teacher. The reasons for opting for peers were it would be more relaxing and more efficient as it saved time as shown in Table 7.

	Teacher (%) 1	(%) 2	(%) 3	(%) 4	(%) 5	Classmate (%) 6
If the seminar remains not counted towards the final grade of this module, who do you prefer to be the evaluator: your teacher or your classmate? (mean=2.65, SD=1.37)	24	29 (n=12)	18	18	12	0

Table 6: Students' perceptions as evaluatees on choice of evaluator (n=17)

Teacher	Classmates
Teacher gives more pieces of professional advice.	It is more relaxing.
Teacher has more knowledge of English	It is more efficient, saves time.
Teacher is quick to find out my mistakes.	
I can learn more concerning grammatical mistakes and pronunciation.	

Table 7: Reasons for choosing teacher or classmate

4.2   Students as evaluators

Table 8 shows positive responses from the students on their experiences as evaluators with a mean of 4.24 enjoying the seminar as evaluators. The respondents enjoyed being the evaluators for various reasons as summarized in Table 9, such as simply liking observation and listening (n=5), learning from others' mistakes (n=4). Two evaluators enjoyed exercising the power as decision makers or evaluators in the classroom and learned how to give comments to peers. A couple of responses showed that making use of the marking form as evaluators could help them 'be smarter' in other seminars. One evaluator mentioned that it was a good experience to evaluate from a fair and objective point of view. With regard to the negative aspects of their experience as evaluators, some responses (n=4) mentioned that it was hard to concentrate, listen and evaluate at the same time. A few (n=4) pointed out that little participation or preparation from the appraised students could dampen their enthusiasm as evaluators, and two responses remarked that they could not point out all the mistakes made by the peers as they were of similar language level (Table 10).

	Strongly disagree (%) 1	Quite disagree (%) 2	Slightly disagree (%) 3	Slightly agree (%) 4	Quite agree (%) 5	Strongly agree (%) 6
I enjoyed the seminar as an evaluator. (mean=4.24, SD=1.01)	0	6	12	53	18	12
				(n=14)		
I was a competent evaluator. (mean=3.35, SD=0.93)	6	12	24	59	0	0
				(n=10)		
I was being lenient. (mean=3.65, SD=0.80)	0	12	24	59	6	0
				(n=11)		

Table 8: Students' perceptions as evaluators (n=17)

Like observation and listening	(n=5)
Learn from others' mistakes	(n=4)
Enjoy exercise power as decision makers	(n=2)
Learn from using the marking criteria	(n=2)
Good experience to evaluate fairly and objectively	(n=1)

Table 9: Reasons for enjoying the seminar as evaluators

Hard to concentrate, listen and evaluate simultaneously	(n=4)
Unable to point out all mistakes due to similar language standard	(n=2)
Bored because little preparation and participation of the evaluatee	(n=1)
A bit boring	(n=1)

Table 10: Reasons for not enjoying the seminar as evaluators

On a scale of harshness to leniency from score 1 to 6 as an evaluator, the mean of 3.67 (Table 8) suggests that the student evaluators considered themselves slightly lenient but rather fair as peer evaluators. This result is supported by qualitative comments as seven students responded that they were serious and responsible in helping their peer improve their inadequacies. Two among them gave unsatisfactory grades to their peers objectively based on their unsatisfactory performance. Three students preferred not to be harsh because they could not be sure if the peers made a mistake as they were of similar standards. Four students ranked themselves to be slightly lenient on the grounds like other classmates not being harsh (n=2), not liking to give out low marks to classmates (n=1) and leniency helping them learn English (n=1). As the design of this question did not provide the mid-point option in order to get a partial picture of student evaluators' fairness, the near mid-range mean of 3.64 and attitudes of the students reflected in open-ended responses indicate that the students tried to be just in evaluation; and this was supported by the data that the overwhelming evaluatees considered the comments fair.

More than half (n=10) of the respondents considered themselves as competent evaluators. The mean of all responses to this question was 3.35, which suggests that they slightly agreed that they were competent evaluators. This might be due to the fact that a few respondents found evaluating a seminar discussion an attention-demanding task, and thought that their speaking ability was similar to that of the evaluatees. This was supported by the areas of improvement to be more competent evaluators suggested by the evaluators themselves: mainly English knowledge of grammar and vocabulary and attention to the evaluatees and discussion as listed in Table 11.

English knowledge, attention to grammar and range of vocabulary	32%
Attention to the evaluatee and discussion	21%
Be more serious, put in effort	16%
Listening, be able to catch the points and write comments	16%
Comments provide details, more advice	11%
To learn how to grade	5%

Table 11: Areas for improvement to be more competent evaluators

5 Discussion

It is indicated in this study that the students positively view the experience of peer evaluation of a seminar as it can enhance their group work, confidence and provide a chance for peer learning. Peer evaluation also gives the evaluators the opportunity to exercise power by using the marking criteria, which to a certain extent helps the students to understand the seminar and the requirements better.

The students generally think that the evaluation and comments given are fair, useful and sufficient for making improvement. Although the students do not receive much explicit training on evaluating their peers, they see themselves as fairly competent evaluators. It seems that the students can learn to master the skills through learning the seminar skills and examining the criteria and evaluation methods in the class, and at the same time implicitly cultivate the ability to observe the presence or absence of the same skills in other learners. This points out the possibility that no or little set-apart evaluation training session is needed for the evaluation of low-stakes seminar discussions. In addition, the findings seem to echo Falchikov's (2001) findings that students on advanced courses tend to be more accurate assessors, since the course of English for Academic

Purposes is rather like an advanced course of General English Language in secondary education with its specific stress on academic context.

The students are slightly lenient as they are reluctant to give adverse ratings and show empathy for their peers. This attitude might not have significantly affected the reliability of the grades, because teachers, when grading, might avoid giving low grades as well or try to be fair when considering the emotional responses of the students. Moreover, as Gopinath (1999, p. 14) suggests that the "benefits of peer assessment extended beyond the question of reliability of the grade," this slight leniency might be acceptable considering the richness of comments and different perspectives of evaluation generated by sharing evaluation authority and the optimum use of classroom time.

Given that the seminar is a low stakes peer evaluation, the students slightly agree that they are competent evaluators, and some find it difficult to pay attention to every part of the 20-minute seminar and give comments concomitantly. They tend to give comments mostly on macro-skills such as interaction skills, content and body language, and few are able to give elaborate or sufficient comments on grammar and vocabulary, which are desired most by the evaluatees to stimulate reflection and change to polish their language skills to a more sophisticated level. This may be due to the lack of experience of the students in peer evaluation and lack of confidence in assessing proficiency of a second language. In spite of the value of the students' evaluation and comments, this study shows that the students as sole evaluators are well aware of some areas of inadequacy. Hence, the students prefer the teacher with recognized breadth of knowledge as a better person to evaluate the seminar discussion if they can choose between a teacher and a peer.

Some research shows that peer evaluation can be reliable substitutes of teachers' grade and suggest that teachers give up grading to do other more educational activities. This study suggests that although the students are generally fair and competent evaluators, and the peers consider the comments useful and the experience enjoyable and useful, the students are concerned about the depth and quality of their work and cherish the proficient and elaborate comments from teachers who possess knowledge and training in the subject. While the inclusion of peer assessment may release the energy and effort of teachers from assessing some aspects of a task, teachers still have an important part to play in assessment.

Most students usually expect good grades but in reality are usually disappointed. In the teacher-only evaluation, it is quite an emotional struggle for students to accept the undesirable grades and the exhortation for improvement from the very same teachers at the same time. This study suggests that as students are capable of evaluating without the need for substantial time investment, teachers can change their role to become partners in grading and present themselves in a favourable light as helpers or guides to students by offering their insightful comments to raise the standard of student performance to a higher level, especially on areas which require professional insight and expertise. Evaluation or assessment is the beginning of another phase or process of learning a task. The enhancement or continuation of student learning after evaluation is as important as the learning before the evaluation. It seems that the richness of critical comments jointly provided by both students and teachers can stimulate students to continue learning after evaluation. As peer evaluation can give learners immediate feedback, the teacher's elaborate feedback may be given later due to class size and time constraint.

Teachers as a learning resource and holder of expertise can focus on giving constructively critical comments and pinpointing areas for improvement. In addition, teachers as managers of assessment and the learning environment inside and outside the classroom can create a supporting and encouraging environment for peer evaluation. Since assigning the whole evaluation work to the students might stir some uncertainty and anxiety in the students, they can assign the students to assess part of the task or let them choose the areas which they might feel capable of or comfortable about managing based on results of rating exercises or checklists in class. They can complement student evaluation by discussing with them the division of evaluation workload and the weighting of different aspects of a task with agreed criteria. In addition, teachers can offer aid and provide skills training to students so that they can gain confidence from participating in the process and satisfactorily complete peer evaluation. For instance, this research shows that the students desire

many instructive comments on language accuracy. The teacher may assist the students by providing a checklist with possible types of grammatical errors and examples for the evaluators to select to enhance the depth and value of their comments. Finally, student evaluators may have self-doubt and reservations about new evaluation procedures, or encounter inter-personal issues like the face problem or free-rider effects, teachers as advisers and friends can listen to students on personal and group issues to develop student evaluators and promote collaborative learning. Hence, "peer evaluation should not be seen as opting out by the tutor, who still has responsibility for the final mark, however it is derived" (Gregory & Thorley, 1994, p. 182). Teachers, therefore, may sometimes be learners alongside the students to find out the optimal evaluation process with easy-to-follow and accurate guidelines for students and give students incentive to exert their autonomy, at the same time maintain constant communication with them.

6   Conclusion

Since this study is a small scale research on 17 undergraduates into seminar discussions and only basic calculations like averages and simple counts are used, generalization is quite limited. However, it is interesting to analyse opinions from the same subjects playing the dichotomy of roles: the evaluator and evaluatee. Also, the qualitative and quantitative data support one another pointing towards the value of peer evaluation on generating timely feedback and making the best use of limited teaching and evaluation time. Also this study shows how students are intrigued and challenged by the experience, how teaching students to master certain skills can simultaneously facilitate appraisal of the same language skills taught, and the desirability of teachers' expertise in the eyes of students who have experienced the power or demand of being an evaluator. It is also revealed that the students require help to be competent evaluators and at the same time are eager to improve their work or performance in the future. Research needs to be undertaken to divulge students' needs as peer evaluators, the actual difference between student and teacher comments in depth and areas of spoken language learning, and how the best practice of sharing evaluation between teachers and students can be achieved especially in attention demanding spoken tasks. If extra time can be gained by teachers through student observation, they may need to work together to ensure that each student can receive individualized insightful feedback to raise their achievement to another level. To the author, from the collected data and brief chats from a couple of students, it appears that students' major concern is primarily on improving their work in the future with instructive comments, and thus the teachers cannot hastily let go of the responsibility in assessments or evaluation without considerable research support.

References

Ainsworth, L., & Viegut, D. (2006). *Common formative assessments*. California: Corwin.

Airasina, P.W. (1997). *Classroom assessment*. New York: McGraw-Hill.

Bligh, D.A. (2000). *What's the point in discussion?* Exeter: Intellect.

Brown, G., Bull, J., & Pendlebury, M (1997). *Assessing student learning in higher education.* London: Routledge.

Brown, H.D. (2004). *Language assessment: Principle and classroom practices.* New York: Longman.

Campbell, K.S., Mothersbaugh, D.L., Brammer, C., & Taylor, T. (2001). Peer versus self assessment of oral business presentation performance. *Business Communication Quarterly, 64*(3), 23–42.

Carson, J.G., & Nelson, G.L. (1996). Chinese students' perceptions of ESL peer response group interaction. *Journal of Second Language Writing, 5*(1), 1–19.

Catterall, M. (1995). Peer learning research in marketing. In S. Griffiths, K. Houston & A. Lazenblatt (Eds.), *Enhancing student learning through peer tutoring in higher education: Section 3 – Implementing* (pp. 54-62). Coleraine, NI: University of Ulster.

Crooks, T.J. (1988). Impact of classroom evaluation on students. *Review of Educational Research, 58*(4), 438–481.

Elliott, N., & Higgins, A. (2005). Self and peer assessment – does it make a difference to student group work? *Nurse Education in Practice, 5*, 40–48.

Falchikov, N. (2001). *Learning together: peer tutoring in higher education*. London: Routledge Falmer.

Gopinath, C. (1999). Alternatives to instructor assessment of class participation. *Journal of Education for Business, 75*(1), 10–15.

Gregory, R., & Thorley, L. (1994). Present challenges. In: R. Gregory & L. Thorley (Eds.), *Using group-based learning in higher education* (pp. 179–186). London: Kogan Page.

Havnes, A., & McDowell, L. (2004). Introduction: Assessment dilemmas in contemporary learning cultures. In A. Havens & L. McDowell (Eds.), *Balancing dilemmas in assessment and learning in contemporary education* (pp. 3–14). London: Routledge.

Inoue, A.B. (2005). Community-based assessment pedagogy. *Assessing Writing, 9*, 208–238.

Jacobs, G., & Zhang, S. (1989). Peer feedback in second language writing instruction: Boon or bane? *Paper presented at the Annual Meeting of the American Educational Research Association*. San Francisco, CA: March 27–31, 1989.

MacAlpine, J.M.K. (1999). Improving and encouraging peer assessment of student presentations. *Assessment and Evaluation in Higher Education, 24*(1), 15–25.

Mitchell, V.W., & Bakewell, C. (1995). Learning without doing – Enhancing oral presentation skills through peer review. *Management Learning, 26*, 353–366.

Nitko, A.J., & Brookhart, S.M. (2007). *Educational Assessment of students*. Upper Saddle River, NJ: Pearson/Merrill Prentice Hill.

O'Donnell, A.M., & Topping, K.J. (1998). Peer assessing peers: Possibilities and problems. In K.J. Topping, & S.W. Ehly (Eds.), *Peer assisted learning* (pp. 255–278). Mahwah, NJ: Lawrence Erlbaum.

Race, P., & Brown, S. (1988). Changing assessment to improve learning. In P. Race (Ed.), *The lecturer's toolkit* (pp. 51–78). London: Kogan Page.

Rendle-Short, J. (2006). *The academic presentation: Situated talk in action*. Hamphire: Ashgate Publishing Limited.

Smyth, K. (2004). The benefits of students learning about critical evaluation rather than being summatively judged. *Assessment and Evaluation in Higher Education, 29*(3), 369–378.

Stanley, J. (1992). Coaching students writers to be effective peer evaluators. *Journal of Second Language Writing, 1*(3), 217–233.

Stefani, L.A.J. (1992). Comparison of collaborative self, peer and tutor assessment in a biochemistry practical. *Biochemical Education, 20*(3), 148–151.

Segers, M., & Dochy, F. (2001). New assessment forms in problem-based learning: The value-added of the students' perspective. *Studies in Higher Education, 26*(3), 327–343.

Strom, P.S., Strom, R.D., & Moore, E.G. (1999). Peer and self-evaluation of teamwork skills. *Journal of Adolescence, 22*, 539–553.

Topping, K. (2003). Self and peer assessment in school and university: Reliability, validity and utility. In M. Segers, F. Dochy & E. Casacallar (Eds.), *Optimising new modes of assessment: In search of qualities and standards* (pp. 55–87). London: Kluwer Academic Publishers.

Appendix

Seminar Discussion: Peer Evaluation
Evaluator:_____   **Evaluatee:**_____

	Very Good	Good	Average	Poor	Very Poor
Content					
Quantity of ideas					
Clarity and development of ideas					
Interactive skills					
Turn-taking skills					
Responsiveness: active/passive					
Accuracy					
Grammatical structures					
Vocabulary					
Pronunciation					
Pronunciation of words					
Intonation: word stress					
Fluency: connected / broken speech					
Referencing: refer to research, sources					
Body language: eye contact, posture, gesture					
Overall performance:					
Comments on accuracy:					
Comments on pronunciation:					
Other comments:					