Improving Peer and Self-assessment for Group Presentations from Chinese Students' Perspective

Caroline T. W. Chan

***Abstract*—Almost everyone agrees that student presentations benefit the students in significant ways. That is why presentation is often required as part of coursework. However, the teachers who implement presentations should experience much problem of how to get the rest of the class "listen" to others' presentations. Without listening actively to the presenters, the audience loses a valuable chance to learn from their peers' work. Although engaging students to assess their peers' work is suggested as a possible method to get students listening and learning from presentations of others, our understanding of the students' perceptions in the context of peer assessment for group presentations is limited. Through a questionnaire survey with 158 engineering sub-degree students in Hong Kong, this study aims to collect primary data on peer assessment for group presentations. The findings are useful to design and develop a user-friendly system to actively engage students as co-assessors for group presentations in the peer assessment process. The findings should provide useful insights to the teachers and researchers, helping them to design an effective assessment tool for group presentations.**

***Index Terms*—Peer assessment, self-assessment, student presentations.**

I. Introduction

In the commercial and professional fields, many competitions and tender evaluations include interview presentations as part of the assessment process. The widespread use of presentations drives us to put more emphasis in training students the required generic skills at schools. Nowadays, student presentation is often an integral component in the context of assessment in higher education.

By doing presentations, students can practise public speaking which is an important generic skill. More importantly, students can learn how to prepare and showcase their work to the class and the teacher. Most teachers who have students do presentations in groups believe that students can learn from listening the presentations, not only from presenting. However, those teachers who implement presentations should experience much problem of how to get the rest of the class "listen" to the others' presentations. The learning potential of student presentations, not just for the presenter, but for the audience, is huge. However, such benefit in most time is not realized.

Manuscript received February 20, 2017; revised May 3, 2017. This work was supported in part by the Teaching Development Grant 2013-14 of the City University of Hong Kong (Project No. 6000498).

Caroline T. W. Chan is with the City University of Hong Kong, Kowloon Tong, Hong Kong (e-mail: caroline.chan@cityu.edu.hk).

II. Potential Benefits of Peer and Self-assessment

Assessment by students is suggested as one of the possible methods to get students listening and learning from the presentations of others [1]. Peer assessment and self-assessment are often applied together as a formative assessment tool that engages students in the assessment process [2], [3]. Falchikov advocates that involving students in the assessment of presentations is "extremely beneficial" to students [4]. The peer and self-assessment process can develop students' self-regulating skills when they analyse their own behaviour.

Learning from peers is considered to be one of the most potentially rewarding teaching and learning methods for teachers and students alike [5]. Numerous pedagogical studies identified the benefits of applying peer assessment in higher education, such as enhancing collaboration, achieving learning goals, sharing of responsibility, encouraging autonomous learning and practicing transferrable skills [3], [6]-[8]. Some past studies evidence that students can develop these positive traits when assessing their peers [9].

Careful planning of the peer assessment, with the integration of feedback (from peer assessment) and reflection (from self-assessment), can lead to deep learning [10]. In case of student presentation, bringing students to fully engage in the assessment process as co-assessors can provide valuable formative and summative feedback to presenters. The entire process can embed a deep learning approach which promotes creativity and knowledge discovery.

III. Concerns of Peer and Self-assessment

Despite of the potential benefits of peer assessment in student presentations, many problems associated with conventional peer assessments, such as peer pressure [11], unhelpful competition [2], personal bias and lack of interest among students [12] are waiting to be solved. Without proper guidelines for teachers and students to follow, using conventional peer assessment in presentations may be viewed as an extra assignment which merely aims to save the teacher's work [12]. In that case, students will not listen to presentations seriously and are not likely to have any sort of learning experience. The presenters will also have great difficulty to present in front of a crowd that pays no attention. As a result, the academic benefits of presentation will be substantially impeded.

doi: 10.18178/ijiet.2018.8.3.1034

IV. The Need for Study

To exploit the potential benefits of peer and self-assessment for presentation assessment, an appropriate design of the assessment processes is paramount. Making the assessment tasks "friendly to peer learning" is the key to achieve satisfactory results [2], and the assessment tool design should be compatible with the characteristics of the coursework such as written reports, oral presentations or group projects. However, existing peer assessment tools mainly focus on the preparation of summative assessment without much consideration in the process design for oral presentation assessment. Among the wealth of studies about peer and self-assessment in higher education, many have been done about the validity and reliability of peer assessments. Most of them suggested a high consistency among peer marks [13], [14] and a high correlation between tutor's marks and peer marks [15]-[18]. There is little in the published literature on how to conduct the assessment for group presentations, when intra-group peer and self-assessment (assessment of the individual's performance within a group) and inter-group peer assessment co-exist. Besides, some studies suggest that Chinese students are shy and reserved to give comments [11]. To design a peer assessment tool for Chinese students, a better understanding of the Chinese students' perceptions is indispensable. The study reported here examines the concerns and the preferred design of peer assessment from the students' perspective, which contributes to an under-researched aspect of peer assessment.

V. Objectives

Using questionnaire survey as a primary data collection method, this study explores the appropriate design of peer assessment tools for group presentations. The aim is to address the concerns and desired features of peer assessment for group presentations among Chinese students. The specific objectives of this study are to evaluate the perceived benefits and concerns of peer assessment for presentations among Chinese students, and to determine the desired features of a peer assessment tool for group presentations. Another objective is to identify any possible gender differences in peer assessment perceptions among Chinese students. The findings will be useful to develop an effective and user-friendly assessment tool that can fully engage students in the co-assessment process, so as to enhance student learning in group presentations. Also, the results generated from this study will be helpful for the design of peer assessment tools that are applicable to Chinese students.

VI. Method

A. Data Collection

Primary data were collected by distributing questionnaires to all year one and year two students who enrolled in the Construction Engineering and Management (CEM) programme at the Associate Degree Level at a university in Hong Kong during 2014-15. This survey was conducted with the approval of the Research Ethics Sub-Committee of the university. Participation in the survey was voluntary. The questionnaire was administered in hardcopy during a class.

An initial pilot study was conducted using a convenience sample of 10 students enrolled in one course at the university. Based on the feedback from the pilot study, minor revision to the wording of two questions was made in the questionnaire. The revised questionnaire was then adopted in the survey.

B. Questionnaire Design

A questionnaire was developed to collect students' perceptions about peer and self-assessment for group presentations. The questionnaire was divided into three parts to address the objectives of this investigation. The first section was related to the demographics of respondents. The second part included eight questions about students' perceptions of the benefits of group presentations and peer assessment. The third part contained twelve questions about the concerns of peer and self-assessment in the context of group presentations. Respondents were requested to rate their agreement against each question in these two parts according to a five-point Likert scale from 1 (which represented strongly agree) to 5 (which represented strongly disagree). The fourth part consisted of fifteen questions related to the preferred design of a peer and self-assessment tool for group presentations. Six out of fifteen questions were answered on a five-point Likert scale based on the level of agreement to which the respondent opined. The rest of them were open ended to understand the views of the students.

VII. Data Analysis

The survey data were analyzed using SPSS 20.0. A reliability test of the data was first conducted followed by descriptive analysis.

A. Cronbach's Alpha

Cronbach's alpha (α) is a measure of the internal consistency or reliability of data collected for questions within each category (Cronbach, 1990). The α value ranges from 0 (perfectly independent) to 1 (perfectly correlated). The higher value of α implies that the questions within a category are reliable and consistently measure the defined construct.

Within the questionnaire, there are four areas of questions, including perceived benefits of group presentations (GP_b), perceived benefits of peer assessment (PA_b), concerns of students about peer assessment (CON) and preferred design of peer assessment for group presentations (PA_d). Cronbach's alpha test was applied to the first three sets (GP_b, PA_b and CON) in which all questions in each set measure a single construct.

Having assessed the internal consistency of the instrument and the profile of the respondents, the data were then analyzed using descriptive analysis.

B. Mean Score

The students were divided into two groups based on gender for analysis. A five-point Likert scale was used to collect students' perceptions. The mean score for each Likert scale question was computed using the equation as shown below (1):

$$\bar{s}' = \frac{\sum(f \times s')}{N} \quad (1)$$

where

= mean score

= score rated to question by respondents, ranging from 1 to 5 (1 = strongly agree and 5 = strongly disagree)

f = frequency of responses to each rating

N = total number of responses concerning each question

C. ANOVA of Gender Differences

The significance of gender differences on the perceived benefits and concerns of peer assessment was of interest in this study. An appropriate test to investigate the existence of significance is an analysis of variance test (ANOVA) [19]. A one-way ANOVA was used to evaluate the gender differences in the mean scores of perceived benefits and concerns in relation to peer assessment.

VIII. Results

The number of students enrolled in the CEM programme during 2014-15 was one hundred and eighty-three. One hundred and fifty-eight valid responses were received. The response rate was 86.3%. The respondents' profile is shown in Table I.

TABLE I: Respondents' Profile

	Male	Female	Total
	65	17	82
	52	24	76
Total	117	41	158

A. Cronbach's Alpha

As a general rule, the criteria of acceptability for Cronbach's reliability coefficients is a minimum of 0.60 [20]. From Table II, the coefficients of GP_b, PA_b and CON groups are above 0.60, indicating that the items within these groups presented an acceptable level of internal consistency. The alpha coefficient of the GP_b group was above 0.70, representing high internal consistency. The grouping of question items into benefits of group presentations, benefits of peer assessment and concerns of peer assessment for analysis was thus statistically acceptable.

TABLE II: Cronbach's Alpha of Each Question Group

	Question Group	Cronbach's alpha	Number of items
1	Benefits of group presentations (GP_b)	0.741	4
2	Benefits of peer assessment (PA_b)	0.617	4
3	Concerns of peer assessment (CON)	0.691	12

B. ANOVAs

A one-way ANOVA of gender differences on the overall perceived benefits of group presentations (GPb), perceived benefits of peer assessment (PAb) and concerns of students when conducting peer assessment (CON) was conducted. As illustrated in Table III, the test showed no significant gender differences in the perceived benefits ($p > 0.05$). However, there is a statistically significant difference ($p = 0.046$) between male and female students with respect to the concerns of peer assessment.

TABLE III: ANOVA of Gender Differences

Benefits of Group Presentations (GP_b) by Gender					
	N	Mean	Standard Deviation	Univariate F	p-value
Male	117	3.479	0.708	0.214	0.644
Female	41	3.421	0.634		
Benefits of Peer Assessment (PA_b) by Gender					
	N	Mean	Standard Deviation	Univariate F	p-value
Male	117	3.665	0.591	2.884	0.091
Female	41	3.476	0.673		
Concerns of Students (CON) by Gender					
	N	Mean	Standard Deviation	Univariate F	p-value
Male	117	2.896	0.453	3.421	0.046
Female	41	3.047	0.437		

C. Mean Scores of Perceived Benefits of Group Presentations and Peer Assessment

Since the ANOVA showed no significant gender differences in the perceived benefits, the mean scores from the two gender groups are not compared. Table IV and Table V report the mean scores for the group presentation benefits and the peer assessment benefits respectively, with the items arranged in descending order of their mean scores.

TABLE IV: Mean Scores of Perceived Benefits of Group Presentations

Rank		Question	Mean score	Standard Deviation
1	GP_b1	Presentation is an effective process to practise / improve presentation skills.	3.684	0.945
2	GP_b2	Other groups' presentations are important to my study.	3.538	0.864
3	GP_b3	I can learn technical knowledge from other groups' presentations.	3.373	0.913
4	GP_b4	I can learn presentation and communication skills from other groups' presentations.	3.259	0.946

TABLE V: Mean Scores of Perceived Benefits of Peer Assessment

Rank		Question	Mean score	Standard Deviation
1	PA_b1	Assessment of group members' contribution is useful in monitoring group work and individual's contribution.	3.646	0.830
2	PA_b2	Peer assessment can improve my generic skills.	3.639	0.876
3	PA_b3	I want to know my performance from my classmates' perspective.	3.614	1.087
4	PA_b4	Peer assessment can improve my academic learning.	3.563	0.794

The mean scores of all the perceived benefits were above 3.0 (ranged from 3.259 to 3.684), which confirmed that students were positive towards the academic benefits of oral presentations and peer assessment. In general, the ratings of the group presentation benefits were comparable to the peer

assessment benefits. Students concurred that oral presentations could improve their presentation skill and technical knowledge. When assessing their peers' presentations, students could also learn from others.

D. Mean Scores of Students' Concern

For the concerns about peer assessment, the mean scores ranged from 2.463 to 3.634 (see Table VI). Out of the twelve concerns, five items scored above 3.0 in the overall means. These items include: *Students are often biased when assessing his/her contribution to the group project* (CON1), *I worry that if I gave low marks or negative comments to the other groups, they might give low marks or negative comments to me as well* (CON2), *I feel both power and pressure when I assess my classmates* (CON3), *I worry that if I gave low marks or negative comments to the other groups, my friendship with peers would be adversely affected* (CON4) and *Most students do not do the peer assessment fairly and responsibly unless it is scored* (CON5). All these items are related to the concern of bias from peers.

TABLE VI: MEAN SCORES OF STUDENTS' CONCERN (BY GENDER)

	Question	Mean scores		
		Overall	Male	Female
1	Students are often biased when assessing his/her contribution to the group project. (CON1)	3.633	3.632	**3.634**
2	I worry that if I gave low marks or negative comments to the other groups, they might give low marks or negative comments to me as well. (CON2)	3.430	3.376	3.585
3	I feel both power and pressure when I assess my classmates. (CON3)	3.215	3.145	**3.415**
4	I worry that if I gave low marks or negative comments to the other groups, my friendship with peers would be adversely affected. (CON4)	3.177	3.197	3.122
5	Most students do not do the peer assessment fairly and responsibly unless it is scored. (CON5)	3.044	*2.983*	**3.220**
6	Students do not have sufficient knowledge and skill to assess their peers. (CON6)	*2.981*	*2.966*	**3.024**
7	The presentations of other groups are not interesting. (CON7)	*2.734*	*2.684*	***2.878***
8	I feel uncomfortable to assess the academic contents of the other groups' presentations. (CON8)	*2.690*	*2.658*	***2.780***
9	I feel uncomfortable to assess my group members' contribution to a group project. (CON9)	*2.652*	*2.700*	*2.512*
10	I feel uncomfortable to assess the presentation skills of other groups' presentations. (CON10)	*2.627*	*2.538*	***2.878***
11	I feel uncomfortable to have my classmates assessed my presentation. (CON11)	*2.601*	*2.444*	**3.049**
12	I cannot assess fairly my contribution to the group project. (CON12)	*2.437*	*2.427*	***2.463***

In general, students agreed that *student's bias when assessing his/her contribution to the group work* (CON1) is the most problematic. Interestingly, *I cannot assess fairly my contribution to the group project* (CON12) scored the lowest in both groups, with an average of 2.437. This indicates that students generally believe that they can assess their contribution fairly but their peers cannot. When considering whether the students have sufficient knowledge and skill to conduct peer assessment (CON6), both groups of students rated moderately (male = 2.966 and female = 3.024). However, all students indicated a stronger disagreement (average mean scores ranged from 2.60 to 2.70) that they have psychological discomfort when doing peer assessment of others (CON8 - CON10). This implies that the students do not feel stressful when conducting peer assessment. Nevertheless, they are not fully confident with their ability to assess their peers.

As highlighted before, the ANOVA results indicated that there were significant gender differences in the concern (CON) construct. Across the twelve concern (CON) items, female students expressed stronger concern than male students in nine items (as shown in bold figures in Table 6). The largest gender difference is observed in item CON11: *I feel uncomfortable to have my classmates assessed my presentation*. The mean score of male students was only 2.444 (ranked 11th within the group) but that of female students was 3.049 (ranked 6th within the group).

E. Preferred Design of Peer Assessment Tools

There were six Likert-scale questions and nine open ended questions related to the preferred design of a peer assessment tool. The mean score and the number of students who responded above 3.0 in each Likert-scale question are presented in Table VII.

TABLE VII: MEAN SCORES OF PREFERRED DESIGN OF PEER ASSESSMENT

Item		Question	Mean score	Number of responses > 3.0
1	PA_d1	It is unfair if the same mark is given to all the members within the group.	3.051	54
2	PA_d2	If inter-group peer assessment is part of the coursework requirement, I prefer to have the peer assessment mark counted in my coursework mark.	2.854	49
3	PA_d3	Assessment results of member' and my contributions should be used to calculate individual's overall score.	3.184	61
4	PA_d4	Ranking my peers' performance is more accurate and easier than giving grades to them.	2.975	49
5	PA_d5	I need more practice to be confident in conducting peer and self-assessment.	3.247	73
6	PA_d6	Progressive peer and self-assessments can provide a better assessment than having one assessment conducted at the project-end only.	3.443	85

Results indicate that the students' readiness of including peer assessment result in the overall mark seems to be mediocre. Most students were neutral to have the same marks awarded to all group members in a group project (PA_d1). The mean score of item PA_d2, *If inter-group peer assessment is part of the coursework requirement, I prefer to have the peer assessment mark counted in my coursework mark*, was only 2.854, confirming that students do not want to have peer assessment results counted into their overall marks. Students are slightly more acceptive of having individual mark

calculated for group projects, based on the individual's contribution (PA_d3). There was no significant preference in either of the peer assessment methods (giving ranks or giving grades) for oral presentations (PA_d4). Students generally agreed that more practice would be desirable to enhance their confidence in conducting peer and self-assessment (PA_d5). In terms of assessment frequency, the participating students hold a positive attitude towards progressive assessments (PA_d6).

Students answered nine open-ended questions about their desired peer assessment tool attributes for group presentations. Table VIII to Table XVI summarise the frequencies and percentages collected from these open-ended questions.

As can be seen in Table VIII, the responding students suggested three main incentives to encourage active listening during oral presentations. Each incentive method was supported by around one-third of the respondents. In general, students tend to pay more attention to their peers' presentations if they are assessors and if the 'quality' of assessment output (e.g. quality of their comments made) relates to the assessors' course work mark.

TABLE VIII: FREQUENCIES AND PERCENTAGES OF STUDENTS' RESPONSES TO METHOD OF ASSESSMENT VS ACTIVE LISTENING

	How to encourage a student to listen to the other groups' presentations seriously?	Frequency	Percentage (N=158)
1	I need to set questions (that are related to the contents of presentations) to each group. The teacher will assess the quality of my questions and adjust my course work mark.	56	35%
2	I need to assess the academic content and presentation skill of each group, and my assessment results will contribute to the assessee's overall mark.	52	33%
3	I have to give formative comments to the other groups' presentations and marks will be added (or deducted) from my overall mark.	50	32%

Table IX summarises the preferred assessment format. *Give comments only to each presenting group* (item 1), which is a formative assessment, was recommended by 40% of the respondents. *Give an overall mark or grade* (item 2) and *rank the groups in the order of their relative performance* (item 3) are summative assessment, supported by 36% and 24% of students respectively.

TABLE IX: FREQUENCIES AND PERCENTAGES OF STUDENTS' RESPONSES TO ASSESSMENT FORMAT

	Format of peer assessment	Frequency	Percentage (N=158)
1	Give comments only to each presenting group	57	40%
2	Give an overall mark or grade to each presenting group	63	36%
3	Rank all groups in the order of their relative performance	38	24%

Regarding to the setting of assessment criteria, the majority of the students (70%) recommended that the assessment criteria should be set and agreed by students (see Table X). This result matches with the literature that the assessment should be made against students' own objectives [21].

Not surprisingly, regarding the question of whether or not the assessor's identity to be disclosed, the majority of the respondents (81%) concurred that the assessor's identity should be made known to the assessee (as in Table XI).

TABLE X: FREQUENCIES AND PERCENTAGES OF STUDENTS' RESPONSES TO CRITERIA-SETTING

	Assessment criteria-setting	Frequency	Percentage (N=158)
1	Assessment criteria to be set by the teacher	48	30%
2	Assessment criteria to be set and agreed by all students	110	70%

TABLE XI: FREQUENCIES AND PERCENTAGES OF STUDENTS' RESPONSES TO ASSESSOR'S IDENTITY

	Assessor's identity	Frequency	Percentage (N=158)
1	Assessor's identity is hidden	30	19%
2	Assessor's identity is made known to the one being assessed	128	81%

Referring to Table XII, despite the popularity of computer and mobile-based applications, only 59% of students preferred to conduct the peer assessment in a web-based environment. Pen and paper is still an effective tool for note-taking and thus welcomed by many students.

Table XIII shows the preferred method to collect peer assessment feedback from the students' perspective. Although a substantial amount of the responding students prefer to use pen and paper to perform the peer assessment, none of them suggested to collect feedback in paper form. Around 57% of the students prefer to collect the feedback through emails, whereas the rest of them prefer to check the results from the web.

TABLE XII: FREQUENCIES AND PERCENTAGES OF STUDENTS' RESPONSES TO ASSESSMENT TOOL

	Assessment tool	Frequency	Percentage (N=158)
1	Through a web-based platform	94	59%
2	By pen and paper	64	41%

TABLE XIII: FREQUENCIES AND PERCENTAGES OF STUDENTS' RESPONSES TO DISTRIBUTION OF ASSESSMENT RESULTS

	Distribution of assessment results	Frequency	Percentage (N=158)
1	Send the results to each student through email	90	57%
2	Check by students from the web-based platform	68	43%

From the results in Table XIV, half of the respondents preferred to finish the peer assessment within the lesson whereas the other half of the respondents suggested to have more time to complete the assessment.

TABLE XIV: FREQUENCIES AND PERCENTAGES OF STUDENTS' RESPONSES TO ASSESSMENT TIME

	Time of assessment	Frequency	Percentage (N=158)
1	During the presentation and before the end of lesson	77	49%
2	Can be done after the presentation lesson and within a set period	81	51%

The last two questions asked about the maximum number of groups and group size for peer assessment. Around 56% of the respondents opined that the maximum number of groups

should be 5 to 6 (see Table XV). For the group size as indicated in Table XVI, almost 70% of the students thought that 5 to 6 students are maximum.

TABLE XV: FREQUENCIES AND PERCENTAGES OF STUDENTS' RESPONSES TO MAXIMUM NUMBER OF GROUPS

	Maximum number of groups	Frequency	Percentage (N=158)
1	8 groups	34	22%
2	7 groups	11	7%
3	6 groups	44	28%
4	5 groups	45	28%
5	4 groups	13	8%
7	3 students	11	7%

TABLE XVI: FREQUENCIES AND PERCENTAGES OF STUDENTS' RESPONSES TO MAXIMUM GROUP SIZE

	Maximum group size	Frequency	Percentage (N=158)
1	10 students	12	8%
2	9 students	1	1%
3	8 students	8	5%
4	7 students	7	4%
5	6 students	61	39%
6	5 students	47	30%
7	4 students	22	14%

IX. DISCUSSION

A. Resistance to Peer Assessment for Group Presentations

Results in Table V indicate that Hong Kong students generally agree with the benefits of peer assessment. However, the idea of including the peer assessment result (the inter-group peer assessment mark) into students' final project mark is not much supported (PA_b2 mean score = 2.854). Whilst the assessment of members' contribution (the intra-group peer assessment) seems to have more support (PA_b3 mean score = 3.184), the mean score is much lower than those constructs that measured the perceived benefits of peer assessment (PA_b: mean scores ranged from 3.563 to 3.646).

Many past studies demonstrate students' resistance to peer assessment [22], [23]. The findings here suggest that the resistance of doing peer assessment is probably attributable to the 'bias from peers' as perceived by the students, rather than their feeling of stress. From Table 6, the students disagreed that they feel uncomfortable to conduct peer assessment (CON8 to CON11; mean score ranged from 2.690 to 2.601). Their indication of worries that their friendship or their peer assessment results awarded by others would be adversely affected (CON2 and CON3) is due to the distrust on others. From the study, the students felt that most of their peers are bias when conducting peer assessment (CON1 mean score = 3.633). Therefore, students are not confident to be assessed, both inter-group and intra-group, by their peers. Furthermore, there is no strong indication that the students think it is unfair if the individual student's mark is the same as that awarded to a group project (PA_d1 mean score = 3.051 in Table 7). As a result, students do not internalise the need for peer assessment.

B. Possibility of Personal Bias

Although many empirical studies advocate that peer assessment results are highly correlated with the teachers' grades [15]-[18], most students believe that their peers' assessments are not consistent or fair. This discrepancy brings a question whether there is any bias in the students' perception. As can be seen in Table 6, *students are often bias when assessing his/her contribution to the group work* (CON1) ranked the 1st among all concerns (mean score = 3.633). However, *I cannot assess fairly my contribution to the group project* (CON12) scored the lowest (mean score = 2.437). Students generally think that they are fair in the peer assessment but their classmates are bias. This perception can escalate students' resistance to peer assessment, which worth further investigation.

C. Gender Differences in Peer Assessment Concerns

Limited past studies on gender differences in peer assessment indicate that there is very little difference in the marks given by male assessors and female assessors [17], [24]. In this study, gender difference was evidenced as significant in the *concerns of peer assessment* (CON) construct (ANOVA $p = 0.046$) but not the perceived benefits (PA_b and GP_b) questions. Female students indicated higher concern level in nine out of twelve questions than male students (see Table 6). In general, we can conclude that the level of concern, or the negative attitude towards peer assessment, was less when compared males with females.

Such significant gender difference may be explained by Fitzpatrick's study that the females perceived everyone as less competent in peer assessment and demonstrated lower perceived self-efficacy than males [25]. Other past studies evidence that female engineering students possess lower self-efficacy perceptions than male students, which influence their behaviour negatively [26], [27]. In this regard, it is not surprising that female respondents felt uncomfortable when they were assessed by their classmates (CON11) (mean score = 3.049) but male students did not indicate such discomfort (mean score = 2.601). Female students also expressed higher level of discomfort when assessing their peers (CON8 and CON10). Although gender differences may not affect the reliability and consistency of the peer assessment results, the possible impact on females' resistance to peer assessment cannot be ignored. Skill training or mastery training is a widely considered as a proven strategy to improve self-efficacy [28]-[30], which can be considered as part of the curriculum.

D. Preferred Design of Peer Assessment Tools for Group Presentations

Based on the findings of the current study, some design features of the peer assessment tool are suggested for group presentations.

At the very beginning, the teacher should explain the objectives and advantages of peer assessment to the class clearly to increase their acceptance and involvement. Also, the assessment criteria are better to be set and agreed by students. Trial or training should be arranged to the class prior to the assessment process. This can improve their understanding and confidence in peer assessment.

Inter-group peer assessment that contributes to students' overall marks can encourage students listen to their peers' presentations. Teachers can ask the students to give a grade

or a mark, or to rank the groups according to their relative performance. Formative feedback such as comments or questions to presenters is very welcomed. Whilst peer assessment should contribute to the overall mark for the group project, it is recommended to limit the contribution to a small percentage in order to reduce the students' resistance. Two intra-group assessments on group members' contribution can be implemented, one in the mid-way and the other one at the end, to motivate the students to work harder after receiving the first peer assessment result.

The choice of assessment tool (either web-based or paper-based) may not be very critical to the students. However, from the teacher's point of view, a web-based tool can facilitate the administration of marks and feedback. Assessors' identity should be made known to the assessees. This helps to improve the mutual trust in the assessment process. Students prefer to receive feedback in the electronic format. Therefore, the feedback should be either sent through email or accessed through the web.

The peer and self-assessment should be completed timely, within a pre-defined period or before the end of lesson. This can be set and agreed with the students at the start of group project. Finally, the group size and the number of groups should be limited, to avoid too much work to the students. The maximum group size is six, whereas the maximum number of groups is six as well.

X. Limitations

As with most researches, this study was subject to limitations. Although Likert scale questionnaires are often used to measure students' perceptions, the possibility of individual's bias or peer influence in the responses cannot be discounted. Some caution should be exercised when interpreting the findings due to the small and unequal sample sizes. Nevertheless, this study has provided some useful insights into the concerns and preferences in the context of peer assessment. These should be of practical use to teachers and researchers.

XI. Conclusion

Peer assessment in group presentations can be a powerful learning technique that teachers should consider. Students not only learn the generic skills of assessment, but also more importantly, learn the academic contents and presentation skills by listening actively to their peers' presentations. This study explores the perceived benefits and concerns, as well as the desired features of peer assessment for group presentations from the Chinese students' perspective. Gender differences were observed as statistically significant in the peer assessment concerns. Although students generally agreed with the benefits of peer assessment, they showed reluctance to do peer assessment. The findings suggest that bias from peers is strongly perceived by all students, which may lead to their resistance to peer assessment. To encourage students' involvement in the peer assessment process, the assessment tasks must be user-friendly to both the students and the teacher. The present analysis has provided some useful pointers relating to what constitutes an optimal context for effective peer assessment. Providing training to students, allowing students to set the assessment criteria, limiting the group size and number of groups and fostering a sense of mutual trust and respect among the students can help to make peer assessment a positive learning experience.

References

[1] M. Baranowski and K. Weir, "Peer evaluation in the political science classroom," *PS: Political Science & Politics*, vol. 44, no. 4, pp. 805-811, 2011.

[2] D. Boud, R. Cohen, and J. Sampson, "Peer learning and assessment," *Assessment and Evaluation in Higher Education*, vol. 24, no. 4, pp. 413-426, 1999.

[3] S. Brown, C. Rust, and G. Gibbs, "Involving students in the assessment process," *Strategies for Diversifying Assessment in Higher Education*, Oxford: Oxford Centre for Staff Development, 1994.

[4] N. Falchikov, *Improving Assessment through Student Involvement: Practical Solutions for Aiding Learning in Higher and Further Education*, 1st Edition. New York: Routledge, 2005.

[5] S. Griffiths, "Teaching and learning in small groups," *Teaching and Learning in Higher Education: Enhancing Academic Practice*, New York: Routledge, 2009, pp. 72-84.

[6] P. Orsmond, S. Merry, and A. Callaghan, "Implementation of a formative assessment model incorporating peer and self-assessment," *Innovations in Education and Teaching International*, vol. 41, no. 3, pp. 273-290, 2004.

[7] A. Zariski, "Student peer assessment in tertiary education: Promise, perils and practice," in *Proc. 5th Annual Teaching Learning Forum*, Perth, Australia, 1996, pp. 189-200.

[8] P. Race, "Practical pointers in peer assessment," *Peer Assessment in Practice (SEDA paper 102)*, Birmingham: SEDA, 1998, pp. 113-122.

[9] L. A. J. Stefani, "Peer, self and tutor assessment: relative reliabilities," *Studies in Higher Education*, vol. 19, no. 1, pp. 69-75, 1994.

[10] A. Davies, "Self, peer and tutor-assessment," presented at the National Conference on The Impact of Modularity on Art and Design in Higher Education, Dartington Hill, UK, September 14-6, 1994.

[11] J. Mok, "A case study of students' perceptions of peer assessment in Hong Kong," *ELT Journal*, vol. 65, no. 3, pp. 230-239, 2011.

[12] C. Brindley and S. Scoffield, "Peer assessment in undergraduate programmes," *Teaching in Higher Education*, vol. 3, no. 1, pp. 79-90, 1998.

[13] D. A. F. Haaga, "Peer review of term papers in graduate psychology courses," *Teaching of Psychology*, vol. 20, no. 1, pp. 28-32, 1993.

[14] G. A. Marcoulides and M. G. Simkin, "The consistency of peer review in student writing projects," *Journal of Education for Business*, vol. 70, no. 4, pp. 220-223, 1995.

[15] I. E. Hughes and B. J. Large, "Staff and peer-group assessment of oral communication skills," *Studies in Higher Education*, vol. 18, no. 3, pp. 379-385, 1993.

[16] R. Kuisma, "Assessing individual contribution to a group project," *Assessment of University Students in Hong Kong: How and Why, Assessment Portfolio, Students' Grading - Evaluation of the Student Experience Project*, vol. 2, City University of Hong Kong, Centre for the Enhancement of Learning and Teaching, 1998, pp. 79-106.

[17] A. M. Langan *et al.*, "Peer assessment of oral presentations: effects of student gender, university affiliation and participation in the development of assessment criteria," *Assessment & Evaluation in Higher Education*, vol. 30, no. 1, pp. 21-34, 2005.

[18] C. Orpen, "Student versus lecturer assessment of learning: A research note," *Higher Education*, vol. 11, no. 5, pp. 567-572, 1982.

[19] S. W. Huck, *Reading Statistics and Research*, 6th ed. Boston: Pearson, 2012, ch. 11.

[20] K. M. Loewenthal, *An Introduction to Psychological Tests and Scales*, 2nd ed. Philadelphia, Pa.: Psychology Press, 2001, ch. 5.

[21] N. Falchikov and J. Goldfinch, "Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks," *Review of Educational Research*, vol. 70, no. 3, pp. 287-322, 2000.

[22] P. Davis, "Computerized peer assessment," *Innovations in Education and Training International*, vol. 37, no. 4, pp. 346-355, 2000.

[23] H. Smith, A. Cooper, and L. Lancaster, "Improving the quality of undergraduate peer assessment: A case for student and staff development," *Innovations in Education and Teaching International*, vol. 39, no. 1, pp. 71-81, 2002.

[24] N. Falchikov and D. Magin, "Detecting gender bias in peer marking of students' group process work," *Assessment & Evaluation in Higher Education*, vol. 22, no. 4, pp. 385-396, 1997.

[25] C. Fitzpatrick, "Students as evaluators in practicum: examining peer/self assessment and self-efficacy," presented at the National Conference of the Association for Counselor Education and Supervision, New Orleans, U.S.A., October 27-31, 1999.
[26] M. Besterfield-Sacre, M. Moreno, L. J. Shuman, and C. J. Atman, "Gender and ethnicity differences in freshmen engineering student attitudes: A cross-institutional study," *Journal of Engineering Education*, vol. 90, no. 4, pp. 477-489, 2001.
[27] S. G. Brainard and L. Carlin, "A six-year longitudinal study of undergraduate women in engineering and science," *Journal of Engineering Education*, vol. 87, no. 4, pp. 369-375, 1998.
[28] S. H. Appelbaum and A. Hare, "Self-efficacy as a mediator of goal setting and performance some human resource applications," *Journal of Managerial Psychology*, vol. 11, no. 3, pp. 33-47, 1996.
[29] P. Hampel, M. Meier, and U. Kümmel, "School-based stress management training for adolescents: Longitudinal results from an experimental study," *Journal of Youth Adolescence*, vol. 37, no. 8, pp. 1009-1024, 2008.
[30] G. Kok, H. Vries, A. N. Mudde, and V. J. Strecher, "Planned health education and the role of self-efficacy: Dutch research," *Health Education Research Theory & Practice*, vol. 6, no. 2, pp. 231-238, 1991.

Caroline T. W. Chan received her PhD degree from Loughborough University, U.K. She is currently a lecturer at the City University of Hong Kong. Her main research interests include project cost estimation, construction project management and pedagogical improvement.