Online peer assessment: effects of cognitive and affective feedback

Jingyan Lu · Nancy Law

Abstract   This study reports the effects of online peer assessment, in the form of peer grading and peer feedback, on students' learning. One hundred and eighty one high school students engaged in peer assessment via an online system—iLap. The number of grade-giving and grade-receiving experiences was examined and the peer feedback was coded according to different cognitive and affective dimensions. The effects, on both assessors and assessees, were analyzed using multiple regression. The results indicate that the provision by student assessors of feedback that identified problems and gave suggestions was a significant predictor of the performance of the assessors themselves, and that positive affective feedback was related to the performance of assessees. However, peer grading behaviors were not a significant predictor of project performance. This study explains the benefits of online peer assessment in general and highlights the importance of specific types of feedback. Moreover, it expands our understanding of how peer assessment affects the different parties involved.

Keywords   Peer assessment · Online assessment · Peer feedback · Peer grading · Cognitive feedback · Affective feedback

Introduction

This paper reports on a study involving the design of online peer assessment activities to support high school students' project-based learning; the study examined the effects of different types of peer assessment on student learning. Assessment has an important influence on the strategies, motivation, and learning outcomes of students (Crooks 1988). Traditionally, too much emphasis has been placed on assessment by grading, and too little on assessment as a way of helping students learn (Crooks 1988; Stiggins 2002). In recent years, peer assessment has been adopted as a strategy for "formative assessment" (Cheng and Warren 1999; Sadler 1989) or "assessment for learning" and for involving students as

J. Lu (✉) · N. Law
Faculty of Education, The University of Hong Kong, Pokfulam Road, Hong Kong, China
e-mail: jingyan@hku.hk

active learners (Gielen et al. 2009; Sadler 1989; Topping et al. 2000). Further, research on peer assessment has amassed substantial evidence on the cognitive (Nelson and Schunn 2009; Tseng and Tsai 2007), pedagogical (Falchikov and Blythman 2001), meta-cognitive (Butler and Winne 1995; Topping 1998), and affective benefits (Strijbos et al. 2010) of peer assessment on student learning (Topping 2003). These efforts have resulted in peer assessment being successfully designed and implemented in K-12 classrooms and in higher-education contexts (Topping 2003).

Online assessment has become increasingly popular since the advent of the Internet and has significantly changed the process of assessment (Tseng and Tsai 2007). Online systems introduce such functions as assignment submission, storage, communication and review management (Kwok and Ma 1999; Liu et al. 2001), and online assessment has a number of advantages over face-to-face assessment (Tsai 2009; Tsai and Liang 2009; Yang and Tsai 2010). Online assessment enables students to communicate with peers and to reflect on and continuously revise their work based on feedback (Yang 2010). Online systems can increase the willingness of students to engage in peer assessment by allowing them to anonymously grade and provide feedback when and where they like (Lin et al. 2001; Tsai 2009). These systems also allow teachers to monitor the online activities and progress of their students more closely (Lin et al. 2001), and enable researchers to collect information about students by automatically recording data about assignments, online participation and communication (Tsai 2009). Finally, teachers can automatically assign students to review more heterogeneous or homogeneous work based on background features such as gender, achievement, and preferences (Tsai 2009).

The highly variable nature of peer assessment practices makes it difficult to determine their effects on learning (van Gennip et al. 2010). Further, as research has tended to focus on university students, we have an under-representation in the literature of these practices in schools. Thus an investigation of the peer assessment practices of high school students will help rectify this unfortunate situation and enable their practices to be compared with those of university students.

This study focuses on how peer grading and peer feedback affect the performance of both assessors and assessees. The paper starts with a discussion of two important assessment components: peer grading and peer feedback. Next, the features of the online assessment system and peer assessment procedures are reported, and the research questions are introduced. This is followed by a description and discussion of the methods of data collection and data analysis adopted. Finally, there is a discussion of the major findings, implications of this study and directions for future work.

Literature review

Many teachers confuse assessment with grading and consider them to be the same thing. Although grading can be seen as a form of assessment, assessment does not necessarily involve grading. This study treats peer grading and peer feedback as two forms of peer assessment and examines their effects on the learning of assessors and assessees.

Peer grading

In peer grading, assessors apply criteria for assigning grades to the work of their peers; many studies have shown this to be a reliable and valid approach (Falchikov and Goldfinch 2000). For instance, a meta-analysis shows a mean correlation of 0.69 between peer- and

teacher- assigned grades, indicating that peer assessment can be reliable (Falchikov and Goldfinch 2000). This finding can perhaps be explained by the fact that teachers often support peer grading by providing students with assessment rubrics to ensure consistent and reliable peer evaluations (Jonsson and Svingby 2007). Assessment rubrics based on descriptive scales show students what is important in assignments and specify the strengths and weaknesses of the work under evaluation (Andrade 2000; Moskal 2000). By applying rubrics to the work of peers, assessors enhance their awareness and understanding of the assessment criteria and, as a result, are likely to apply these to their own work more reflectively and attentively.

Studies show that rubric-supported peer grading enhances student learning. Students become more reflective and their learning outcomes improve when they are involved in defining marking rubrics (Stefani 1994). Reports from undergraduates have indicated that, although peer grading is challenging and time-consuming, it is also beneficial as it enables students to think more critically and to learn more effectively (Falchikov 1986; Hughes 1995; Orsmond et al. 1996). While peer grading is widely used and its effects on university students have been extensively examined, its effects on high school students have seldom been investigated.

Peer feedback

Despite the positive reports on the impact of peer grading on students' learning, many researchers have argued that evaluative feedback is more important than assigning grades (Ellman 1975; Liu and Carless 2006) and that peer feedback is more effective than grading in peer assessment.

Peer feedback refers to giving comments on the work or performance of peers, which involves reflective engagement (Falchikov and Blythman 2001). Some researchers have argued that peer feedback enhances learning by enabling learners to identify their strengths and weaknesses, and by receiving concrete ideas on how to improve their work (Crooks 1988; Rowntree 1987; Xiao and Lucking 2008). Although, peer feedback may not achieve the quality of teacher feedback, it can often be given in a more timely manner, more frequently, and more voluminously (Topping 1998). Assessees may also see feedback from peers as less threatening and so perhaps be more willing to accept it (Ellman 1975).

Cognitive and affective feedback

Nelson and Schunn (2009) differentiate feedback into cognitive and affective categories. Cognitive feedback targets the content of the work and involves summarizing, specifying and explaining aspects of the work under review. Affective feedback targets the quality of works and uses affective language to bestow praise ("well written") and criticism ("badly written"), or uses non-verbal expressions, such as facial expressions, gestures and emotional tones. In online learning environments, emoticons are often used for giving affective feedback.

The effects of cognitive feedback vary with the types given or received, the nature of the task, the developmental stage of the learners, and mediating conditions. When assessors give cognitive feedback they summarize arguments, identify problems, offer solutions, and explicate comments. Different forms of cognitive feedback can affect the performance of assessees in different ways. Earlier feedback is more effective than later feedback, and feedback to correct responses is more effective that feedback to incorrect ones (Hattie and Timperley 2007). Specific comments are more effective than general ones (Ferris 1997),

as are timely comments emphasizing personal progress in mastering tasks (Crooks 1988). Short, to-the-point explanations are more effective (Bitchener et al. 2005) than lengthy, didactic ones (Tseng and Tsai 2007). Finally, comments that identify problems and provide solutions are especially effective (Nelson and Schunn 2009). Whereas peer feedback studies have overwhelmingly focused on the effects of written feedback on the writing assignments of university students (either as assessors or as assessees), this study focuses on the effects of online feedback on high school students, both assessors and assessees, within Liberal Study projects.

Praise is usually recommended and is found to be one of the most common features present in the feedback provided by undergraduates (Cho et al. 2006; Nilson 2003; Sadler and Good 2006). There are many studies of praise given by teachers (Brophy 1981; Burnett 2002; Elawar and Corno 1985; Elwell and Tiberio 1994; Page 1958; White and Jones 2000; Wilkinson 1981), but few of them focus on praise that is given by peers. More importantly, the effects of praise on learning are unclear and highly debated. Some studies report positive effects of praise: high school students report that online positive feedback significantly contributes to their learning (Tseng and Tsai 2007); first year college students respond favorably to praise in their writing performance where it helps them improve work quality (Straub 1997); and college students generally perceive positive comments as motivating (Duijnhouwer 2010). Conversely, some literature indicates that praise is ineffective for task performance, either for students at high school level (Crooks 1988), or at college level (Ferris 1997), especially where the required performance is a cognitively demanding task (Ferris 1997; Hattie and Timperley 2007; Kluger and DeNisi 1996). Some researchers argue that the constructive effects of affective feedback, whether positive or negative, are limited because they do not address the cognitive content of the work under review. Hattie and Timperley (2007), for instance, find that affective feedback fails to bring about greater task engagement and understanding because it usually provides little task-related information. Surprisingly, negative effects of praise are found in Cho and Cho's (2010) study, which reports university students' physics technical research drafts tend to be of a lower quality when students receive more praise. Furthermore, most research focuses on the effects of praise on assessees; little has been done on the effects on assessors, except for the work by Cho and Cho (2010), who report that, the more positive comments students give, the more the writing quality of their own revised drafts tends to improve. However, the coding on praise comment is combined with other cognitive feedback, so the unique effect of praise is hard to differentiate. Most research on praise is conducted in traditional classroom settings, where students receive praise from the teacher; reports on the effects of feedback from peers given through online systems are limited. In this study we investigate the effects of praise and cognitive feedback on both assessors and assessees, as students usually play both roles in the same task in a given situation.

Effects of peer feedback on assessors and assessees

Peer feedback can improve the learning of both assessors and assessees (Li et al. 2010; Topping and Ehly 2001; Xiao and Lucking 2008) by sharpening the critical thinking skills of the assessors and by providing timely feedback to the assessees (Ellman 1975). It increases the time spent thinking about, comparing, contrasting and communicating about learning tasks (Topping 1998). Furthermore, assessors review, summarize, clarify, diagnose misconceived knowledge, identify missing knowledge, and consider deviations from

the ideal (Van Lehn et al. 1995). Assessors who provide high quality feedback have better learning outcomes (Li et al. 2010; Liu et al. 2001). When pre-service teachers assess each other's work in Tsai et al.'s study (2001), those who provide more detailed and constructive comments perform better than those who provide less. Topping et al. (2000) also report positive effects of peer assessment on assessors: students involved in peer assessment not only improved the quality of their own work but also developed additional transferable skills.

Some studies provide mixed findings with regard to the usefulness of peer feedback for assessees. Olson (1990) reports that sixth graders who received both peer and teacher feedback produced better quality writing in their final drafts than those who received only teacher feedback. Some students show mistrust of peer assessment. Brindley and Scoffield (1998) report that some undergraduate students expressed concerns about the objectivity of peer assessment due to the possibility of personal bias when comments are given by peers, and some students considered assessment to be the sole responsibility of teachers (Brindley and Scoffield 1998; Liu and Carless 2006; Wen and Tsai 2006).

Although research indicates that peer feedback is more beneficial to assessors than to assessees, most of the research has focused on college students. Moreover, studies comparing the effects of peer feedback on assessors and assessees and on primary and secondary school students are rare.

Research questions

Despite the number of studies on peer assessment, it is difficult to pinpoint the following: What contributes to the effects of peer assessment (van Zundert et al. 2009)? Who benefits most from it, assessors or assessees? And what role does online assessment play? A limited number of quantitative studies compare and measure the effects of different types of assessment on student learning; very few have explicitly compared the effects of peer assessment on both assessors and assessees (Cho and Cho 2010; Li et al. 2010). Researchers have also identified a number of concerns about online assessment, especially with regard to high school students (Tsai 2009). For instance, assessors may have insufficient prior domain knowledge to judge the work of peers or an inability to provide neutral comments, therefore assessees may have difficulties accepting and adapting to feedback from peers. Another possible concern is a lack of computer literacy: both parties may lack the requisite technological skills to fully negotiate the online systems. This study was designed to explore these issues. Students participating in this study were provided with rubrics to help them to grade the work of their peers. The students' academic achievements, reflecting their prior domain knowledge and computer skills, were collected to investigate the extent to which these factors affectend their peer assessment behavior.

This study employed an exploratory method to investigate who will benefit from peer assessment, how, and why. As an empirical investigation to determine whether or not online assessment activities are related to the learning performance of high school students it posed two questions:

1. Are peer grading activities related to the quality of the final project for both assessors and assessees?
2. Are different types of peer cognitive and affective feedback related to the quality of the final projects for both assessors and assessees?

Methods

Subjects

All 181 thirteen- to fourteen-year-old students studying at secondary two level in a publicly funded school in Hong Kong were invited to participate in this study. The school was chosen as a convenient sample from schools that participated in a university-school partnership project involving the use of an online platform in the teaching and assessment of the Liberal Studies (LS). The students studied in five different classes at this school and were taught by different teachers who had worked together to prepare the school-based teaching plans for the LS syllabus. The students worked on a six-week LS project, and, given that all five teachers followed the same teaching schedule, online activities and assessment rubrics, the schedules of the learning tasks and progress of all five classes were equivalent. Students were encouraged to engage in peer assessment for all subtasks so that they could become involved in a continuous process of reflection and improvement. Nevertheless, assessment activities were not compulsory and students had the freedom to choose which subtask(s) to assess.

Online assessment platform and task description

This study is part of a larger project (Law et al. 2009) in which an online learning platform, known as the Interactive Learning and Assessment Platform (iLAP), was developed to support teaching and assessment of LS. LS is a new core course implemented in secondary schools in Hong Kong, where students are provided 'with opportunities to explore issues relevant to the human condition in a wide range of contexts and enables them to understand the contemporary world and its pluralistic nature' (Education and Manpower Bureau 2007). The teaching of the new subject poses many challenges to high school teachers, not only because it is new but also because it advocates an inquiry-based learning approach that is novel to both teachers and students. The online learning platform iLAP was developed to facilitate inquiry-based learning and to help teachers manage, support, and assess students' learning processes and inquiry outcomes.

The functions of the iLap platform are fourfold: (1) assignment submission, (2) rubric development, (3) assessment implementation, and (4) performance monitoring. Students can upload assignments to the platform. iLAP provides some sample assessment rubrics for different kinds of learning/inquiry tasks that teachers can adopt or modify for adoption in their own teaching practice. iLAP also allows teachers to create brand new rubrics on their own by specifying task dimensions and rating criteria for varying levels of performance. Figure 1 is a screen-capture of the iLAP interface that illustrates rubrics creation by teachers: each row is a rubric dimension (or criterion) and the performance levels are represented by the columns. The online rubric on iLAP not only supports the use of teacher assessment but also online self assessment and peer assessment.

For peer assessment, students were divided into small groups of four or five, according to their student ID number (so that group assignment was random). Individual assignments submitted online were made available for review and comment by fellow group members. The peer assessment interface of iLAP comprises two parts: a rubric area and a comment area. When an assignment is selected, the rubric created by the teacher appears to the assessors. For instance, the rubric in Fig. 2 has five criteria: (1) topic researchability, (2) clarity, (3) relevance and background information, (4) methodology, and (5) writing quality. Each criterion is further divided into four levels of quality: 'cannot be scored'
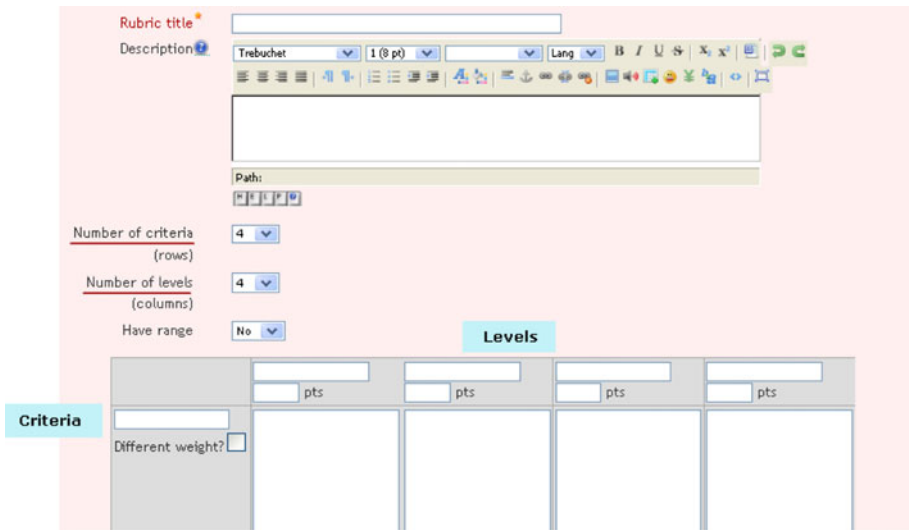
Fig. 1 Screenshot of rubric development on iLAP

(score 0), 'needs improvement' (score 1–3), 'satisfactory' (score 4–6), and 'excellent' (score 7–9). To assign a score to their peers, students click on drop-down menus to choose the level of quality for each category and select the cell colors according to the assigned score. iLAP then automatically calculates the total score for all the assessed dimensions. In addition to peer grading assigned from rubrics, students can also comment on the assessed assignment in the comment area.

iLAP, like many other online peer assessment systems, also allows teachers to easily monitor student performance and progress (Lin et al. 2001). Teachers can access summaries of assignment submissions and assessments of student performance (see Fig. 3). Student work is listed with peer assessment information, including grades and comments.

While LS is a new core course for senior secondary students in Hong Kong, some schools have chosen to introduce it to students as early as junior secondary level, so as to provide a smooth transition for both teachers and students. The curriculum context of this study was the subject "Liberal Study and Humanities", which was a new subject for the secondary two students taking part in this study. Participants worked on a project based on one of the three topics covered in their Humanities course in the previous semester, which was about Consumer Education, City Transportation and Development, and City Economy. Their teachers divided the project into five subtasks: (1) collect three topic-relevant photographs, (2) draft project plans, (3) prepare interview questions, (4) draft questionnaires, and (5) write up projects. Each subtask has a different focus. Subtask 1 focused on finding three photographs related to the topic and formulating a research question. Subtask 2 focused on drafting a plan for an empirical investigation of the research question. Subtask 3 focused on preparing interviews, including identifying possible interviewees and constructing interview questions. Subtask 4 focused on drafting a questionnaire for the investigation. Subtask 5 focused on writing up the project reports. Students could collect empirical data through interviews or surveys with their identified informants, who could be peers, professionals, or members of the general public, as deemed appropriate by the students based on the research question they wanted to answer.
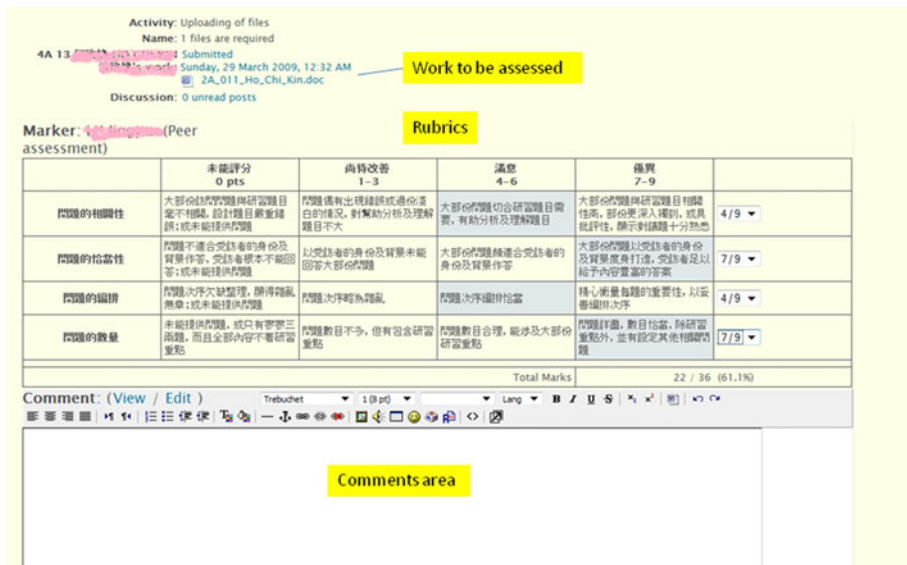
Fig. 2 Screenshot of peer assessment on iLAP

The teachers assigned a specific rubric for students to use in the peer assessment of each subtask. Students were asked to grade the work of their peers in their own group based on teacher-provided rubrics and were also asked to give comments of their own volition. As soon as grades and comments were provided by peers, they became immediately available to assessees through iLAP. None of the subtasks used self-assessment. The teacher's final grading of the project, rather than peer grading scores, formed the basis of each student's final score.

Data sources, coding schema, and data analysis

iLAP records how often students give and receive grades and feedback, to whom these are sent and from whom these are received, and the content of the feedback exchanged. Achievement scores for the Computer Literacy and Humanities course in the previous semester were collected as control variables in the analysis.

Peer feedback was qualitatively analyzed (See Table 1 for coding scheme details and examples). The authors collaboratively developed a coding scheme based on relevant literature and the peer assessment context of this study. Based on Nelson and Schunn (2009), and Tseng and Tsai (2007), comments were first coded as affective and/or cognitive. Affective comments were further coded as positive (e.g. "very good") or negative (e.g. "badly written"). Cognitive comments were categorized as (1) identify problem; (2) suggestion; (3) explanation; and (4) comment on language. The first author and a research assistant coded the work independently with an inter-rater reliability of .83. Comments that were neither cognitive nor affective (e.g. "can't read your project, cannot comment") were classified as 'other' and were later excluded as there were very few of these.
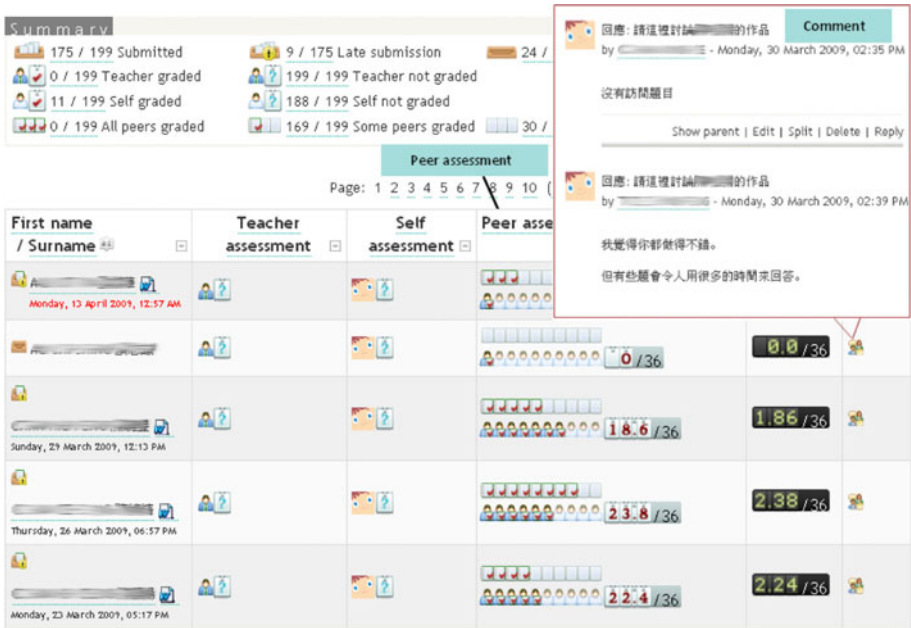
Fig. 3 Screenshot of the teacher's view of student assessment

Table 1 Coding scheme for peer feedback

Categories	Definition	Example
Cognitive		
Identifying problems	Addressing specific issue or dimension of the essay	You don't have enough survey questions!
Suggestions	A method is suggested to deal with the problem	"The topic is too broad, should narrow down the topic"
Explanation	Explanation or elaboration on the problems identified or suggestions provided	Your project lacks of reliable evidence. If you don't collect people's opinion, how can you know it (subway) is convenient?
Language	Comments addressing the writing in general	Your writing is to colloquial.
Affective		
Negative	Give criticism	"You didn't even do it!" "Badly written"
Positive	Praise the work	"Very good" "Quite good overall"

Student grading and assessment activities were coded, quantified, and summed for all five subtasks. Data from the five classes were collapsed for analysis because: (1) students from the five classes did not differ significantly in their academic performances in general based on the achievement data collected from the previous semester; (2) the materials and tasks on iLAP were identical for all five classes; (3) the five classes did not differ

significantly in their final project scores based on the ANOVA test results ($F = 1.46$, $P > 0.05$); and (4) multilevel intra-class correlation on class dependence (Raudenbush and Bryk 2002) was 0.01, which indicating no statistically significant variance difference among the five classes. This implied that teachers' differences were not a factor.

Statistical analyses were used to explore relationships between students' online assessment activities and LS performance. Hierarchical multiple regression was used to examine the relationship of these elements with project performance (final project scores) controlled by student grades in previous Computer Literacy and Humanities courses prior to the commencement of the study. By entering blocks of independent variables in sequence, we identified additional variance explained by newly introduced variables in each step. In the first step, two control variables, examination scores in the Humanities and Computer Literacy courses, were entered into the regression equations. Peer grading measures to assessees and by assessors were entered in the second step. Peer feedback measures to assessees and by assessors were entered in the third step. The independent predictor variables included in this study are the number of grades given to and received from peers, number of suggestions given to and received from peers, number of explanations given to and received from peers, number of comments on identifying the problems given to and received from peers, number of comments on language issues given to and received from peers, number of items of positive and negative emotional feedback given to and received from peers.

Results

Summary statistics

Data from 181 students was included in the analysis. Data from nine other students was excluded because their final project scores were missing and they were not involved in any assessment activities. Table 2 summarizes students' online peer assessment activities. Table 3 provides the correlations of all variables. According to Table 2, students on average gave and received similar amounts of the same types of feedback. Note that the most frequently given and received types of cognitive feedback were Identifying problem [Mean (given) = 5.84, Mean (received) = 5.83], followed by Suggestion [Mean (given) = 1.10, Mean (received) = 1.12], then Comment on language use [Mean (given) = .71, Mean (received) = .70], and Explanation [Mean (given) = .52, Mean (received) = .55]. The number of peer grading received by participants (mean = 13.19) was similar to the number they gave to their peers (mean = 13.94). The variance of assessment activities also varies across different types of feedback. The higher the range and means of peer assessment variables the larger the variance. However, within the same type of feedback, variances between giving and receiving feedback are about the same.

Multiple regression

The multiple squared correlation coefficient was .128 indicating that students' online collaboration and evaluation activities accounted for about 12.8% of project performance variance. Table 4 shows that the two control variables were significant predictors of final project scores ($\Delta R^2 = 0.037$, $p < 0.05$). Students' performance in the Humanities course (PHS) was a significant predictor of final project scores ($t = 2.58$, $p < 0.05$) while their performance in the Computer Literacy course (PCLS) was not ($t = -.27$, $p > .05$). Adding

Table 2 Descriptive statistics of major variables ($N = 181$)

		Min	Max	Mean	S.D.
1	SugTP	0	9	1.10	1.76
2	SugBP	0	8	1.12	1.56
3	ExpTP	0	4	.52	.84
4	ExpBP	0	5	.55	.87
5	IdeTP	0	17	5.84	4.06
6	IdeBP	0	14	5.83	3.13
7	LanTP	0	13	.71	1.39
8	LanBP	0	6	.70	1.16
9	NETP	0	9	.52	1.09
10	NEBP	0	10	.54	1.15
11	PETP	0	23	3.26	3.33
12	PEBP	0	13	3.33	2.71
13	TGTP	0	27	13.98	5.74
14	TGBP	0	26	13.34	4.98
15	PCLS	18	90	68.64	10.92
16	PHS	20.45	81.36	50.83	11.14
17	PS	0.7	13.08	5.81	2.65

SugTP suggestions to peers, *SugBP* suggestions from peers, *ExpTP* explanations feedback to peers, *ExpBP* explanations from peers, *IdeTP* feedback on identifying problems to peers, *IdeBP* feedback on identifying problems by peers, *LanTP* giving feedback on language (sues), *LanBP* receiving feedback on language (sues), *NETP* Negative emotional feedback to peers, *NEBP* Negative emotional feedback from peers, *PETP* Positive emotional feedback to peers, *PEBP* Positive emotional feedback from peers, *TGTP* Total grades given to peers, *TGBP* Total grades from peers, *PCLS* Previous Computer Literacy score, *PHS* Previous Humanities score, *PS* project score as the dependent variable

the number of peer grading received or given to the model produced no significant change in variance, while adding peer feedback activities did (adjusted $\Delta R^2 = .142$, $p < .01$). Controlling for students' prior Humanities scores, Computer Literacy scores, and number of grading, we identified three variables with significant contributions to the model: Giving suggestions to peers ($t = 2.17$, $p < .05$), giving feedback on identifying problems to peers ($t = 2.16$, $p < .05$), and positive affective feedback from peers ($t = 2.25$, $p < .05$). However, the strength of prediction of Humanities performance became insignificant when peer feedback activities were added ($t = .75$, $p > .05$). Other online assessment activities were not significant predictors of the final project scores.

Discussion

This study examined whether the involvement of high school students in online peer assessment predicted their performance on LS projects. We focused on peer grading and peer cognitive and affective feedback, and their effects on both assessors and assessees. Online peer assessment was found to significantly affect the quality of students' project learning outcomes, probably because it provided opportunities to evaluate the work of others. We discuss the effects of grading, followed by cognitive and affective feedback.

Table 3 Zero-order correlations among major variables ($N = 181$)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. IdeTP																
2. IdeBP	.30**															
3. SugTP	.54**	.19*														
4. SugBP	.27**	.33**	.31**													
5. ExpTP	.31**	.10	.34**	.16*												
6. ExpBP	.12	.10	.23**	.25**	.11											
7. LanTP	.31	.08	.24**	−.01	.12	−.10										
8. LanBP	.05	.10	.04	.09	.01	.10	.15*									
9. NETP	0	−.10	−.05	−.13	−.09	−.10	−.10	−.10								
10. NEBP	−.16*	−.05	−.19**	−.18**	−.09	−.10	−.10	−.10	.22**							
11. PETP	.30**	0	.24**	−.03	.12	.01	.06	−.10	.21**	.04						
12. PEBP	.10	.11	.06	.16*	.01	.12	−.10	−.10	.16*	−.08	.36**					
13. TGTP	.38**	.22**	.32**	.22**	.09	.10	.04	.03	.18*	.04	.25**	.14				
14. TGBP	.14	.36**	.17*	.34**	.01	.09	−.10	.01	.02	.12	.20**	.29**	.53**			
15. PCLS	.22**	.21**	.24**	.19*	.17*	.07	−.10	−.10	.09	.03	.14	.18*	.27**	.29**		
16. PHS	.42**	.27**	.42**	.35**	.38**	.18*	.14	.02	−.01	−.20**	.12	.18*	.32**	.22**	.32**	
17. PS	.32**	.10	.31**	.07	.10	.11	.17*	.11	−.09	−.20**	.14	.18*	.11	.10	.05	.22**

SugTP suggestions to peers, *SugBP* suggestions from peers, *IdeTP* feedback on identifying problems to peers, *IdeBP* feedback on identifying problems by peers, *ExpTP* explanations feedback to peers, *ExpBP* explanations from peers, *LanTP* giving feedback on language (sues), *LanBP* receiving feedback on language (sues), *NETP* Negative emotional feedback to peers, *NEBP* Negative emotional feedback from peers, *PETP* Positive emotional feedback to peers, *PEBP* Positive emotional feedback from peers, *TGTP* Total grades given to peers, *TGBP* Total grades from peers, *PCLS* Previous Computer Literacy score, *PHS* Previous Humanities score, *PS* project score as the dependent variable

* $p < .05$

** $p < .01$

Table 4 Hierarchical multiple regression analysis in predicting project score ($N = 187$)

	R^2	R^2_{adj}	ΔR^2	ΔF	β	t
Step 1	.047*	.037*	.037*	4.412		
PCLS					−.02	−.27
PHS					.22**	2.89
Step 2	.050	.033	.002	.428		
PCLS					−.03	−.39
PHS					.21**	2.63
TGTP					.05	0.65
Step 3	.191*	.128*	.142*	2.93		
PCLS					−.05	−.60
PHS					.07	.75
TGTP					−.04	−.49
SugTP					.20*	2.17
IdeTP					.20*	2.16
PEBP					.16*	2.25

PCLS Previous Computer Literacy score, *PHS* Previous Humanities score, *TGTP* Total grades given to peers, *SugTP* suggestions to peers, *IdeTP* feedback on identifying problems to peers, *PEBP* Positive emotional feedback from peers

* $p < .05$

** $p < .01$

Effects of peer grading

Although a modest significant correlation was found, the number of peer assigned or received grades did not predict LS project scores. It was hypothesized that more grading would improve assessors' understanding of the rubric and consequently their own performance. Though students gave and received more grades than feedback, the effect of the former was less significant than that of the latter. This finding echoes studies that suggest peer grading alone is less effective than peer grading plus feedback. Xiao and Lucking (2008) found that students involved in peer grading plus qualitative feedback were more satisfied and showed greater improvements in their writing task performance than those involved only in peer grading. Ellman (1975) even suggests that peer grading is an unfortunate reinforcement of the already overemphasized practice of grading. Thus, providing meaningful, constructive, and concrete comments to fellow students appears to be more conductive to learning than solely assigning grades.

Effects of cognitive and affective comments

Different types of cognitive comments

Although research findings on the effects of different types of feedback are readily available in the literature, the different effects these have on assessors and assessees are not (Li et al. 2010). Moreover, while peer feedback is a common practice in schools

and is thought to be helpful, there is little agreement as to which types of feedback are most effective. In this study, we find that students benefit more as assessors than assessees, particularly with regard to comments that identify problems and make suggestions.

The more problems assessors identified, and the more suggestions they made, the better they performed in their own LS projects. It seems reasonable that identifying problems and making suggestions leads assessors to engage in activities with higher cognitive demands. Identifying problems, the type of cognitive feedback most frequently given, varied from such specific comments as "the photos are not clear" to general comments such as "the inquiry problem should be clearer". Suggestions were the next most frequently given type of cognitive feedback. Assessors develop clearer and deeper understandings of assessment criteria and project tasks by judging and commenting on the quality of peer projects (Liu et al. 2001; Reuse-Durham 2005). They gained insights into their own projects by assessing the projects of peers (Bostock 2000). Our findings replicated Chen and Tsai' research (2009) on in-service teachers in that the more peer feedback given to peers, the more likely students were to make improvements to their research proposal. Giving explanation and feedback on language use were not found to be significant predictors of project performance. This may be due to the low frequencies of these two kinds of feedback. Future studies should aim to encourage students to employ these forms of feedback and to study their effects on learning performance.

We found that cognitive feedback did not have the same effects on assessees, which contradicts Gielen et al.'s (2009) finding of a positive relationship between receiving constructive comments, especially justified ones, and learning. Many researchers have observed that constructing feedback that is useful to assessees is a very complex issue. The difference between the correlations of research project scores with feedback given and feedback received, respectively, indicate that those who tend to give more feedback gain more from the project process at the same time, while weak students do not benefit from receiving more feedback. In other words, feedback affects assessors rather than assessees, and peer feedback may not help weak students. This could be because weak students have difficulties understanding, interpreting, and consequently integrating the feedback received. Secondly, not all feedback is equally useful. A review of the relevant literature reveals that different types of feedback may have different effects. Our findings also showed that the feedback students gave and received was not equally distributed. The third issue is feedback implementation—the intermediate step that occurs between feedback and performance improvement (Nelson and Schunn 2009). Simply put, feedback is useful to recipients only when they act on it (Topping 1998). Thus, feedback affects assessees *indirectly* through the mediation of understanding (Nelson and Schunn 2009). When learners receive feedback, they first need to fully comprehend the problem(s) pinpointed or suggestion(s) offered. Unfortunately, explanations were rarely provided by students in this study (the means were .52 for assessors and .55 for assessees). This could make it more difficult for the weaker students to benefit from peer feedback as they did not know how to improve their work when receiving qualitative feedback without sufficient explanation. Another important issue is perceived validity (Straub 1997). Assessees have to evaluate the usefulness of feedback. They may only act on feedback that they perceive as valid and useful for improving their performance. Or they may value only feedback from teachers rather than peers. The fact that so many factors can influence the extent to which assessees accept and implement feedback helps to explain why cognitive feedback failed to enhance their performance.

Explanations and language comments did not predict final project performance. Students gave very little feedback of this type, and consequently received few of them. In addition to the low frequencies of occurrence, language comments may have failed to predict project performance because they were too general, such as "you should improve your grammar," or perhaps because language played a minor role in final project performance.

Affective comments

Positive affective feedback was a significant predictor for the performance of assessees on LS projects in this study. The more positive comments students received, the more likely they were to perform well on their projects. Evidence in the literature for the usefulness of praise is conflicting and inconclusive. Some researchers have found that praise has a negative affect (Cho and Cho 2010), or no affect, on student performance (Crooks 1988; Ferris 1997), while others have found that it has a positive effect on student performance (Straub 1997; Tseng and Tsai 2007). Henderlong and Lepper (2002) found that praise enhances intrinsic motivation and maintains interest when it encourages performance attributions, promotes a sense of autonomy, and provides positive information about personal competence. Affective comments that provoke positive feelings help boost student interest, motivation, and self-efficacy, even when they are not task-focused or informative. Another possible explanation is that the affective language used in feedback might color a person's perception of the reviewer and the feedback received (Nelson and Schunn 2009). Affective comments provoking positive emotions in learners could lead to a favorable outlook on the comments and consequent implementation. This study found that students receiving positive feedback were more likely to perform better on their final LS projects than those who did not. However, no information was collected asking how assessees interpreted this positive feedback and how such feedback affected their learning performance. Future study should investigate how such feedback might affect their learning outcomes. Praise did not have positive effects on assessors, which contradicts the finding by Cho and Cho (2010). This could be a result of the different method of coding used within this study: we coded praise and cognitive feedback separately, while Cho and Cho coded them together.

Prior knowledge

Two control variables: Prior Humanity Scores (PHS) and Previous Computer Literacy Scores (PCLS), affected students' LS project performance differently. PCLS was not significantly related to LS projects scores, perhaps because the computer skills needed for LS projects were minimal and below the levels of variance. Student project performance had a low correlation with PHS. Although the three topics: "consumer education", "transportation and city development", and "economy" were related to topics in the previous Humanities course, they made only a small contribution to Project scores (3.7% variance change). The fact that PHS became insignificant when cognitive feedback was added to the model implied great overlap of variance on the two variables. Effects of PHS were transferred to or represented by the impact of cognitive feedback activities. This suggests that students who were good at Humanities also made more cognitive feedback. However, cognitive feedback also had its unique and significant contributions to project scores, as exemplified by the significant change of variance of 14.2%.

Conclusion

This study expands our understanding of the relationship between peer assessment and learning performance. It proposes a possible explanation for the benefits of online peer assessment on student learning performance. Students benefit more as assessors than as assessees. Delving into how peer grading and feedback influence assessors and assessees, respectively, three variables stand out. Thoughtful feedback on specific problems and suggestions are strong predictors of how assessors perform on their final project. Positive affective feedback predicts a higher performance of assessees. Peer grading behaviors have very limited effects on the learning performance of both assessors and assessees.

Our findings have several implications for teachers and educational researchers interested in designing and implementing peer assessment. First, teachers need to be sensitive to the fact that peer assessment works differently for assessors and assessees. Different types of peer assessment, either peer grading or peer feedback, set in motion different learning processes on the part of assessors and assesses, which can lead to different outcomes. Second, modeling or training should be provided prior to or during the task because peer assessment is not easy. Teachers should ask students to be specific in their feedback, particularly with regard to the problems in assessees' work, and to provide suggestions. The findings of our study could assist teachers in developing strategies to amplify the effectiveness of peer assessment for all students. Peer grading appears to be less effective than peer feedback for assessors, because giving feedback to peers activates crucial cognitive processes that contribute to learning gains of assessors. Students should be encouraged to give thoughtful and meaningful comments rather than simply assign grades to peers. Students can be asked to explain why they assigned particular grades to peers. Third, teachers should scaffold assessment processes with scaffolding tools, particularly for the weak students. As weak students could neither give nor benefit from qualitative feedback they need to be given specific instructions on the types of feedback they should give to peers; they should also be encouraged to reflect on and implement the feedback they receive. Fourth, students should be encouraged to exchange affective comments that give socio-emotional support to peers and recognize peers' achievement. Our results suggest that positive affective comments are not just about making other people feel good. They can help boost the motivation, interest, and self-efficacy of assessees, which in turn can enhance their performance.

Enhancing the effects of feedback on assessees is a complex issue. Effects on assessees may be hard to see since we might need to see how they change their attitudes with respect to feedback for it to be successful. For example, students should be asked to respond to feedback by explaining why (or why not) and how they would implement suggestions for their work. In future studies, we suggest using mixed methods to investigate how peer assessment affects different kinds of participants and under what conditions or circumstances. Such studies can inform practitioners and researchers about how to cultivate "mindful reception" of feedback (Bangert-Drowns et al. 1991) on the part of assessees.

References

Andrade, H. G. (2000). Using rubrics to promote thinking and learning. *Educational Leadership, 57*(5), 13–19.

Bangert-Drowns, R. L., Kulik, C. L. C., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*(2), 213–238.

Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing, 14*(3), 191–205.

Bostock, S. (2000). Student peer assessment Retrieved May 5, 2010, from http://www.palatine.ac.uk/files/994.pdf.

Brindley, C., & Scoffield, S. (1998). Peer assessment in undergraduate programmes. *Teaching in Higher Education, 3*(1), 79–90.

Brophy, J. (1981). Teacher praise: A functional analysis. *Review of Educational Research, 51*(1), 5–32.

Burnett, P. C. (2002). Teacher praise and feedback and students' perceptions of the classroom environment. *Educational Psychology, 22*, 5–16.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 65*(3), 245–281.

Chen, Y.-C., & Tsai, C.-C. (2009). An educational research course facilitated by online peer assessment. *Innovations in Education and Teaching International, 46*(1), 105–117.

Cheng, W., & Warren, M. (1999). Peer and teacher assessment of the oral and written tasks of a group project. *Assessment & Evaluation in Higher Education, 24*(3), 301–304.

Cho, Y., & Cho, K. (2010). Peer reviewers learn from giving comments. *Instructional Science*. doi:10.1007/s11251-010-9146-1.

Cho, K., Schunn, C. D., & Charney, D. (2006). Commenting on writing: Typology and perceived helpfulness of comments from novice peer reviewers and subject matter experts. *Written Communication, 23*(3), 260–294.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*(4), 438.

Duijnhouwer, H. (2010). *Feedback effects on students' writing motivation, process, and performance.* Universiteit Utrecht. Retrieved from http://igitur-archive.library.uu.nl/dissertations/2010-0520-200214/UUindex.html.

Education and Manpower Bureau. (2007). Background. *Liberal Studies: Currlculum and Assessment Guide (Secondary 4-6).* Hong Kong.

Elawar, M. C., & Corno, L. (1985). A factorial experiment in teachers written feedback on students homework: Changing teachers behavior a little rather than a lot. *Journal of Educational Psychology, 77*(2), 162–173.

Ellman, N. (1975). Peer evaluation and peer grading. *The English Journal, 64*(3), 79–80.

Elwell, W. C., & Tiberio, J. (1994). Teacher praise: What students want. *Journal of Instructional Psychology, 21*(4), 322.

Falchikov, N. (1986). Product comparisons and process benefits of collaborative peer group and self assessments. *Assessment & Evaluation in Higher Education, 11*(2), 146–166.

Falchikov, N., & Blythman, M. (2001). *Learning together: Peer tutoring in higher education* (1st ed.). New York: 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*(3), 287.

Ferris, D. R. (1997). The influence of teacher commentary on student revision. *TESOL Quarterly, 31*(2), 315–339.

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2009). Improving the effectiveness of peer feedback for learning. *Learning and Instruction, 20*(4), 304–315.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112.

Henderlong, J., & Lepper, M. R. (2002). The effects of praise on children's intrinsic motivation: A review and synthesis. *Psychological Bulletin, 128*(5), 774–795.

Hughes, I. E. (1995). Peer assessment. *Capability, 1*, 39–43.

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*(2), 130–144.

Kluger, A. N., & DeNisi, A. (1996). Effects of feedback intervention on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119*(2), 254–284.

Kwok, R. C. W., & Ma, J. (1999). Use of a group support system for collaborative assessment. *Computers & Education, 32*(2), 109–125.

Law, N. W. Y., Lee, Y., van Aalst, J., Chan, C. K. K., Kwan, A., Lu, J., et al. (2009). Using Web 2.0 technology to support learning, teaching and assessment in the NSS Liberal Studies subject. *Hong Kong Teachers' Centre Journal, 8*, 43–51.

Li, L., Liu, X., & Steckelberg, A. L. (2010). Assessor or assessee: How student learning improves by giving and receiving peer feedback. *British Journal of Educational Technology, 41*(3), 525–536.

Lin, S., Liu, E., & Yuan, S. (2001). Web-based peer assessment: Feedback for students with various thinking styles. *Journal of Computer Assisted Learning, 17*(4), 420–432.

Liu, N.-F., & Carless, D. (2006). Peer feedback: The learning element of peer assessment. *Teaching in Higher Education, 11*(3), 279–290.

Liu, E. Z. F., Lin, S. S. J., Chiu, C. H., & Yuan, S. M. (2001). Web-based peer review: The learner as both adapter and reviewer. *IEEE Transactions on Education, 44*(3), 246–251.

Moskal, B. M. (2000). Scoring rubrics: What, when and how. Retrieved from http://pareonline.net/getvn.asp?v=7&n=3.

Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science*, *37*(4), 375–401.

Nilson, L. B. (2003). Improving student peer feedback. *College Teaching, 51*(1), 34–38.

Olson, V. L. B. (1990). The revising processes of sixth-grade writers with and without peer feedback. *Journal of Educational Research, 84*(1), 22.

Orsmond, P., Merry, S., & Reiling, K. (1996). The importance of marking criteria in the use of peer assessment. *Assessment & Evaluation in Higher Education, 21*(3), 239–250.

Page, E. B. (1958). Teacher comments and student performance: A seventy-four classroom experiment in school motivation. *Journal of Educational Psychology, 49*(4), 173–181.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis method* (2nd ed.). Newbury Park, CA: Sage.

Reuse-Durham, N. (2005). Peer evaluation as an active learning technique. *Journal of Instructional Psychology, 32*(4), 328–345.

Rowntree, D. (1987). *Assessing students: How shall we know them?* London: Kogan Page.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*(2), 119–144.

Sadler, P. M., & Good, E. (2006). The impact of self-and peer-grading on student learning. *Educational Assessment, 11*(1), 1–31.

Stefani, L. A. J. (1994). Peer, self and tutor assessment: Relative reliabilities. *Studies in Higher Education, 19*(1), 69–75.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan, 83*(10), 758–765.

Straub, R. (1997). Students' reactions to teacher comments: An exploratory study. *Research in the Teaching of English, 31*(1), 91–119.

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*(4), 291–303.

Topping, K. J. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*(3), 249–277.

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

Topping, K. J., & Ehly, S. W. (2001). Peer assisted learning: A framework for consultation. *Journal of Educational and Psychological Consultation, 12*(2), 113–132.

Topping, K. J., Smith, E. F., Swanson, I., & Elliot, A. (2000). Formative peer assessment of academic writing between postgraduate students. *Assessment & Evaluation in Higher Education, 25*(2), 149–169.

Tsai, C.-C. (2009). Internet-based peer assessment in high school settings. In L. T. W. Hin & R. Subramaniam (Eds.), *Handbook of research on new media literacy at the K-12 level: Issues and challenges* (pp. 743–754). Hershey, PA: Information Science Reference.

Tsai, C.-C., & Liang, J.-C. (2009). The development of science activities via on-line peer assessment: The role of scientific epistemological views. *Instructional Science, 37*(3), 293–310.

Tsai, C.-C., Lin, S. S. J., & Yuan, S.-M. (2001). Developing science activities through a networked peer assessment system. *Computers & Education, 38*(1–3), 241–252.

Tseng, S. C., & Tsai, C. C. (2007). On-line peer assessment and the role of the peer feedback: A study of high school computer course. *Computers & Education, 49*(4), 1161–1174.

van Gennip, N. A. E., Segers, M. S. R., & Tillema, H. H. (2010). Peer assessment as a collaborative learning activity: The role of interpersonal variables and conceptions. *Learning and Instruction, 20*(4), 280–290.

Van Lehn, K. A., Chi, M. T. H., Baggett, W., & Murray, R. C. (1995). *Progress report: Towards a theory of learning during tutoring*. Pittsburgh, PA: Learning Research and Development Centre, University of Pittsburgh.

van Zundert, M., Sluijsmans, D., & van Merriënboer, J. (2009). Effective peer assessment processes: Research findings and future directions. *Learning and Instruction, 20*(4), 270–279.

Wen, M. L., & Tsai, C. C. (2006). University students' perceptions of and attitudes toward (online) peer assessment. *Higher Education, 51*(1), 27–44.

White, K. J., & Jones, K. (2000). Effects of teacher feedback on the reputations and peer perceptions of children with behavior problems. *Journal of Experimental Child Psychology, 76*(4), 302–326.

Wilkinson, S. S. (1981). The relationship of teacher praise and student achievement: A meta-analysis of selected research. *Dissertation Abstracts International, 41*(9-A), 3998.

Xiao, Y., & Lucking, R. (2008). The impact of two types of peer assessment on students' performance and satisfaction within a Wiki environment. *The Internet and Higher Education, 11*(3–4), 186–193.

Yang, Y.-F. (2010). A reciprocal peer review system to support college students' writing. *British Journal of Educational Technology*. doi:10.1111/j.1467-8535.2010.01059.x.

Yang, Y.-F., & Tsai, C.-C. (2010). Conceptions of and approaches to learning through online peer assessment. *Learning and Instruction, 20*(1), 72–83.