Lee, H.-J., & Lim, C. (2012). Peer Evaluation in Blended Team Project-Based Learning: What Do Students Find Important? *Educational Technology & Society*, *15* (4), 214–224.

Peer Evaluation in Blended Team Project-Based Learning: What Do Students Find Important?

Hye-Jung Lee and Cheolil Lim
Department of Education, Seoul National University, Korea // hyejung@snu.ac.kr // chlim@snu.ac.kr

ABSTRACT

Team project-based learning is reputed to be an appropriate way to activate interactions among students and to encourage knowledge building through collaborative learning. Peer evaluation is an effective way for each student to participate actively in a team project. This article investigates the issues that are important to students when evaluating their peers in team project-based learning. A message analysis framework was inductively derived for the study, and data collected from the team-project learning process were categorized within this framework. Each message type was analyzed with respect to the students' peer evaluation results. The results showed that managerial, procedural, and social messages, rather than academic messages, significantly predicted peer evaluation results. These results indicate that students find social contributions, such as organizing or coordinating managerial abilities, more important than cognitive contributions when they evaluate peers. Additional results and the significance of their implications are discussed.

Keywords

Peer evaluation, Students' evaluation, Blended team project-based learning, Group learning evaluation

Introduction

In recent years, there has been an increasing demand for the development of effective teaching methods and instructional strategies to improve the quality of university education. Most university Centers for Teaching and Learning (CTLs) have invested in program development to improve their teaching methods. Instructional design that facilitates student-student interaction is considered to be an effective strategy to maximize learning through active student participation, which nurtures various social competencies during academic knowledge building.

Team project-based learning is one of the most commonly used methods to activate interactions among students. Team project-based learning has been introduced and is increasingly used as a teaching and learning method in higher education to promote knowledge building through social interaction (Von Kotze & Cooper, 2000). The Korean National Human Resources Development Council, an organization devoted to cultivating creative human resources, reports that team project-based learning has attracted educators' attention as an alternative teaching method for both improving the quality of teaching and enhancing learning effectiveness in higher education through social learning (Jung, 2001). In addition, team project-based learning promotes higher learning skills including cooperative ability, critical reasoning, creative thinking, responsibility, and communication (Moursund, 2003). A social learning context thus promotes both students' social and cognitive learning. Team project-based learning allows students to engage in the practice of knowledge building through a process of social investigation in a meaningful context. Therefore, teamwork competencies including communication, leadership, collaboration, and interpersonal relations, can be acquired during team-based social activities rather than in lectures or in individualized tasks. Many reports have been published outlining the advantages of online (or blended) team-based learning, e.g., student participation and interaction (Pena-Perez, 2000), social knowledge building (Stahl, 2000; Gunawardena et al., 1997), and critical thinking in online learning (Bullen, 1998; Newman et al., 1995).

However, there are also noted disadvantages of this type of team-based learning. Social loafing is the phenomenon of people exerting less effort to achieve a goal when they work in a group than when they work alone (Karau & Williams, 1993). Its two common manifestations are (1) Free-rider effect, where some members do not put in their share of work under the assumption that others' efforts will cover their shortfall, and thus cause (2) Sucker effect, where the other (fully performing) members lower their efforts in response to the free-riders' attitude (Kerr & Bruun, 1983; Salomon & Globerson, 1987). This type of social laziness indicates that team-based learning does not guarantee effective interactions in the classroom. Additionally, instructors are often overloaded with tasks to provide prompt and timely feedback for students to lessen social laziness because providing enough feedback is very effort-intensive (Dunlap, 2005; Ertmer, et al., 2010). The negative side effects of team-based learning may be difficult to resolve because an instructor may not observe all the processes occurring within the student groups. Typically, instructors evaluate the quality of the final product without knowledge of the team work process.

Peer evaluation may be a good strategy to monitor the dynamics within the group. Peer evaluation is an effective way of allowing every student to participate in team-based learning and monitor the process, as well as the product, of team learning. Peer evaluation is useful in higher education contexts because it is also expected to decrease instructors' workloads (Ertmer et al., 2010). However, peer evaluation has not been used extensively due to instructors' perceptions that such evaluations lack credibility. Student evaluation, although it effectively motivates students, has been considered less reliable than instructor evaluation. This controversy may arise from the fact that teachers and students evaluate different things. If instructors and students in fact evaluate different aspects of the learning process, this controversy should not be a matter of reliability or credibility. However, research on whether students and teachers evaluate the same criteria is scarce.

Therefore, this study investigated what students find important in their peer evaluations in team project-based learning. To facilitate data collection enabling the tracking of student activities, the team project-based learning in this study was implemented in a blended e-learning environment so that students' interactions could be recorded on the website easily. A blended team-based learning mode is similar to offline team-based learning except that students' interactions - both synchronous and asynchronous - are easily recorded on the website. This system enables students to reflect on what they have performed and allows teachers to track students' learning processes more easily. In our case, student activities recorded on the class website were primarily in the form of text messages. These messages were categorized and coded using a novel message analysis framework developed for this research. Regression analysis was run on the categorized messages and peer evaluation results in order to determine what components students consider to be important when evaluating their peers. The results have significant implications for quality team-based learning in higher education.

Theoretical Background

Peer Evaluation

Peer evaluation is a process in which students evaluate their peers' performance during group work or in class (Kane & Lawler, 1978). Research on peer evaluation within student groups has mostly focused on whether it is as credible as teacher evaluation. The credibility and effectiveness of peer evaluation are supported by many studies showing that peer evaluation is as valid as instructors' evaluations (Falchikov & Goldfinch, 2000; Wever et al., 2011). Gerard (2002) insisted that peer evaluation is a predictor of long-term success and the best indicator for evaluating performance. Peer evaluation is also reported to encourage individual responsibility (Liu et al., 2002) and to provide students the opportunity to reflect on their own work and to compare their work with that of others, which enhances students' meta-cognitive perception (Topping et al., 2000).

Peer evaluation, however, has also been criticized with respect to fairness. Brindley and Scoffield (1998) report that when the peers know one another, they do not assign each other poor scores, even when their performance is poor. Chang et al. (2011) also reported peer evaluation is not valid and reliable in their experimental study. A major weakness of peer evaluation was reported to be grade inflation or deflation (Balint et al., 2002). Students may take advantage of the system to raise their own grade. Conversely, some students may also be overly critical of their own performance and deflate their own scores (Blanche, 1988). In addition, the discrepancy between the ratings of students and experts is reported to be higher in group-work assessments then in individual-work (Sung et al., 2010).

These arguments collectively suggest that there are often differences in the results of evaluations performed by a teacher and those completed by students. However, many studies have apparently overlooked the features of students' performance that are being evaluated. Differences in evaluation results, if any, may be related to the fact that the evaluators have different points of view. Students may evaluate different aspects of their peers' performance than teachers do, precisely because they have a different vantage point from which to measure other students' work. Therefore, judging that peer evaluation lacks credibility based on only a teacher's viewpoint might be unfair. If students and teachers simply evaluated different things, the difference in evaluation results should not be a matter of reliability, and then student evaluations should not be given less weight than teacher evaluations.

Message Analysis Frameworks

Existing message analysis frameworks were not adequate for analyzing messages concerning students' interactions for this study. Message transcripts are a useful resource for investigating psycho-social dynamics, as Henri (1992) has indicated. Many message analysis frameworks have been presented in previous studies (Bullen, 1998; Fahy, 2001; Gunawardena et al., 1997; Henri, 1992; Newman et al., 1995; Zhu, 1996). In the early stage of message analysis framework development, Henri (1992) attempted to analyze student cognition from an interaction analysis and developed a content-analysis framework with five categories: participatory, social, interactive, cognitive, and meta-cognitive. However, although her framework provoked research on content analysis, these five categories proved to be ambiguous because they are not completely independent, and thus the categories do not have equal levels of orthogonal dimensions. For example, the 'participatory' and 'interactive' categories are behavioral, whereas the 'cognitive', 'meta-cognitive', and 'social' categories are thematic. Cognitive messages could be interactive and participatory, and of course, vice versa. Gunawardena et al. (1997) also reported problems using Henri's (1992) model to distinguish between cognitive and meta-cognitive activities in conferences due to the lack of precise criteria defining each category.

In fact, many message analysis frameworks have been criticized for category ambiguity, e.g., multiple postings in few categories (Gunawardena, et al., 1997) or messages coded into more than one category (Zhu, 1996). According to Fahy (2001), discriminant capability and reliability among users have been major problems in previous transcript analysis work. Discriminant capability refers to the function of a coding instrument that permits the direct and unambiguous placement of message content into discrete and useful categories. A lack of discriminant capability directly effects reliability, since unclear categories lead to discrepancies in coding. Furthermore, the previously developed instruments have tended to be too complex and contain too many codes, which hinder their application. Gunawardena et al. (1997) represented 20 categories in five phases, Cookson & Chang (1995) developed 16 codes, and Rourke et al. (2001) used 12 indicators. In addition, inter-coder reliability has often not been clearly delineated.

Regarding message content type, there are some content analysis frameworks specifically designed for capturing critical thinking processes (Bullen, 1998; Fahy, 2001; Newman et al., 1995). However, the cooperative interactions in team project-based learning differ from critics occurring in debates. Furthermore, the units of content analysis are diverse, including sentences, messages, and thematic units. For example, Fahy (2001) insisted that the data unit of a transcript analysis should be the sentence. Fahy appears to focus more on the format of the transcript rather than on the 'meaning' itself. However, we sought to uncover what students consider to be important in interactions with their peers. Therefore, an analysis based on units of meaning rather than sentence format was deemed more valid for this study precisely because there may be several meanings encoded in one sentence or only one meaning spanning several sentences.

In summary, the message analysis frameworks reported in the previous literature are conceptual frameworks derived deductively based on the researchers' logic. Those frameworks are not applicable to real message coding in our case study because they are not drawn from the actual coding of the messages found in an online team project-based learning context. Therefore, we inductively developed a new message analysis framework to clarify the analytic categories, which is described in the following section.

Research Methodology

Developing a Message Analysis Framework

To develop a message analysis framework, a thematic unit was defined as a coding unit. A primary coding scheme was then extracted from the existing literature (Bullen, 1998; De Wever, et al., 2006; Fahy et al., 2001; Gunawardena et al., 1997; Henri, 1992; Newman et al., 1995; Oren et al., 2002; Pena-Shaff & Nicholls, 2004; Zhu, 1996). Among the previous works consulted, the framework of Oren et al. (2002) was deemed to be the best fit for the macro categories analyzed in this study. Herein, cognitive messages are related to academic processes, whereas social messages are not focused on the academic content itself but instead concern building interpersonal relationships that foster smooth progress in a collaboration. Procedural messages include scheduling, timetables, to-do lists, and other various procedures related to the learning process. Each researcher independently conducted primary coding using these three macro-categories. Coding revisions were iterated for each category to better delimit the message categorization.

Table 1. The message analysis framework developed for this study

Categories of Oren et al. (2002)	Newly elaborated categories in this research	Definition	Student Examples
Cognitive (academic)*	Intellectual	Academic intellectual messages including activities for cognition, reflection, judgment, structuring, knowledge building, criticism, reasoning, and those related to the academic content.	*The author seemed to suppose that there were two types of instructional design and management styles. However, most practice lies in the middle of two extremes. Why do we need to compare structured instruction and interactive instruction? We need to implement both as much as we can. It is not a problem of selection. In my opinion, we need to add this point to our work. What do you think?*
	Informative	Academic, learning-content related but not related to the original intellectual activity, instead derived mainly from information from other resources and references.	*Attached file is the history of modern France royal family which I found on the internet for reference…* *Attachments : france-history.doc*
Social (relational)*	Social	Messages used mainly for social purposes, not related to academic learning content.	*Hi, all, the weather outside is too good for us to study all day long, isn't it? -,-;; But we're almost here, to the final goal, hurray!! Everybody, cheer up! ^_^…*
	Diplomatic	Messages not used primarily for social purposes but rather to facilitate smooth communication; additional diplomatic sentences at the beginning or end of the main message.	*I really appreciate your opinion. I'm so impressed by your insightful point. I agree some part of your suggestion. However,…*
Procedural (operational) *	Managerial	Messages regarding who does what and how, task allocation, coordination, organization, and mediation.	*..Ok, then let's wrap up our talking. Now A take part 1 and B take part 2. and C 'd better review A & B's work upon their request. If there is any problem, we can talk on chatting board and get to consensus, all right? …*
	Procedural	Messages related to procedures: scheduling, deadline management, assignment due date of the coursework.	*…When is our assignment due?* *Where should we upload our progress on the class site?* *See you at 11:00 pm on the messenger…*
	Technical	Messages about technical issues related to online learning activities	*…Download the attached file and unzip it. Then click front.html. If you know how to use PHP, then it'll be a lot helpful…*

* Articulation of Oren et al. (2002)'s categories for this research are within the parentheses.

During the coding revision process, we found that the three macro-categories of Oren et al. (2002) required further elaboration. For example, cognitive messages on academic process should be differentiated into *intellectual* messages, i.e., those conveying students' own ideas, and *informative* messages, representing the collection of ideas or references from on- and offline resources. In addition, social messages were also sub-categorized into multiple types. Oren's social messages are related to general social relationships removed from academic content. However, we found that some messages directly served the purpose of building rapport while other messages relayed diplomatic courtesies before or after the main (either cognitive or procedural) message; these messages were not considered to be intended primarily for social relationship building. For our purposes, the former were classified as

social messages, whereas the latter were classified as *diplomatic* messages. Finally, procedural messages were differentiated into managerial, procedural, and technical messages. We defined *managerial* messages as messages concerned with who does what and how, task allocation, coordination, organization, and mediation. *Procedural* messages specifically concerned scheduling, deadlines, the assignment process or the coursework, etc., whereas *technical* messages were focused on technical issues related to online learning activities. *Managerial, procedural,* and *technical* messages are all related to 'operation' of the team process, whereas *intellectual* and *informative* messages are about 'academic' quality of the team product. *Social* and *diplomatic* messages contribute the whole process more smoothly and communally, so it can be articulated as 'relational' category.

In summary, seven categories were inductively derived from Oren's original three categories to comprise the content-analysis framework for this study (See Table 1): Cognitive (Academic: Intellectual, Informative), Social (Relational: Social, Diplomatic), and Procedural (Operational: Managerial, Procedural, Technical). Three external experts (PhDs in Educational Technology) reviewed all categorizations for verification. Inter-coder agreement among the external coders, based on Cohen's Kappa Coefficient, was .87. Detailed development process of the message analysis framework is reported separately on Lee & Kim (2011).

Data Collection and Analysis

Thirty-two undergraduate students enrolled in Instructional Methods and Educational Technology at Seoul National University participated in this study. Their majors were diverse, including humanities, social science, natural science, and other subjects. The students in the class were organized into eight teams of four students each. Each team was assigned two project tasks: a theoretical review task and a development task. Although they were given the same task requirements, each team developed their own unique topic. Students interacted with their team members mainly through asynchronous online discussion boards several times a week (synchronous chatting board was also used but not as often as the asynchronous site). All students were very familiar with online communication. Student teams sometimes had offline meetings, but more often they met online, and all of their discussions and brainstorming ideas were recorded by a member of each group and uploaded on the team project website so that researchers were able to track and analyze students' interactions. The instructor was able to read all of the recorded interactions on each team's board but did not intervene during team communications. These team interactions included only communications among students within each team. If there were any general questions for the instructor, students posted their questions on the class Q/A board or asked the instructor directly via email.

Peer evaluations can involve peer nomination, peer rating, and peer ranking (Kane & Lawler, 1978). Peer nomination, for example, consists of nominating peers with the highest or the lowest performance in each of the evaluated items. This method is reputed to display the highest reliability and validity. Peer rating, known as the most useful method with the least bias, refers to simply rating peers on each of the evaluation items. Peer ranking involves ranking peers in relative order from the best performers to the worst performers. This last method may be the best way to differentiate each of the peer evaluation scores but has hardly been investigated. In this study, peer rating was used in the initial implementation, but the rating scores were translated into peer ranking because there were wide variations in the individual perceptions of the best and worst performers.

Peer evaluation was performed after two major team projects for each group during the semester. The peer evaluation report was not made public. The peer evaluation form used in this research included the following items: 1) Participated in group project or meetings; 2) Helped keep the group focused on the task; 3) Contributed useful ideas; 4) Quantity of work performed; and 5) Quality of work performed. Students were instructed to award up to five points for each item, for up to twenty-five total points on the evaluation form. Students were not permitted to assign the same score for every student Relative peer ratings, rather than the direct raw scores, were used for the analysis because the relative importance of student perceptions rather than the absolute scores was deemed more significant in this work. Additionally, each ranking was weighted. Because there were four members in each team, the rankings were from first to fourth, and the corresponding weighting values were from four to one. Thus, the first-ranked student received four points for each item, the second-ranked student received three points, the third-ranked student received two points, and the lowest-ranked student received only one point. These peer-ranking points were analyzed with respect to the content of the messages related to each category to determine what types of messages the higher-ranking students posted.

Results and Discussion

Finally, 773 postings were posted on the asynchronous board. These messages were categorized using the content-analysis framework and subsequently analyzed for correlation with the results of the peer evaluation by regression analysis. There were 1,814 thematic message units (there could be several thematic message units per one posting); including 560 intellectual, 256 informative, 161 social, 277 diplomatic, 108 managerial, 381 procedural, and 71 technical message units. The intercorrelations and descriptive statistics for all study variables are summarized in Table 2.

Table 2. Intercorrelations and Descriptive Statistics for Study Variables

	Intellectual	Informative	Social	Diplomatic	Managerial	Procedural	Technical
Peer evaluation	.041	.066	.279	.037	.405**	.209	-.072
Intellectual	1	.582**	.619*	.545**	.305	.592**	.112
Informative		1	.344	.493**	.204	.390*	.218
Social			1	.517**	.532**	.258	.025
Diplomatic				1	.400*	.597**	.132
Managerial					1	.225	.014
Procedural						1	.009
Technical							1
No. of units	2-63	2-61	2-8	2-64	2-12	2-69	2-5
Mean	17.50	8.00	5.03	8.66	3.38	11.91	2.22
(SD)	(14.92)	(11.32)	(2.12)	(11.37)	(2.20)	(14.81)	(0.75)

Notes. No. of units indicates the minimum and the maximum number of the thematic message units. Mean is the average number of the thematic message units by each student. *p<.05, **p<.01

Managerial, Procedural, and Social Contributions

Students evaluated their peers' managerial, procedural, and social contributions as being more important than cognitive contributions. Table 3 presents a summary of the regression results, showing that managerial, procedural, and social messages of student activities were significantly correlated with the peer evaluation results, whereas cognitive messages were not significant. This difference indicates that the students who conveyed more managerial, procedural, and social messages received higher scores from their peers, whereas cognitive message contribution did not significantly influence peer evaluation results.

Table 3. Regression Analysis Results: Message Types Predicting the Overall Peer Evaluation Score

Variables	**Standardized coefficient**		**F**
	β	**Standard Error**	
Intellectual	-1.250	.692	3.269
Informative	.304	.228	1.781
Social	**.571**	**.315**	**3.285****
Diplomatic	-.511	.522	.957
Managerial	**.580**	**.199**	**8.488***
Procedural	**1.144**	**.575**	**3.960****
Technical	-.293	.204	2.063

Note: Dependent variable is peer evaluation
*p<.05, ** p<.01

Social messages acted as a type of lubricant that allowed the collaboration process to function more smoothly. Procedural and managerial messages led the progress of the team project and kept all team members working – one goal. The results indicate that students perceived the social role of harmonious collaboration and steady progress to be more important than cognitive contributions. Students seemed to perceive collaborative competency as more necessary for successful team learning than individual cognitive excellence.

This finding has significant implications for the debate on peer vs. instructor evaluation. When grading students, instructors usually cannot observe what occurs within each group. Therefore, their grading must be more focused on the academic achievement evident in the final *product*, regardless of the nature of the team's collaborative *process*. Instructors rarely evaluate the process of teamwork or other social competencies, which, in addition to academic knowledge in each discipline, universities seek to nurture through their curriculum. Our results therefore imply that peer evaluation may be a good strategy for *process* evaluation and can supplement instructor evaluation.

Intellectual Messages (Quality) vs. Informative Messages (Quantity)

Analyzing these relationships in further detail (Table 4), we found that the numbers of informative and procedural messages specifically predicted scores on the peer evaluation question "participation in group project or meetings". Procedural and diplomatic messages influenced the scores on "helped keep the group focused on the task"; social messages predicted "contributed useful ideas" scores. Informative, diplomatic, and procedural messages were related to the quantity of the work performed, whereas intellectual messages predicted the quality of the completed work.

As shown in Table 4, an interesting result emerged between informative messages and intellectual messages. Students' evaluations of "informative" and "procedural" messages were predicted by the quantity rather than the quality of the work, whereas they considered "intellectual" messages to represent the quality of the work. This distinction supports two conclusions: 1) the findings justify the use of the newly modified message analysis framework in this study, in which "informative" messages are differentiated from "intellectual" messages, whereas both were categorized broadly into "academic cognitive" messages in the previous literature; and 2) student evaluation is quite reasonable because the students distinguished "intellectual" messages from others and credited these messages toward the quality of work.

In the previous literature, credibility or reliability issues have been reported as factors in the reluctance to adopt a wider use of peer evaluation, because it may often be perceived as an unfair evaluation method. However, the results of this study showed that students clearly distinguished quality contributions from quantity contributions and evaluated their peers quite reasonably in each area. This finding implies that peer evaluation is a valid and fair strategy to evaluate team members' efforts and contributions.

Furthermore, although students evaluated "intellectual" messages as being related to the "quality of the work", they did not count them as critical contributions, unlike the sum of other social and procedural messages. In other words, students considered "managerial", "procedural", and "social" messages to be the most important, although these were not "intellectual" messages. Social messages such as compliments and appreciation for other's efforts were also evaluated positively in this research. Therefore, students seem to perceive that the most critical factor in team collaboration is not each individual's intellectual participation but rather each team member's social collaboration, which facilitates the progress of the project. Instructors would not be able to evaluate the various types of participation of each student within the group because instructor evaluations are usually based on the final product. The results of this study imply that peer evaluation can complement instructor-only evaluation.

Table 4. Regression Analyses Results: Message Types Predicting Scores on the Five Components of the Peer Evaluation

Peer evaluation question (Dependent variables)	**Independent variables**	**Standardized coefficient**		**F**
		β	**Standard Error**	
Participated in group project or meetings	Informative	.606	.213	8.098**
	Procedural	1.174	.597	3.869*
Helped keep the group focused on the task	Diplomatic	.606	.213	5.743**
	Procedural	1.174	.597	6.039**
Contributed useful ideas	Social	.443	.213	4.325*
	Informative	.487	.210	5.392**
Quantity of work performed	Diplomatic	-.831	.361	5.301*
	Procedural	.579	.292	3.939*
Quality of work performed	Intellectual	1.150	.567	4.119*

*p<.05, ** p<.01

Proactive vs. Reactive

Students also evaluated proactive participants who posted messages as more important than passive members who only read messages. Table 5 shows the results of regression analysis examining how posting and reading activities predicted peer evaluation results. The results indicate that the frequency of posting predicted the peer evaluation results significantly, whereas the frequency of reading did not. That is, the more frequently a student posted messages, the higher score the student received from peers in their evaluations. However, reading frequency did not contribute to their score in the team project-based learning context.

Table 5. Regression Analyses Results: Posting and Reading Frequencies Predicting Peer Evaluation Scores

Independent variables	Standardized coefficient		F
	β	Standard Error	
Frequency of Posting messages	**.844**	**.312**	**7.305****
Frequency of Reading messages	.182	.302	.363

Note: Dependent variable is peer evaluation
** p<.01

In individual learning activities, both reading and posting messages have been reported as critical learning achievement factors (Lee, 2008; Fulford & Zhang, 1993). In instructor evaluation, "reading-only" activity as well as "also-posting" activity showed positive effects on academic achievement, especially in individual receptive learning where knowledge reception, rather than critical or creative idea generation, is required more importantly. On the other hand, in team project-based learning, observable proactive participation appeared to be more significant for the students. This difference may be attributed to the interaction-based communal learning context. In individual cognition-based receptive learning, students can acquire knowledge sufficiently well by reading the text alone, without observable participation. In contrast to the receptive learning, the students were required to actively participate in learning activities, leading and participating proactively in interaction-based communal learning. Therefore, for more effective team-based collaborative learning, instruction should be designed to allow students to accurately perceive what constitutes proactive contributions and how to encourage those in addition to self-learning.

Summary and Conclusions

This research sought to determine what components of peer interaction students consider to be important when they evaluate their peers' contributions in blended team project-based learning. For this research, first, the content-message analysis frameworks outlined in the previous literature were reviewed. Then, because the majority of the previously developed frameworks were not appropriate for coding the contents of the messages collected in this work, an applicable framework for content-message analysis was empirically developed through iterative revision of a customized coding scheme. Three main categories adapted from previous works were elaborated into seven new categories to create a novel message analysis framework. The new framework included the following categories: Intellectual, Informative, Social, Diplomatic, Managerial, Procedural, and Technical. Next, we examined which message categories predicted the peer evaluations. Peer evaluations within each group were measured after two major team projects were completed in one semester. Students were required to post every activity and interaction on their group website; this record enabled teachers to easily track and analyze the students' activities. A total of 773 messages including 1,814 thematic message units posted by 32 students (eight four-member groups) were categorized using the message analysis framework. In order to determine which categories students use to evaluation their peers, we ran regression analysis to determine what message categorizations predicted the actual peer evaluations that students had given to their peers.

The results showed that the more managerial, procedural, and social messages students posted during the team project-based learning, the higher score they received from the other team members in their peer evaluations. Moreover, students clearly differentiated quantity contributions from quality contributions. Students evaluated informative messages, such as resources or references, as being related to the quantity of the work, whereas intellectual messages, such as an individual's own ideas, opinions or thoughts, were related to the quality of the work. Although the students perceived intellectual messages as being related to the quality of their teamwork, they valued social and managerial contributions more significantly than cognitive contributions in the comprehensive evaluations.

These results imply that students evaluate their peers on contributions (managerial, procedural, and social) that an instructor would not easily be able to observe. Instructors' evaluations tend to place a greater emphasis on the outcome or product of teaming rather than on the process of teaming itself (Wang, 2011). Our results thus suggest that controversy regarding the fairness of peer evaluation should not be a matter of credibility or reliability. The difference, if any, between peer and instructor evaluation must instead be based on the differing evaluation criteria. Students' peer evaluations were considered fair because students perceived each of the different types of contributions clearly, implying that student peer evaluation can be employed as a complementary strategy, especially in communal learning. Students focus on social competencies, such as organizing or coordinating abilities to keep the project moving forward, as important peer contributions. Teachers are unable to deeply access the process of team collaboration, especially when there are many students in a class, although team project-based learning is specifically designed to nurture social competency as well as other intellectual achievements. Therefore, the finding that students can evaluate the social competencies that an instructor cannot readily observe suggests that student peer evaluations constitute a useful, credible, and complementary strategy for instructors in assessing students' social competencies beyond individual learning competencies. As an instructional strategy to encourage students' active interaction within each team, teachers could announce at the beginning of the course that peer evaluation results may be included in the final grading.

In addition, the finding that the students who proactively post messages are preferred over reactive and receptive students who only read messages in team project learning indicates that strategies for team project-based learning should differ from strategies for receptive individual learning. Therefore, for more effective instructional design in team project-based learning, instructors should clearly explain in detail what types of participation and contributions are expected in their team project-based learning. Students must be aware of the differences in necessary engagement types between individual receptive learning and communal interactive learning.

Finally, this study highlights the importance of social competencies for a communal society. A report from the Korean Human Resources Research Center (2009) shows how employers' and professors' perceptions differ regarding the core competencies that college students should acquire. In the report, employers indicated that social skills(19%) are the most critical abilities for college students to have, followed by attitude(18%), logical thinking skill(15%), leadership(14%), learning ability(12%), creativity(11%), and academic knowledge(11%), whereas professors stated academic knowledge(27%) is the most critical ability for college students to nurture, followed by creativity(27%), attitude(19%), logical thinking skill(13%), leadership(9%), learning ability(3%), and social skill(2%) as the least critical. Another recent news article reported that students at Seoul National University (the top university in Korea) lack group sociability (Yonhap News, 2012); the director of the Career Development Center at SNU stated, "…According to a survey of employers, SNU students show outstanding capabilities in academic knowledge and logical thinking but poor ability in collaboration, interpersonal relationship, and leadership, so how to improve these sociabilities is a critical issue at SNU…" This study contributes to the literature through the novel finding that students do not regard individual cognitive ability to be as critical as managerial, procedural, and social contributions which appear to be more influential in collaborative learning, especially in improving group relations and sustaining group work, unlike in individual learning.

The results and implications of this study show that peer evaluation can facilitate the authentic goal of team project-based learning. The learning goal of team project-based learning is best achieved when teams are effectively collaborating. If one smart student handles most of the team project and produces a high quality end product with little collaboration, it is missing the point of meaningful team project-based learning. In fact, a previous study demonstrates that high-achieving students tend to dominate the team assignment for their good grade, rather than sharing and collaborating socially (Lee et al., 2011). To facilitate collaboration from every student, evaluation of the team working *process* would be necessary. This study shows peer evaluation can contribute to the *process* evaluation. Therefore, peer evaluation can be recommended as a useful strategy to encourage and support social competencies, especially in higher education, because the professors are not sufficiently aware of the necessity and significance of social abilities, whereas employers consider sociabilities to be of critical importance.

There are some limitations of this study and suggestions for future research. First, this study is about how *students, not instructors*, evaluate their peers in team project-based learning. Future work may want to compare the relationships between learning outcome by instructor's evaluation, peer evaluation, and interaction message types. Second, researchers may want to not only study cognitive learning outcome, but also social learning or acquisition of social skills in team project learning, since they were closely related to the interaction message types and peer

evaluation. Third, this is a case study of one type of course with limited number of students. Researchers should examine additional cases of team project-based learning with broader backgrounds of students before making broad generalizations of the findings in this study.

References

Balint, M., Bodner, D., & Evans-Nachi, H. (2002). Self- and Peer- Evaluation of Group Discussion. *Proceedings of JALT 2001* (Shizuoka).

Blanche, P. (1988). Self-assessment of foreign language skills: Implications for teachers and researchers. *RELC Journal, 19*, 75–96.

Brindley, C., & Scoffield, S. (1998). Peer assessment in undergraduate programmes. *Teaching in Higher Education, 3*(1), 79–89.

Bullen, M. (1998). Participation and critical thinking in online university distance education. *Journal of Distance Education, 13*(2), 1–32.

Chang, C.-C., Tseng, K.-H., Chou, P.-N., & Chen, Y.-H. (2011). Reliability and validity of Web-based portfolio peer assessment: A case study for a senior high school's students taking computer course. *Computers & Education, 57*(1), 1306–1316.

Cookson, P. S. & Chang, Yu-bi. (1995). The Multidimensional Audioconferencing Classification System (MACS). *The American Journal of Distance Education, 9*(1), 18–36.

De Wever, B., Schellens, T., Valcke, M., & Van Keer, H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & Education, 46*(1), 6–28.

Dunlap, J. C. (2005). Workload reduction in online courses: Getting some shuteye. *Performance and Improvement,44*(5), 18–25.

Ertmer, P.A., Richardson, J.C., Belland, B., Camin, D., Connolly, P., Coulthard, G., Lei, K., & Mong, C. (2010). *Impact and Perceived Value of Peer Feedback in Online Learning Environments.* Retrieved June 7, 2010, from http://www.edci.purdue.edu/ertmer/docs/AECT05_Proc.pdf .

Fahy, P.J. (2001). Addressing some common problems in transcript analysis. *International Review of Research in Open and Distance Learning*, *1*(2), Retrieved June 23, 2010, from http://www.irrodl.org/index.php/irrodl/article/view/321/530.

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), 287–322.

Fulford, D. P., & Zhang, S. (1993). Perceptions of interaction; the critical predictor in distance education. *The American Journal of Distance Education, 7*(3), 8–21.

Gerard, R, J. (2002). *Peer evaluation as a predictor of future success*. Unpublished doctoral dissertation, Pennsylvania State University.

Gunawardena, C., Lowe, C., & Anderson, T. (1997). Analysis of a global on-line debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research, 17*(4), 395–429.

Henri, F. (1992). Computer conferencing and content analysis. In A. Kaye (Ed.), *Collaborative learning through computer conferencing: The Najaden papers* (pp.117–136). London: Springer-Verlag.

Jung, S. (2001). *Effects of Basic Thinking Skills and Project Method on Creativity and Project Performance Ability of Elementary School Children*. Doctoral Dissertation, Kyungsung university Press.

Kane, J., & Lawler, E. (1978). Methods of peer assessment. *Psychological Bulletin, 85*(3), 555–586.

Karau, S. J. & Williams, K. D. (1993). "Social loafing: A meta-analytic review and theoretical integration". *Journal of Personality and Social Psychology*, *65* (4), 681–706.

Kerr, N. L., & Bruun, S. E. (1983). Dispensability of member effort and group motivation losses: Free rider effects. *Journal of Personality and Social Psychology, 44*, 78–94.

Korean Human Resources Research Center Report (2009). *Core competencies to be nurtured in college education: stated by employers and professors.* Seoul National University.

Lee, H.-J. (2008). A Learning Process Mechanism in CSCL(Computer Supported Collaborative Learning). *Proceedings of ICCE(International Conference of Computers in Education) 2008* (Taiwan).

Lee, H.-J. & Kim, I. (2011). Development of a Content Analysis Framework in Online Group-Based Learning. *Journal of Lifelong Learning Society, 7*(3), 109–127.

Lee, H.-J., Kim, H., & Byun, H. (2011). How are high-achieving college students doing in team project-based learning?: An institution-wide analysis at a Korean top University. *Proceedings of (ICER) International Conference of Educational Research*.

Liu, E. Z. F., Lin, S. S. J., & Yuan, S. M. (2002). Alternatives to instructor assessment: A case study of comparing self and peer assessment with instructor assessment under a networked innovative assessment procedures. *International Journal of Instructional Media, 29*(4), 10.

Moursund, D. G. (2003). *Project-based learning using information technology*. Eugene, OR: International Society for Technology in Education.

Newman, D., Webb, B., & Cochrane, C. (1995). A content analysis method to measure critical thinking in face-to-face and computer supported group learning. *Interpersonal Computing and Technology: An Electronic Journal for the 21st Century, 3*(2), 56–77.

Oren, A., Mioduser, D. & Nachmias, R. (2002). The development of social climate in virtual learning discussion groups. *International Review of Research in Open and Distance Learning, 3*(1). Retrieved March 15, 2011, from http://www.irrodl.org/index.php/irrodl/article/view/80/154

Pena-Perez, J. (2000). *Participation, interaction and meaning construction in a university-level course using a computer bulletin board as a supplement to regular class discussions: A case study*. Unpublished doctoral dissertation, Cornell University, Ithaca, NY.

Pena-Shaff, J., & Nicholls, C. (2004). Analyzing student interactions and meaning construction in Computer Bulletin Board (BBS) discussions. *Computers and Education, 42*, 243–265.

Rourke, L., Anderson, T., Garrison, D. R., & Archer, W. (2001). Methodological issues in the content analysis of computer conference transcripts. *International Journal of Artificial Intelligence in Education, 12.*

Salomon, G. & Globerson, T. (1987). When teams do not function they way they ought to. *International Journal of Educational Research, 13*, 89–100.

Stahl, G. (2000). A Model of Collaborative Knowledge-Building. In B. Fishman & S. O'Connor-Divelbiss (Eds.), *Fourth International Conference of the Learning Sciences* (pp. 70-77). Mahwah, NJ: Erlbaum.

Sung, Y.-T., Chang, K.-E., Chang, T.-H., & Yu, W.-C. (2010). How many heads are better than one? The reliability and validity of teenagers' self- and peer assessments. *Journal of Adolescence, 33*(1), 135–145.

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

Von Kotze, A., & Cooper, L. (2000). Exploring the transformative potential of project-based learning in university adult education. *Studies in the Education of Adults, 32*(2), 212.

Wang, J. (2011). Work in progress - A feedback system for peer evaluation of engineering student teams to enhance team effectiveness. *Proceedings of Frontiers in Education Conference (FIE),* pp.S4C-1–S4C-5.

Wever B. D., Keer, H. V., Schellens, T., & Valcke, M. (2011). Assessing collaboration in a wiki: The reliability of university students' peer assessment. *The Internet and Higher Education, 14*(4), 201–206.

Yonhap News (2012, February 17). *SNU students lack sociability*. Retrieved from February 17, 2012, from http://news.naver.com/main/read.nhn?mode=LSD&mid=sec&sid1=101&oid=001&aid=0005518315.

Zhu, E. (1996). Meaning negotiation, knowledge construction, and mentoring in a distance learning course. *Proceedings of Selected Research and Development Presentations at the 1996 National Convention of the Association for Educational Communications and Technology.* Indianapolis: Available from ERIC Documents: ED 397 849.