Analysis of Peer Evaluation Essay in Problem-Based Learning

Shin-Il Han

Department of Education, Sungkyunkwan University, Seoul, Korea
Email: sihan@skku.edu

Abstract. The purpose of this study is to analyze peer evaluation essays to determine what college students consider important when evaluating their peers in a PBL class. Discovering and comparing students' comments regarding high achievers and low achievers during peer evaluation are the specific study objectives. Eleven students taking a class that utilized PBL participated in this study. The participants were tasked with submitting peer evaluation essays after the completion of PBL group activities. Each student was asked to first rank his or her peers and then to explain in essay form the reasons why he or she ranked them in that order. The analytic process for the peer evaluation essay involved qualitative research methods comprised of (1) examining every sentence in the students' essays and classifying similar words or phrases into a provisional concept, (2) a review by an external educational expert, (3) and member-checking by participating students. Triangulation and member-checking were used for the validation of the study. Upon grouping the sentences in the peer evaluation essays according to similar meanings, words, or concepts, a convergence on 6 to 8 common factors emerged. Eight factors (participation, cognition, preparation, communication, leadership, responsibility, affectivity, skills) were observed when the criteria for the high achievers at peer evaluation were examined. Six factors (participation, cognition, preparation, communication, diligence, cooperation) were identified when the criteria for the low achievers at peer evaluation were analyzed. In addition, factors of cognition, leadership, and participation were most frequently and positively mentioned by students when evaluating high achievers within their groups, while participation and diligence were most frequently and negatively mentioned when evaluating low achievers. Based on the findings, various subject matters are discussed.

Keywords: Peer evaluation essay, problem-based learning, qualitative research, higher education, cooperative learning.

1   Introduction

Cooperative learning is an instructional paradigm in which teams of students work on structured tasks under conditions that meet five criteria: positive independence, individual accountability, face-to-face interaction, appropriate use of collaborative skills, and regular self-assessment of team functioning (Kaufman, Felder, & Fuller, 2000). According to their assertion, well-practiced cooperative learning improves information acquisition and retention, higher-level thinking skills, interpersonal and communication skills, and self-confidence. Problem-based learning (PBL), one of the major cooperative learning methods, is a student-centered approach in which students learn material in small groups by addressing or solving an authentic and complex problem (Elder, 2009). Collaborative and PBL strategies are theorized to be effective methods for strengthening various courses (Curran, Carlson, & Celotta, 2013). PBL helps students learn skills for problem-solving, collaboration, and self-directed learning, and increases their adaptive knowledge and mastery motivation (Hmelo-Silver, 2004). In addition, the cooperative learning assessment methods of PBL have been shown to be fairly potent in their own right, so incorporating those methods into PBL may add value and extend the advantages intrinsic to PBL, regardless of the assessment method (Albanese & Dast, 2014).

  One of the tantalizing problems in cooperative learning environments such as PBL is that it is difficult to evaluate the contributions of individual students to team activities. As one solution to the problem, peer evaluation is often recommended. Peer evaluation has been used for the determination of individual contributions to small-group projects (Goldfinch, 1994). Felchikov (1986) used peer group

process assessment questionnaires to help promote competence, confidence, creativity, coping, and cooperation in a four-person group for a film-marketing project. Ohland, Laton, Loughry, and Yuhasz (2005) mentioned that peer raters often provide better, more accurate feedback, suggesting that instructors should carefully evaluate the benefits and drawbacks of using peer evaluations to adjust students' grades if the main objective is formative assessment. DeNisi, Randolph, and Blencoe (1983) also said that peers make finer distinctions among different aspects of performance than supervisors do; consequently, feedback from peers is more effective than supervisor ratings in eliciting behavior changes. Furthermore, a study by Strachan and Wilcox (1996) revealed that when peer assessment was used, students perceived it to be fair, valuable, enjoyable, and helpful in developing transferable skills in research, collaboration, and communication. Especially, in terms of achieving learning goals, peer evaluation can facilitate the authentic goal of team-based learning, because the goal of team-based learning is best achieved when teams are effectively collaborating (Lee & Lim, 2012).

In peer evaluation, group members judge their fellow members on specific traits, behaviors, and achievements (Kramer, 1990). The number of reviews performed per reviewer and the academic performance of the students on the assignment are reported to be important in peer assessment (Jeffery, Yankulov, Crerar, & Ritchie, 2016). Additionally, the traits, behaviors, and achievements are mostly suggested by instructors in the form of a questionnaire. The peer evaluation questionnaire developed by Segers and Dochy (2001) included items related to contribution to group discussions, quality of the contributions, preparedness for involvement in tasks, and actual contributions to teamwork. Also, Duzer and McMartin (2000) developed an assessment tool for self-evaluation and the evaluation of team members. Their questionnaire included the following items: (1) failed to do an equal share of the work, (2) kept an open mind/ was willing to consider the ideas of others, (3) was fully engaged in discussions during meetings, (4) took a leadership role in some aspects of the project, (5) helped group outcome differences to reach effective solutions, (6) often tried to excessively dominate group discussions, (7) contributed useful ideas that helped the group succeed, (8) encouraged the group to complete the project on a timely basis, (9) delivered work when promised/needed, (10) had difficulty negotiating issues with members of the group, (11) communicated ideas clearly/effectively. Another peer evaluation form developed by Ofori, Pai, and Layton was for an engineering course, and included the items (1) attended meeting regularly, (2) contributed to discussions, (3) had good communication skills, (4) committed to group goals, (5) listened effectively, (6) took responsibilities seriously, (7) accepted criticism gracefully, (8) performed significant tasks, (9) tasks had technical content, (10) completed tasks on time (Ohland, Laton, Loughry, & Yuhasz, 2005).

On the other hand, a two-year consecutive study of peer evaluation by Higgitt (1996) involved a list of peer evaluation criteria generated by the entire class, not by the instructors or researchers. His goal for using peer evaluation in the course was to enhance students' appreciation with student-led participation. The peer assessment criteria in the first year in his study were ranked in the following order: participation and enthusiasm, cooperation and flexibility, reliability, originality (ideas and interest), originality (design flair), self-motivation, collection of data, and explanation. However, the peer assessment criteria in the second year were quite different from the previous year. They were ranked in the following order: integration (listening and participation), collection and analysis of information, ideas, enthusiasm and contribution, contribution to design and style, effort, competence, and co-operation on final report.

As can be seen from previous studies, most of the criteria for peer evaluation questionnaires were either developed by instructors/researchers or were polled opinions from the whole class. Using these criteria, the personal opinions or evaluation criteria of individual students are not reflected in the results of peer evaluation in cooperative learning, even though most cooperative learning is based on collaboration between small numbers of students. In other words, smaller scale class activities need to be based on the opinions of each of the group members. From this point of view, the peer evaluation forms described earlier are not suitable for identifying individual members' unique opinions, which are important for evaluating the traits, behaviors, and achievements of fellow members. This study was motivated by the need to understand the variety of peer evaluation criteria used by individual students.

The purpose of this study is to analyze peer evaluation essays in order to identify what individual students consider important when evaluating their peers in a college level PBL class. Especially, discovering and comparing students' comments for high achievers and low achievers during peer evaluations are the more detailed objectives for the analysis of peer evaluation essays.

2   Methodology

2.1   Study Design

This study used a qualitative research method to analyze students' peer evaluation essays. All the comments made by students in the peer evaluation essays were grouped respectively by the rank order, i.e., all comments made by a student regarding the individual they ranked first within their group were grouped together and the next set of comments were grouped in the same manner. Finally, all the comments for the high achievers of each group were collected, examined, and classified as positive peer evaluation comments for high achievers. On the other hand, all the comments for the low achievers of each group were gathered, looked over, and categorized as negative peer evaluation comments for low achievers. The results of the summarized and classified comments for high and low achievers in the peer evaluation were compared and explained.

2.2   Participants and Data Collection

The study participants consisted of 11 college students taking a general education requirement course that utilized PBL as an important method of learning. A peer evaluation essay had to be submitted by the participants right after completion of the PBL group activity. Each of the PBL group members was asked to first rank his or her group members and then to give reasons in essay form why he or she ranked them in that order. In order to keep the students' peer evaluation essays confidential, they were collected through the instructor's e-mail. Demographic information is displayed in Table 1.

Table 1. Descriptive Statistics for Participants.

Variables		School Year		Total
		Junior	Senior	
Gender	Male	3	5	8
	Female	1	2	3
Total		4	7	11

2.3   Data Analysis

The analytical process for the peer evaluation essays began with the analysis of every sentence in the students' essays and the classification of similar words or phrases into a provisional concept by the principal researcher. Over the course of the analysis, previously analysed phrases or words were reclassified or the provisional concepts were redefined. The first draft of the classification table for students' peer evaluation essays was generated following repetition of the analysis process.

Second, an educational expert that specialized in qualitative studies and education was invited to review the first draft of the classification table. Based on the result of the expert's examination, the first draft was revised, and the second draft was created.

Next, the second draft was sent to two students that participated in the class. The students were asked to examine the second draft and provide feedback. The feedback contained detailed explanations of the students' written comments, and deepened the understanding of the intent behind the comments. Finally, by revising the second draft based on the feedback from the two students, the third and final version of the analysis of the students' peer evaluation essay was performed.

2.4   Validation of Data

Triangulation and Member-checking are employed in this qualitative study to corroborate the findings from the analysis. Triangulation refers to a qualitative research strategy using multiple methods or data sources to develop a comprehensive understanding of phenomena (Carter, Bryant-Lukosius, DiCenso, Blythe, & Neville, 2014; Patton, 1999). The major forms of triangulation are data triangulation,

investigator triangulation, theory triangulation, and methodological triangulation. Investigator triangulation in this study consisted of recruiting two educational and qualitative experts to examine the classification of comments in the peer evaluation essays. Investigator triangulation refers to the participation of two or more researchers in the same study to provide multiple observations and conclusions, and ensures both the confirmation of findings and different perspectives, adding breath to the phenomenon of interest (Denzin, 1978).

Member-checking is the process of sending ideas back to research participants for confirmation and collecting materials to expand researcher's categories (Charmaz, 2006). Through this process, member validation is performed by obtaining the views of participants on the accuracy of the data gathered, descriptions, or even interpretations. Such checks are performed through the return of transcripts or interpretative summaries to individual participants for verification and confirmation of the accuracy, or through participant focus groups (Doyle, 2007; Harvey, 2015). In this study, two students were asked to review the second draft of the classification table or the peer evaluation essays. Based on their feedback, the final version of both the positive peer evaluation comments for high achievers and the negative peer evaluation comments for low achievers was created.

3   Findings

Students in the cooperative learning class that utilized PBL used various criteria to evaluate their peers. When all the sentences in the peer evaluation essays were grouped by similar meanings, words, or concepts, 6 to 8 common factors emerged. Eight factors were observed when the criteria for the high achievers were examined. These are presented in Table 2. On the other hand, six factors were observed when the criteria for the low achievers were analyzed and these are displayed in Table 3.

Table 2. Positive peer evaluation comments for high achievers.

Factors	Comments
participation	* passionate participation * active participation
cognition	* idea suggestion/ idea developer * informative * expertise
preparation	* hard work * hard effort
communication	* good listener that provides timely and meaningful information or knowledge * reflecting peer opinion
leadership	* leadership (decisive/ persuasive)
responsibility	* responsible presentation
affectivity	* enthusiastic * deep interest * caring * trustworthiness
skills	* excellent skills in presentation * technical skills in making power point files

Table 3. Negative peer evaluation comments for low achievers.

Factors	Comments
participation	* passive
cognition	* providing no opinion
	* poor understanding
preparation	* attending without review of previous lesson
communication	* simply listening without responding
diligence	* absence
	* being late
	* not punctual
	* no preparation for collaborative work
cooperation	* selfish
	* not voluntary

The four factors, participation, cognition, preparation, and communication, were common evaluating criteria for both high and low achievers, even though the style of expression was opposite. For high achievers, the tone of students' expressions was positive, whereas the tone was negative for low achievers. Positive peer evaluation comments for high achievers read:

I was amazed to see the high achiever participate passionately in all discussion sessions. It is safe to say that specific information, expertise, and creative ideas were suggested by the high achiever. On the other hand, negative comments for low achievers included:

The low achiever missed a discussion with no assigned research results presented, and his part of the presentation seemed weakly prepared.

The low achiever was so silent in discussion sessions that we hardly felt that he was one of our team members.

The other four factors for high achievers were leadership, responsibility, affectivity, and skills. Examples of comments from students are:

As our team leader, he tried hard to organize a teamwork plan that ensured all members were as satisfied as possible. When we met, the high achiever continually set the next goal and tried not to go astray. The high achiever added new technological ideas and information to the presentation material, and finalized the presentation format.

The other factors for low achievers were diligence and cooperation. Examples of the students' comments are as follows:

The low achiever missed the first meeting.

Even when we asked the low achiever to be a supporting presenter for our final presentation, he did not reply at all to any of our group members.

Table 4. Positive and negative peer evaluation comments frequency table.

Positive peer evaluation comments			Negative peer evaluation comments		
Factors	Frequency	Percent	Factors	Frequency	Percent
participation	7	13.4	participation	7	33.3
cognition	17	32.7	cognition	2	9.5
preparation	2	3.8	preparation	2	9.5
communication	3	5.8	communication	2	9.5
leadership	9	17.3	diligence	6	28.6
responsibility	4	7.7	cooperation	2	9.5
affectivity	4	7.7			
skills	6	11.5			
Total	52	100		21	100

4   Discussion and Conclusion

This study investigates peer evaluation essays in a cooperative learning environment at a university to understand what individual students think are important when they evaluate peers. Discovering and comparing students' comments for high achievers and low achievers in peer evaluation are the main goals of this qualitative research. The common factors in the peer evaluation essays for both high achievers and low achievers are participation, cognition, preparation, and communication. For the high achievers, there are many complimentary comments such as passionate participation, idea developer, expertise, hard effort, good listener, or reflecting the opinions of their peers. On the other hand, for the low achievers, most of the comments from their fellow students are complaints about passive participation, no opinion suggestion, poor knowledge, no preparation for discussion, or half-hearted communication. These four factors were also emphasized by Felchikov (1986) and Duzer and McMartin (2000), Higgit (1996), and Segar and Dochi (2001). The peer group assessment questionnaire created by Felchikov (1986) included competence and creativity, which fall under cognition. Segar and Dochi (2001) also developed peer assessment tools including items related to contribution to group discussions and preparedness for involvement in group work. In addition, Higgitt (1996) reported that peer assessment criteria generated by the experimental group include participation, enthusiasm, originality (ideas and interest), analysis of information, ideas, contribution, and effort.

Peer evaluation comments about leadership, responsibility, affectivity, and skills appear only for high achievers. Those comments are related to persuasive leadership, responsive presentation, caring and trustworthiness, excellent presentation skill, or technical skills in using software. These factors were also selected by Duzer and McMartin (2000) as essential elements of peer and self-evaluation. They created a peer evaluation tool including consideration of others' ideas, taking a leadership role, and helping the group. In contrast, comments about diligence and cooperation appeared only in negative peer evaluation for low achievers. They contained words, phrases, or sentences about absence, not being punctual, selfish, involuntary participation, etc. Ohland et al (2005) constructed a peer evaluation form including items such as completing tasks on time.

The most frequently mentioned factors in the peer evaluation comments for high achievers are their intellectual ability, passionate leadership, and enthusiastic participation, while those for low achievers are passive participation and lack of diligence such as being late, absent, or unprepared.

The analysis of the peer evaluation essays written by students involved in group-based activities provides a lot of insight into students' attitudes toward group members. As shown in many studies across different ages, the emphasis in cooperative learning environments is on whether students participate in group activity with a passionate attitude or not. In addition, the study reminds students not to ignore the importance of leadership, intelligence, and diligence at cooperative work with classmates. All the peer evaluation criteria originate from the students, not from instructors, and no guiding direction was suggested in advance. Therefore, the identification of student's perspectives through peer evaluation here is raw and unfiltered. This fact makes this study meaningful.

In spite of the benefits of this study, it has several limitations. First, this study was conducted with a small number of participants, so extensive generalization of the results is not recommended in a learning environment. An unequal gender ratio for the participants also makes it difficult to represent a typical classroom environment. Second, because the major analytical method used in this study is based on qualitative research, which means that all the data analysis is influenced by the subjectivity of the involved researchers and reviewers, clear objectivity in the study process is not guaranteed. Jabeen (2013) reported that the weakness of the qualitative research method includes the inability to generalize, issues related to the use of certain techniques and instruments, data analysis issues such as multiple meanings of words and power dynamics among others. Finally, too much optimism about peer evaluation as a means of assigning grades to students should be restrained because doubts remain regarding the accuracy of peer assessment in general (Furman & Robinson, 2003). Students, in particular, do not trust the results of peer assessments (Furman & Robinson, 2003), so careful application of the results of peer evaluation by instructors is necessary.

References

1. Albanese, M. A., & Dast, L. (2014). Problem-based learn¬ing: Outcomes evidence from the health professions. Journal on Excellence in College Teaching, 25(3&4), 239-252.
2. Carter, N., Bryant-Lukosius, D., DiCenso, A., Blythe, J., & Neville, A.J. (2014). The use of triangulation in qualitative research. Oncology Nursing Forum, 41(5), 545-547.
3. Charmaz, K. (2006). Constructing grounded theory: A practical guide through qualitative analysis. London:Sage..
4. Curran, E., Carlson, K., & Celotta, D.T. (2013). Journal of the Scholarship of Teaching and Learning, 13(2), 49-71.
5. DeNisi, A.S., Randolph, W.A., & Blencoe, A.G. (1983). Potential problems with peer ratings. Academy of Management Journal, 26, 457-467.
6. Denzin, N.K. (1978). Sociological methods: A sourcebook. New York, NY: McGraw-Hill.
7. Doyle, S. (2007). Member checking with older women: A framework for negotiating meaning. Health Care for Women International, 28(10), 888-908.
8. Duzer, E.V., & McMartin, F. (2000). Methods to improve the validity and sensitivity of a self/peer assessment instrument. IEEE Transactions on Education, 43(2), 153-158.
9. Elder, A.D. (2009). Using a brief form of problem-based learning in a research methods class: Perspectives of instructor and students. Journal of University Teaching & Learning Practice. 12(1), 1-12.
10. Felchikov, N. (1986). Product comparisons and process benefits of collaborative peer group and self-assessments. Assessment and Evaluation in Higher Education, 11, 146-166.
11. Furman, B., & Robinson, W. (2003). Improving engineering report writing with Calibrated Peer Review. In D. Budny (Ed.), Proceedings of the 33rd Annual Frontiers in Education Conference. Piscataway, NJ: IEEE Digital Library.
12. Goldfinch, J. (1994). Further developments in peer assessment of group projects. Assessment and Evaluation in Higher Education, 19, 29-35.
13. Harvey, L. (2015). Beyond member-checking: a dialogic approach to the research interview. International Journal of Research & Method in Education, 38(1), 23-38.
14. Higgitt, D.L. (1996). The effectiveness of student-authored field trails as a means of enhancing geomorphological interpretation symposium. Journal of Geography in Higher Education, 20(1), 35-44.
15. Hmelo-Silver, C.E. (2004). Problem-based learning: What and how do students learn? Educational Psychology Review, 16(3), 235-266.
16. Jabeen, T. (2013). Qualitative research methods in children related research. PUTAJ-Humanities and Social Sciences, 20, 103-110.
17. Jeffery, D., Yankulov, K., Crerar, A., & Ritchie, K., (2016). How to achieve accurate peer assessment for high value written assignments in a senior undergraduate course. Assessment & Evaluation in Higher Education. 41(1), 127-140.
18. Kaufman, D.B., Felder, R.M., & Fuller, H. (2000). Accounting for individual effort in cooperative learning teams. Journal of Engineering Education, 89(2), 133-140.
19. Kramer, J.F. (1990). Perceived similarity and accuracy of peer ratings. Journal of Educational Psychology, 82, 213-218.
20. 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.
21. Ohland, M.W., Laton, R.A., Loughry, M.L., & Yuhasz, A.G. (2005). Effects of behavioral anchors on peer evaluation reliability. Journal of Engineering Education, 94(3), 319-326.
22. Patton, M. Q. (1999). Enhancing the quality and credibility of qualitative analysis. Health Science Research, 34, 1189-1208.
23. Segers, M., & Dochy, F. (2001). New assessment forms in problem-based learning: the value-added of the students' perspective. Studies in Higher Education, 26(3), 327-343.
24. Strachan, I.B., & Wilcox, S. (1996). Peer and self-assessment of group work: Developing an effective response to increased enrollment in a third-year course in microclimatology. Journal of Geography in Higher Education, 20, 343-353.