Using Peer Assessment in conjunction with Project-Based Learning: A Comparative Study

Gabriel Badea, Elvira Popescu
Computers and Information Technology Department
University of Craiova
Craiova, Romania
gbadea@software.ucv.ro, popescu_elvira@software.ucv.ro

***Abstract* – Peer assessment is a valuable approach in education, which can increase students' engagement and foster critical thinking and reflection. In this paper we investigate the potential of peer assessment for project-based learning settings, by means of a comparative case study. More specifically, we analyze and compare the outcomes and experiences of the students in two similar scenarios: one course where the peer assessment activity is compulsory and one course where it is optional. A dedicated peer assessment platform, called LearnEval, is used in both cases, which supports the teacher in configuring and monitoring the process and the students in reviewing peers' work and visualizing their activity and personal scores. We provide a description of the project settings and compare the unfolding and results of the two studies; we analyze the involvement and reviewing skills of the students, as well as their subjective appraisal of the learning experience. Our study adds to the body of literature on peer assessment, while the lessons learned can prove useful for the practitioners in devising and refining their instructional scenarios.**

Keywords – peer review; project-based learning; experience sharing; comparative case study

I. INTRODUCTION

Peer assessment (PA) has been applied in a variety of educational contexts and scenarios, due to its well-known benefits, especially in the last decade [6, 7, 10]. The literature reports on the use of the peer assessment approach as an effective component to complement and enhance the learning process, promoting highly desirable skills such as critical and analytical thinking or problem-solving abilities [8]. On the other hand, project-based learning (PBL) is a popular instructional method which also fosters problem solving and decision making skills, together with the application of knowledge in novel contexts [12]. Therefore, joint use of PBL and PA has the potential to enhance the learning experience, as reported in the literature [6, 9, 10].

Our aim in this paper is to further contribute to this body of knowledge, by providing a comparative case study between two different courses employing PBL in conjunction with PA. More specifically, a dedicated peer assessment platform is used for both courses, called LearnEval [1]. The system is versatile and highly configurable in terms of workflow, allowing the teacher to tailor the peer assessment process based on the context of the course [2]. As far as the students are concerned, the platform offers important affordances such as submission system, reviewing forms, automatic grade assignment, personal scores and reputation system, automatic notifications, open learner model and graphical visualizations [1].

In this context, we report on our experience of using peer assessment in two distinct PBL settings by: presenting the course background; offering an overview of the scenarios applied (section III); describing the unfolding and results of the studies; discussing the encountered differences; comparing the involvement, competence and reviewing capabilities of the students; investigating the quality and usefulness of the reviews provided by the learners (section IV); presenting students' opinions regarding the system usability as well as their subjective appraisal of the learning experience (section V). We end the paper with conclusions and lessons learned, outlining also some future research directions (section VI).

II. RELATED WORK

We start with an overview of related studies which report on the use of PA in conjunction with PBL.

A project management course for a computer science degree that combines three approaches: PBL, spiral learning (SL) and PA, is presented in [10]. The course requires students to develop four incremental projects whose complexity increases as learners gain new knowledge. PA does not affect the final grades and is performed based on Google Forms rubrics. The study analyzes the influence of PA on PBL and SL referring to the quality of the projects developed by the students. Furthermore, it explores whether there are significant differences in learning outcomes when applying such approaches depending on the type of learner. 154 students were involved in the study and the findings show that PBL combined with SL increased the quality of the projects, while PA raised it even further.

Peer assessment was also applied in conjunction with self-assessment (SA) in a higher education course on language teaching along five PBL modules [9]. 36 senior students were enrolled in the course. For each session, the student had to perform an SA and up to seven PAs. Positive correlations were found between SA and PA scores and the results show that students were consistent in assessing

2161-377X/20/$31.00 ©2020 IEEE
DOI 10.1109/ICALT49669.2020.00119

themselves and their peers. Overall, the authors conclude that the PA process was more accurate than SA.

Another report on the experience of PA in a Web Design project is presented in [8]. A total of 37 master and 26 bachelor students were involved in the study, having the task to develop a web application in the form of a personal blog. Students reviewed the programming activities of their peers, who were then allowed to revise their work based on the feedback received before submitting it for teacher evaluation. The findings show that peer review leads to an improvement in the project results and is deemed useful by the learners. Furthermore, strong correlations were found between the overall student rating based on peer reviews and the exam score, as well as between the teacher's evaluations and the students' reviews.

Somewhat different results are reported in [11], where SA, PA and tutor assessment are employed in a medical education course. The study compares the grades assigned by peers and tutors during tutorials involving 349 first year medical students, along 7 semesters. The findings show that the SA and PA grades were consistently higher than the grades assigned by the tutor. Another interesting discovery is that SA grades and PA grades mean increased slightly from the first until the last semester. The authors conclude that although SA and PA grades may be reliable, they are not valid for the PBL tutorial process.

Finally, a different approach is used in [5], where students are divided in two groups: one using PA and one not using PA. The context is an experimental physics PBL course aimed to enhance students' professional skills in LED (light-emitting diode) design, involving 73 junior students. The results show that the PA group achieved better results in concept clarification and enhancement of LED design skills in well-structured problem solving, while no significant difference between the two groups was recorded with respect to ill-structured problem solving. In addition, students from the PA group reported enhanced inquiry learning and reflective thinking skills.

To sum up, PA and PBL have been applied in various contexts over the time, ranging from computer science courses to medical education and language teaching, from studies employing dozens to several hundreds of students, and in conjunction with different other approaches, such as self-assessment or spiral learning. The current paper adds to the literature by exploring and comparing the outcomes and experiences of the students in two similar, yet distinct scenarios involving PA and PBL: one course where the PA activity is compulsory and one where the PA activity is optional. Furthermore, we believe this experience sharing can prove useful for the practitioners, helping to create enhanced instructional scenarios based on PA and PBL.

III. STUDY SETTINGS

In what follows we present the instructional scenarios of the two courses in which we applied PBL in conjunction with PA. The first study was in the context of a Multimedia Technologies in E-Learning course held during the first semester of 2018-2019 academic year and involved 41 undergraduate 4th year students following a Multimedia Systems Engineering program at the University of Craiova, Romania; we will call this Study A for the rest of the paper. The second study was in the context of a Web Applications Design course held during the second semester of 2018-2019 academic year and involved 109 undergraduate 3rd year students following a Computer Science program at the same university; we will call this Study B for the rest of the paper.

A PBL approach was used in both studies, with students learning by means of an individual project that they had to design and implement throughout the semester. In study A, students' task was to develop an interactive Informatics lesson for high school level, consisting in a website incorporating multimedia content, such as images, audio and video tutorials, animations, educational games, simulations and assessment tests; more details are available in [3]. In Study B, students had to develop a more complex web application on a topic of their choice, such as an online library, an auctions platform, a virtual store, an online management system, a platform for movie reviews etc.

The projects were organized on several milestones, at which students had to present intermediary deliverables. In Study A there were 3 such milestones: the first two were concerned with the structure, layout and style of the website and the creation of the basic educational material (involving mainly HTML5 and CSS3 code), while the third addressed the more complex and interactive educational resources (involving mainly JavaScript). In Study B the number of deliverables was higher, as more aspects had to be covered such as requirements engineering, database design or security integration. The first milestone required students to define the functionalities of the web application and model it using various UML diagrams. The second milestone was related to the data storing part of the application, where students had to define the database schema, the entities for the application and create the connection between the application and the database. For the third milestone, students had to design the architecture of the application and implement a subset of the functionalities. For the fourth milestone, students had to address the security of the application and integrate authentication and authorization functionalities. The last milestone required students to complete the implementation of the web application, integrating all the functionalities defined at the start of the project. From a technical point of view, students used mainly ASP.NET Core, together with HTML, CSS, Bootstrap, JavaScript & jQuery.

In both studies, the PBL approach was supplemented with a peer assessment component: for each milestone, the students had to review the deliverables developed by three peers. The main difference between them was that in Study A the PA task was mandatory and represented a part of the final grade, while in Study B it was optional and rewarded

with extra points in the final grade.

The peer assessment process was supported by the dedicated LearnEval platform, which provides various functionalities both for teachers and students. Thus, the instructor can configure the PA workflow to the specific course requirements, create and edit assignments, monitor learner activity, rely on the automatic score computing mechanism as well as the various graphical visualizations of course and student data [2]. Students can use LearnEval to upload their deliverables, submit reviews for peers' work, access reviews received and provide back-reviews, visualize detailed personal scores, statistics and charts [1].

For each milestone, one PA session was created in the platform; each student project deliverable was automatically allocated to three peers, with various reviewing skills. Students then had a timeframe for submitting the evaluations: two weeks in Study A and one week in Study B (as the number of milestones was higher). The instructor provided a set of assessment criteria for each milestone (3 for Study A, a variable number between 2 and 8 for Study B); students had to write a comment for each criterion in addition to assigning a summative grade (on a scale from 1 to 10). Once the review deadline was reached, each deliverable was automatically assigned a grade and a confidence factor based on the assessments received from the peers.

IV. Peer Assessment Activity and Results

In what follows, we present the unfolding of the two studies, comparing the PA activity exhibited by the students in the two courses.

As already mentioned, in Study B participation to peer assessment was not mandatory; nevertheless, a high number of students registered in LearnEval (89 out of 109, or 82%[1]), showing an initial interest in the PA activity. Subsequently, only a part of these students actually participated in the peer assessment process, as summarized in Table I. A total of 203 deliverables were submitted for review, with a higher number in the first two sessions and a lower number in the last three sessions. Students also provided more reviews in the first two sessions compared with the last three, with a total of 315 assessments submitted. Potential explanations for the lower number of deliverables and reviews consists in the higher complexity of those milestones, the nearby spring holidays (for the third session), and the nearby final exams (for the fifth session), when students chose to dedicate more time for exam preparation rather than for an optional activity.

Interestingly, we can notice that students followed a different trend in the two courses: while the number of project deliverables submitted remained relatively constant for the three sessions in Study A (with a slightly lower number in the second session), in Study B this number decreased from one session to the other (with the highest value in the first session and the lowest in the last session). A similar trend can be noticed for the number of reviews provided by the students. The rationale behind this behavior could be that students lose their interest over time when an activity is not mandatory and tend to give up when the tasks become more difficult.

TABLE I. Numbers and percentages of deliverables and reviews submitted in each study

PA session	Deliverables		Reviews	
	Study A	***Study B***	***Study A***	***Study B***
Session I	33 (80%)	61 (56%)	86 (87%)	112 (61%)
Session II	27 (66%)	50 (46%)	52 (64%)	94 (63%)
Session III	33 (80%)	35 (32%)	80 (81%)	34 (32%)
Session IV	N/A	35 (32%)	N/A	53 (50%)
Session V	N/A	22 (20%)	N/A	22 (33%)

LearnEval platform also computes an involvement score for each student, based on the number of deliverables, reviews and back-reviews submitted [1]. A comparison of the students' involvement scores between the two studies can be seen in Table II (only students who registered in LearnEval are included). As can be seen, learner involvement was significantly higher in Study A than Study B; furthermore, a relatively large number of students in Study B did not submit any deliverables or reviews (18 out of 89 who registered in LearnEval and 38 out of 109 overall).

TABLE II. Student involvement scores overview

Study	Involvement Score		
	>= 8 (high involvement)	***>= 5 (medium or high involvement)***	***= 1 (no involvement)***
Study A	23 (56%)	34 (83%)	3 (7%)
Study B	16 (18%)	43 (48%)	18 (20%)

As far as the quality of the submitted project deliverables is concerned, the average scores per session are included in Table III. The values are quite high and similar for the two studies. It is worth noting that for Study B, while the number of deliverables decreased from one session to the other, their quality generally increased (with the exception of session IV, which was perceived as more difficult by the students). This can be explained by the fact that learners with lower performance gradually lost their interest and decided to skip the last PA sessions, which were only attended by the highly motivated students. By contrast, in Study A most students participated in all the sessions, regardless of their performance, as the PA activity was mandatory.

An additional score is computed, which takes into account also the missing deliverables, i.e., for each student, an average overall deliverable score represents the sum of the marks assigned to the submitted deliverables divided by

[1] Percentages are rounded to the nearest integer throughout the paper

the total number of milestones (missing deliverables being graded with 1). A comparison between the distribution of scores in the two studies is presented in Table IV (again, only students who registered in LearnEval are included). While in Study A the majority of the students obtained passing scores (i.e., higher than 5), in Study B many students obtained the lowest score (i.e., 1), as they did not submit any deliverables.

TABLE III. AVERAGE PROJECT DELIVERABLE SCORES PER SESSION

Study	Session I	Session II	Session III	Session IV	Session V
Study A	8.49	8.66	8.06	N/A	N/A
Study B	8.14	8.41	8.55	7.67	8.8

TABLE IV. DISTRIBUTION OF OVERALL DELIVERABLE SCORES

Study	Average overall deliverable score		
	>= 8	*>= 5*	*= 1*
Study A	13 (32%)	30 (73%)	3 (7%)
Study B	17 (19%)	44 (49%)	21 (24%)

Finally, we take a look at the content of the textual feedback provided by the students; for each assessment criterion defined by the teacher, the learners had to write a comment to justify their score. The number of comments and their distribution according to length is included in Table V.

TABLE V. NUMBERS AND PERCENTAGES OF STUDENT COMMENTS ACCORDING TO LENGTH

Study	Comments	Number of characters			
		>= 50	*>= 100*	*>= 200*	*>= 400*
Study A	654	180 (28%)	68 (10%)	28 (4%)	11 (2%)
Study B	1600	433 (27%)	157 (10%)	50 (3%)	9 (1%)

As can be seen, the majority of the comments are very short (less than 50 characters); the percentages of longer comments are very similar for the two studies. This shows students' tendency to provide brief and sometimes not very helpful feedback to their peers. Examples of such short comments which do not provide sufficient insight to the learner include: *"it is a complete diagram", "the schema is too simple", "too poor", "seems ok", "could be extended more", "very well explained", "acceptable, but not complete"*. On the other hand, there are also helpful and detailed comments, such as: *"The class diagram seems incomplete considering the application is supposed to allow the customer to either order coffee online or to actually walk into the coffee shop to get it. The provided class diagram seems to have the classes necessary only for the customer walking to the shop. Either more classes need to be added or more methods for each class should be implemented"*. This kind of comments can actually draw students' attention on the weaknesses of their work and help them improve their projects for the next milestones. Hence, it is essential to integrate approaches for motivating students to provide detailed and relevant feedback to their peers.

V. LEARNER SATISFACTION

At the end of both studies, the students were asked to fill in two opinion surveys: one regarding the usability of the LearnEval system and one regarding their perceived satisfaction with the learning experience. The results are reported and compared in the next subsections.

A. Usability of LearnEval System

The System Usability Scale (SUS) [4] was applied in order to gauge students' opinion on the usability of LearnEval. 38 students from Study A filled in the questionnaire (i.e., 93%), compared to 48 students from Study B (i.e., 54% of the total number of students who registered in LearnEval, or 68% of those students who had at least some activity in the system). The results are included in Table VI.

TABLE VI. RESULTS OF SUS QUESTIONNAIRE

SUS item [4]	Study A	Study B
1. I think that I would like to use this system frequently *(relative and strong agreement)*	42%	69%
2. I found the system unnecessarily complex *(relative and strong disagreement)*	40%	61%
3. I thought the system was easy to use *(relative and strong agreement)*	74%	71%
4. I think that I would need the support of a technical person to be able to use this system *(relative and strong disagreement)*	63%	79%
5. I found the various functions in this system were well integrated *(relative and strong agreement)*	68%	83%
6. I thought there was too much inconsistency in this system *(relative and strong disagreement)*	34%	69%
7. I would imagine that most people would learn to use this system very quickly *(relative and strong agreement)*	63%	69%
8. I found the system very cumbersome to use *(relative and strong disagreement)*	53%	79%
9. I felt very confident using the system *(relative and strong agreement)*	42%	69%
10. I needed to learn a lot of things before I could get going with this system *(relative and strong disagreement)*	71%	85%

Students in both studies agree that the system was easy to use (item 3) and that most people would learn to use it very quickly (item 7). However, for all other items, the results are significantly better in Study B; most students were confident using the system and would like to use it frequently; they did not find it too complex, cumbersome, requiring additional knowledge or the support of a technical person; they also believed that functionalities were well integrated in LearnEval, without too much inconsistency.

B. *Learning Experience and PA Satisfaction*

An additional questionnaire was applied to the students in order to evaluate their learning experience and satisfaction with the PA process. This questionnaire was filled in by 38 students from Study A (i.e., 93%) and 50 students from Study B (i.e., 56% of the total number of students who registered in LearnEval, or 70% of those students who had at least some activity in the system). In what follows we discuss the results for some of the most relevant survey items.

The vast majority of students in Study B (98%) believed it was a good or very good idea to evaluate peers' projects, compared to 58% in Study A. Similarly, 96% of students in Study B thought that being evaluated by their peers is a good or very good idea, compared to 58% in Study A. A high percentage of learners in Study B considered helpful or very helpful to provide assessment and feedback to their peers (86%), as well as to receive assessment and feedback from peers (82%). By contrast, only 34% of students in Study A agreed with these statements.

As far as the PA settings are concerned, most students in both studies were satisfied with the time allotted for submitting solutions (71% in Study A vs. 76% in Study B), the time allotted for providing reviews (71% in Study A vs. 82% in Study B), and the number of assessment criteria (87% in Study A vs. 92% in Study B). Most of the students agreed that the PA process should be double-blind (82% in Study A vs. 86% in Study B).

Regarding the quality of the reviews received, perception was more positive in Study B: the percentages of students who appraised the evaluations to a large or very large degree were as follows: i) objective and unbiased - 56% in Study B vs. 34% in Study A; ii) complete and detailed - 46% in Study B vs. 21% in Study A; iii) useful for improving the project - 62% in Study B vs. 34% in Study A.

Furthermore, the PA process was found motivating to a large or very large extent by 56% of the students in Study B vs. 40% in Study A. At the same time, it was considered stressful or very stressful by only 8% of the students in Study B and 26% of the students in Study A; and time-consuming by only 6% of the students in Study B and 24% of the students in Study A. Finally, the majority of the students were satisfied or very satisfied with the LearnEval platform: 84% in Study B and 58% in Study A.

Overall, we could notice that results are more positive in Study B, with students being more attracted to the PA process when it is not a mandatory activity.

VI. CONCLUSION

The paper explored the use of peer assessment in project-based learning settings, by means of a comparative case study. Two similar instructional scenarios are described, one in which the PA activity is compulsory (Study A) and one in which the PA activity is optional (Study B). As expected, the level of learner engagement is higher in Study A, with more students providing reviews for peers' work. However, learner satisfaction is increased in Study B, when students can choose whether to participate in the PA process or not.

A drawback in both scenarios consisted in students' tendency to provide short and sometimes superficial feedback to their peers. Hence, efforts should be made by the instructors to increase the quality and level of detail of the textual part of student assessments. Providing examples of high quality reviews and/or templates to be followed, setting a minimum required length for the text, as well as focusing on the review quality when awarding grades are some potential recommendations. Students themselves can foster the provision of adequate feedback by using the back-review mechanism to signal superficial or useless reviews.

Overall, students' experience with LearnEval platform was a positive one and we plan to extend its use to various courses and pedagogical scenarios. We also aim to further explore the potential of peer assessment for project-based learning and devise mechanisms to increase students' engagement and motivation with the process.

REFERENCES

[1] G. Badea and E. Popescu, "A Web-Based Platform for Peer Assessment in Technology Enhanced Learning: Student Module Prototype", Proc. ICALT 2019, IEEE CS Press, pp. 372-374, 2019.

[2] G. Badea and E. Popescu, "Instructor Support Module in a Web-Based Peer Assessment Platform", Proc. ICSTCC 2019, IEEE, pp. 691-696, 2019.

[3] G. Badea and E. Popescu, "Using LearnEval Peer Assessment Platform in Project-Based Learning Settings: A First Experience Report", Proc. ITHET 2019, IEEE, 2019.

[4] J. Brooke, "SUS: A Retrospective", Journal of Usability Studies, vol. 8(2), pp. 29-40, 2013.

[5] S.-H. Chang, T.-C. Wu, Y.-K. Kuo, and L.-C. You, "Project-Based Learning with an Online Peer Assessment System in a Photonics Instruction for Enhancing LED Design Skills", Turkish Online Journal of Educational Technology, vol. 11(4), pp. 236-246, 2012.

[6] Y.-H. Ching and Y.-C. Hsu, "Peer Feedback to Facilitate Project-Based Learning in an Online Environment", The International Review of Research in Open and Distance Learning, vol. 14(5), pp. 258-276, 2013.

[7] M. De Marsico, F. Sciarrone, A. Sterbini, and M. Temperini, "Supporting Mediated Peer-Evaluation to Grade Answers to Open-Ended Questions", EURASIA Journal of Mathematics, Science and Technology Education, vol. 13(4), pp. 1085-1106, 2017.

[8] V. Dropcova, M. Homola, and Z. Kubincova, "May I Peer-Review Your Web Design Project?", Proc. ICSLE 2015, Springer, pp. 243-249, 2015.

[9] T. Erdogan, I. Yurdabakan, and N. Senemoglu, "Self- and Peer Assessments in PBL: A Higher Education Example", Proc. IAC-GETL 2018, pp. 53-61, 2018.

[10] A. Jaime, J.M. Blanco, C. Domínguez, A. Sánchez, J. Heras, and I. Usandizaga, "Spiral and Project-Based Learning with Peer Assessment in a Computer Science Project Management Course", Journal of Science Education and Technology, vol. 25(3), pp. 439–449, 2016.

[11] J.L.M. Machado, V.M.P. Machado, W. Grec, V.R. Bollela, and J.E. Vieira, "Self- and Peer Assessment May Not Be an Accurate Measure of PBL Tutorial Process", BMC Medical Education, vol. 8, article 55, 2008.

[12] J.W. Thomas, "A Review of Research on Project-Based Learning", bie.org/research/study/review_of_project_based_learning_2000, 2000.