SHIFTING RESPONSIBILITIES: USING PEER ASSESSMENT IN SENIOR, ENGINEERING DESIGN TO PROVIDE EFFECTIVE SUPPORT AND MEANINGFUL FEEDBACK DESPITE LARGE CLASS SIZE

Andrea Bradford and Julie Yale
School of Engineering, University of Guelph
abradfor@uoguelph.ca

Abstract – Urban Water Systems Design is a required, senior engineering design course for undergraduate students in the Water Resources and Environmental Engineering programs at the University of Guelph. A central component to this course has been a challenging, stormwater management design and simulation project. Recently, enrollment has increased from approximately 50 to 90 students. With these increased numbers, it is no longer feasible for the instructor to provide individualized, rich and robust feedback on the project. Rather than eliminating or simplifying the project, which is a highly valuable learning activity, peer assessment was investigated as an option.

In Fall 2016, peer assessment was implemented for two term tests and a design project. In addition to addressing resource constraints, this shift in responsibilities takes advantage of the active, collaborative, learning opportunities provided by grading tests in class and giving feedback to and receiving feedback from others on design work and report writing. To achieve the largest benefit possible, best practices suggested in the literature were followed, such as training peer assessors, including developing a rubric with the class to enhance understanding of expectations; using multiple assessors to address student concerns about fairness; and incorporating reflection on the peer assessment activities.

Data were collected through two surveys administered before and after the course’s peer evaluation activities and through graded reflections on peer evaluation activities. Most students thought they learned as much or more than they would have without the incorporation of peer evaluation. Based on predominantly positive student comments and fair grading, peer grading of tests will be used in future with a few minor modifications. Most students also found the peer assessment of design reports to be fair, with a reasonable time commitment. Some students were troubled by the variability of grades given by peers. Enhanced training was suggested to help students grade more consistently and provide more effective feedback.

Keywords: Peer assessment, Collaborative learning, Engineering design.

1. INTRODUCTION

Urban Water Systems Design is a required, senior engineering design course for undergraduate students in the Water Resources and Environmental Engineering programs at the University of Guelph. With increased enrollment in the environmental engineering program, class size has increased to between 80 and 100 students.

The course is delivered with two 80 minute lectures and one 110 minute “tutorial” each week. In the tutorial, students learn to use simulation tools and work on a design project. In the most recent offering, with approximately 90 students, there were two tutorial sections and one teaching assistant (TA). Both the TA and instructor are present during the tutorials. The TA’s remaining hours are usually devoted to helping students with modelling challenges and marking occasional test questions. Since TAs with design experience in urban water systems are not consistently available, the instructor historically did most of the test grading and all of the design project grading.

A challenging stormwater management design and simulation project is a central component to the course. With the increased class size, it is not feasible for the instructor to provide individualized, rich and robust feedback on the project.

The course was split into two offerings to mitigate the large class size, but resource constraints no longer allow for this split. Another mitigation strategy was the use of design teams, but the use of teams shifted the learning outcomes. In particular, this course’s individual project is seen to have an important place in the school’s curriculum since students have few remaining
opportunities for individual design and report writing within their undergraduate programs.

The learning value of the project was indisputable, so peer assessment was investigated as an option to allow the preservation of the project and associated desirable outcomes. While instructor feedback on individual work is reduced, peer assessment provides students with timely feedback and allows the instructor to focus on supporting students, monitoring learning, and providing substantive feedback to the class as a whole. Although motivated by resource constraints, it was recognized that peer assessment offered the opportunity to enhance learning.

Providing sufficient feedback, both often enough and in enough detail, and timely feedback are important to student learning [2]. A contrasting view is that there is actually little evidence that effort and resources put into enhancing teacher feedback actually results in improved student learning [3]. Approaches that offer frequency and speed, such as peer assessment, may compensate for poorer quality or less individualized feedback [2].

Wilson, Diao and Huang [4] question whether it is “justifiable and appropriate to require undergraduate students to mark other students’ assessments” and whether “students are equipped with the necessary competencies and motivation” and consider student concerns about the “inherent procedural unfairness of peer-to-peer review” to be an underexplored issue. Van Zundert, Suijsmans, and van Merrienboer [5] examined four categories of outcome variables: psychometric qualities of peer assessment (related to validity and reliability), domain-specific skills (e.g. quality of writing), peer assessment skills (related to type and quality of student feedback) and student attitudes towards peer assessment (e.g. confidence in assessing peers, perceived learning benefits) and identified the important influence of peer assessor training.

Generating comments about the strengths and weaknesses of the work of peers promotes deep learning and self-evaluation as students simultaneously consider the merits of their own work [3]. However, further research is needed to develop best practices for implementation such that the full learning potential of peer assessment is achieved, while addressing student concerns about fairness, workload and the quality of feedback from peers.

The research project investigates which of the potential benefits of peer assessment were realized in this offering and gauges student receptiveness to the approach (when effort is made to communicate the pedagogical rationale), with the goal of refining some of the best practices for the application of peer assessment in engineering design courses. This paper presents preliminary observations and results. The originality lies in the application of peer assessment in a design-intensive course and to individual work, rather than team work.

2. METHODOLOGY

2.1. Implementation of Peer Assessment

In Fall 2016, peer assessment was implemented for two term tests, each worth 20% of the course grade, and three design project deliverables. The first test, which was simulation-based, had two versions completed in tutorial sections on different days. Both versions were simultaneously graded by the class during a single 80 minute lecture period the following week. Acceptable answers were projected on a screen at the front of the class and the instructor responded to queries from markers for each test question. Peer grading of the second test was similar, but the grading was completed in the first lecture period following the test, there was only one version of the test, and it was slightly shorter than the first test. For both tests, one question was given after the peer grading and it was necessary to participate in the activity to earn marks on the question. This helped to ensure that all students were in attendance for the peer grading of the tests.

The design project deliverables included two short reports and a final presentation. Peer evaluation of the two design reports was facilitated by the PEAR (Peer Evaluation, Assessment and Review) Tool. Five peers used a four item rubric to evaluate each of the preliminary reports (worth 5% of the course grade) and submitted review comments via the PEAR tool.

For the second submission (worth 20% of the course grade), the class developed a more comprehensive rubric and four peers evaluated and provided feedback on each report. Students earned up to 2% and 5% for grading their peers first and second reports, respectively.

The final phase of the project was done in groups, with oral presentations of the design and simulation results. The instructor evaluated presentations and provided feedback, although peers also provided feedback on the presentations for three other groups as well as their own group members. Students earned up to 5% of the course mark for this peer grading activity.

2.2. Data Collection

Data were collected through two surveys administered before and after the course’s peer evaluation activities and through graded reflections on peer evaluation activities.

The first 20 questions on each of the surveys comprised the revised, two-factor, study process questionnaire (R-SPQ-2F) to assess deep and surface approaches to learning [1]. Since students’ learning strategies can adjust to the particular context and course
being taught, the R-SPQ-2F questions were included in both surveys. The first survey also included 5 questions related to background and preparedness of students and 12 questions related to perceptions of peer assessment (e.g. fairness, workload). The second survey included 16 questions related to the peer assessment experience. It also asked students to rank learning activities in terms of their value. Most of the questions were asked on a Likert style scale.

The first graded reflection was given as a take home question after in-class grading of Test 1. Students were asked to provide a one paragraph, thoughtful, description of their views on the peer grading of Test 1 as a whole class learning activity. Students were invited to provide constructive criticism but since the submissions were not anonymous, they were advised that they would not be penalized if they chose not to do so. For the second reflection, students were asked to provide a description of their views on the peer evaluation of the design project.

Participation in the surveys was optional. Reflections were mandatory. For those students that consented, the study design allowed student perceptions and experiences to be linked to learning strategy and performance on course assessments.

3. RESULTS AND DISCUSSION

Initial analysis of the data indicates that students had some reservations, but showed an understanding and appreciation of the pedagogical rationale for peer assessment. There were 50 and 36 respondents on the pre-activity and post-activity surveys, respectively. Most students assigned high value to the design project as a learning activity, so efforts to maintain this course component are justified. Consistent with the findings of other studies, most students thought they learned as much (53% of respondents), or more (12% of respondents), than they would have without the incorporation of peer assessment. There were 50 and 36 respondents on the pre-activity and post-activity surveys, respectively. Most students assigned high value to the design project as a learning activity, so efforts to maintain this course component are justified. Consistent with the findings of other studies, most students thought they learned as much (53% of respondents), or more (12% of respondents), than they would have without the incorporation of peer evaluation. Analysis of the survey data is ongoing.

3.1 Tests

From the instructor’s perspective, the in-class grading of Test 1 (based on simulation of a water distribution system) went well. Students did have to be asked to quiet down a number of times. Due to lack of time, one of the most “open-ended” questions was graded by the instructor. This gave the instructor an opportunity to check the peer grading. Only very occasional errors were noted.

The students had a variety of perspectives. Some students were frustrated by the lack of time available to ask clarifying questions during the marking while others were frustrated with the number of questions asked and the lack of judgment exhibited by some peers. The following quotes illustrate this range of views:

“In a way I felt like it was rushed and some of the questions being asked by other peers were not heard or shared to the rest of the class.”

“The unfortunate downside to this method of learning/teaching is that a number of people do not use their own personal judgement when marking the tests at times where the answer was not fully correct. This resulted in a lot of wasted time clearing up questions that should be able to be answered on their own.”

Some students found that they were preoccupied with trying to remember how they had responded on their own test. A few students were uncomfortable when one of their responses was shared and deemed to be wrong, even though the process was anonymous. A common suggestion was to include more acceptable responses, as well as incorrect responses, on the projected test.

In-class grading of Test 2 (design of a section of a sanitary sewer system) went very smoothly. A handful of students misinterpreted the flow directions within the sewer system. The instructor pulled these tests out ahead and graded the affected questions. Two students who had clearly not learned the design process were identified. The instructor graded these tests in advance and the students were advised that they would be given their own tests to follow during the in-class grading. A design table with the correct values was projected on a screen at the front of the room. Although students were required to check calculations (projected on a second screen), the summary table made any problem areas obvious. The last 20 minutes of the class were used to re-test a portion of the design process and the instructor followed up with any students that still did not fully understand the design process.

At the end of the course most students commented that they were comfortable with the peer grading of the tests. Many students found that it helped them to learn the theory and design processes better. Many found it valuable to hear other correct responses and the reasons that some responses were incorrect. The following quotes provide a sample of the student comments:

“It allowed me to be able to understand what a good answer was, and how a poorly articulated answer could have been improved.”

“I think reading the answers of others and being able to determine whether they were correct or on the right track makes for a much better learning exercise. It is one thing to know the answer but
another to be able to explain why an answer is correct. This is something I found myself doing as I read the answers of my peers and it was definitely beneficial for my own learning.”

“It was very interactive and engaging which made for a much more enjoyable and effective learning environment than sitting through a lecture. I feel as though my grasp on several water distribution concepts was solidified during this exercise.”

Despite these positive comments, some students identified the “loss” of a lecture as a drawback. From the instructor’s perspective, sacrificing a traditional lecture to engage different learners and to reinforce lecture content was worthwhile. Instructor and teaching assistant time can also be reallocated to other course activities.

Based on predominantly positive student comments and fair grading (as judged by the instructor), peer grading of tests will be used in future offerings of the course. One of the refinements will be to limit the clarifications from the class on each question; students that remain uncertain will instead be encouraged to flag the response for review by the instructor.

3.2 Design Project

Compared to the peer evaluation of tests, there was a broader range of perspectives on the peer evaluation of design reports. Even though there was more support for peer assessment at the end of the course, and even with the explanation that simplification or elimination of the design project is a more likely option than instructor grading of individual reports, several students maintained that grading of the reports should be done by the instructor.

With respect to workload, students found peer assessment took more time than they expected (Fig. 1). When asked how the peer assessment process could be made more time-efficient, some students indicated that the time required was reasonable:

“I think it is fair based on the grades which are allocated to each peer assessment activity.”

“I feel it is not possible for the process to be more time-efficient. Students need time to properly assess others especially because they are still learning the course themselves.”

“I don’t think reading reports should be made “time-efficient”. The peer assessment took a long time, but we were rewarded with marks for doing so. I think the marks allocated for it were more than enough for the time it took.”

![Fig. 1. Workload expectation and experience](image1.png)

With respect to fairness, a slightly larger proportion of respondents found their work was graded unfairly than expected this to be the case (Fig. 2). For those students that did not find that their work was marked fairly, the survey probed for more information.

![Fig. 2. Fairness expectation and experience](image2.png)

Some students recognized that the average grade given on their work was about right, even when individual evaluations may have been inappropriate. Other students seemed more troubled by a large variation in marks. In some cases, students felt their own grade was appropriate, but still found it unfair that others may have received higher grades for similar work.

“I didn’t agree with each evaluation but I agree that my overall grade and feedback was fair (balanced out). “

“My work had its share of mistakes, and was marked on a semi-harsh to fair range in my opinion. The problem lies in others making the same
mistakes or WORSE, and receiving a much higher mark because they had easy markers.”

“Someone gave me perfect on the last report. I shouldn’t have gotten that. It speaks of many student’s desire to improve grades by encouraging everyone to mark nice. I don’t agree with that. On that same deliverable, another student docked tonnes of marks for ludicrous reasons. In the end, it all evened out and I feel that I got the mark I deserved. Creating a structure where the grade is an average of those provided by students is very effective and should result in a fair grade. Individually some students were horrible though.”

Students had the most reservations about peer grading of the project report which was worth 20% of the course mark. In particular, strong students questioned how weaker students could adequately assess their work. In some cases, students hesitated to provide feedback that might not be correct.

Survey respondents found that feedback was provided in a timely manner and all but four respondents found the feedback on writing and technical content to be neutral or somewhat helpful. Lack of feedback, or inappropriate feedback, on design reports were commonly cited concerns.

All but one respondent found that the PEAR tool was effective for the design project and found that adequate time was provided for the assessments. The process was intended to be anonymous; however, after submission of the first reports, it was discovered that the majority of the class did not know that it was necessary to remove personal information before submitting their work. Furthermore, after grading of the second reports, it was discovered that some students would have gone back and adjusted the grades they gave but did not know that it was possible to do so. Anonymity seemed to be important to the process (Fig. 3).

Students were asked why they would have given different marks and feedback if the process had not been anonymous. They had a range of perspectives, as illustrated by the following comments. However, it is clear that many students do not want a friend or classmate to be upset with them because of a poor grade or critical comment. Some indicated that they might have been more careful about comments if it has not been anonymous.

“I wouldn’t feel comfortable giving grades to students who could identify me. I would be less confident and would find it challenging to balance social interactions and fair grading.”

“Although I prefer to be a nice guy, if someone submits a poor report I will call them out on it via the grade and constructive comments I give. If they resent that it is their problem not mine. The subpar grade would be their fault, with me merely the messenger. Sure some may consider this harsh but I can justify my actions as the alternative undermines our education system.”

“Pretty much everyone in class knows each other, so I think it is better to keep it unbiased by making it anonymous.”

Critical feedback and low grades were given by a few reviewers, but in almost all cases, the instructor found these to be justified. The instructor did find some reviews to be insufficiently critical, which may give students a misleading impression of their level of achievement and knowledge. This was addressed in feedback from the instructor to the class as a whole.

Students provided valuable suggestions for improving the peer assessment process. Many thought that enhanced training would help students grade more consistently and provide more effective feedback, even though a rubric for the 20% project report was constructed with the class. Involving peer assessors in the construction of marking rubrics has been recommended [4] and arguably, senior engineering students should be able to identify elements of good designs and reports. However, neither instructor nor students found the collaborative development of the rubric to be an effective use of a lecture period. Students want expectations to be set out early in the process but at that point, they have not thought about the assignment enough to contribute to setting out the expectations.

A number of students felt that it would be an appropriate compromise for the instructor to set out the expectations (as appropriate for a senior engineering course) and then to use a lecture period, after students had submitted their reports, to discuss weak, average, and strong exemplars. Students thought that this training would better help them to grade consistently and provide constructive feedback. Although the training would not
help them to understand expectations in time to improve their own submissions, students at this level recognize that providing exemplars in advance that can be “copied” does not encourage the desired behavior. Still, the training and reviewing of peers’ work provide valuable learning opportunities.

4. CONCLUSIONS

Students confirmed the value of the design project so it is worthwhile to investigate strategies that would allow this project and the associated desirable outcomes to be maintained. Most students thought they learned as much or more than they would have without the incorporation of peer evaluation. Thus, efforts to refine this approach to achieve its full learning potential while addressing student concerns about fairness, workload and the quality of feedback from peers are also worthwhile.

Peer grading of tests will be used in future with a few minor modifications. Some students were concerned about the “loss” of a lecture period; however, far more students identified valuable outcomes associated with this activity. Most students also found that reviewing and providing feedback on the design reports of peers was a valuable learning experience. Comments suggested that the workload and the grades and time allocated for the assessment activities were reasonable. The use of multiple assessors improved the fairness of the process although some students were troubled by the variability of marks even if the average grade they received was appropriate. Students suggested enhanced training to help them grade more consistently and provide more effective feedback.

References


