# Utilizing Peer-review Assessments of Submissions in Large Classes

Jan S. Sneep[1], Angelica J. Campigotto[1]
[1]*Queen's University, Kingston, ON*

Multiple peer-review approaches have been utilized to conduct peer-review assessments in a Capstone design course in the Mechanical and Materials Engineering program, at Queen's University, Kingston, ON. The need to improve the timeliness of feedback provided to students on their Individual Draft of the Final Design Report assignment was addressed using a peer-review process. Validation of the process was done using a low-stakes Motivation assignment using the Aropä online system. Each time the process was used the majority of the class were able to demonstrate a reasonable level of effort in providing both quality and quantity of feedback, as well as subjective assessments of their peer's work using a rubric.

Keywords: Peer-Review, Aropä, Feedback Fruits, Assessment

## Background

Capstone Design project teams, in the Mechanical and Materials Engineering department of Queen's University, Kingston, ON are usually composed of five, fourth-year students assigned to work collaboratively on an open-ended design challenge. Historically at the end of the tenth week of the twelve-week course, a mandatory "Individual Draft of the Final Design Report" assignment has been included in the Capstone Design course[i], referred to as the "Individual Draft" assignment. It requires each student on the team to write-up their portion of the Final Design report, such that it would be possible, with appropriate cutting and pasting, to create a first draft of the team's Final Design report. The Final Design report is a team submission. It is up to the team to decide which individuals will write each portion of the report. The two primary objectives of the Individual Draft assignment are to help reduce the anxiety of the team members, because every member of the team is expected to write something and also to provide an additional individual writing assessment within the teamwork-based environment.

At 36% of the course grade the Final Design report is the most heavily weighted single deliverable. With such a high-stakes deliverable it is not uncommon for students to feel some anxiety, not only about their own portion of the report, but also the work of their teammates. By forcing the entire class to work on a low-stakes draft of the Final Design report, worth 2% of the course grade, two weeks before the end of the course, anxiety may be ameliorated. It is not expected that all content will be fully completed by the end of week 10, however, the majority of the work required for the project should be at a stage where writing can begin.

A secondary objective of the Individual Drafts assignment is to provide individual feedback about their writing and to offer specific suggestions for improving their submissions. Accomplishing the second objective has proven to be challenging with only one Teaching Assistant (TA) and an enrolment that has averaged 150 +/– 26 students over the last five years. Typically, the TA and the course Instructor required 14 – 16 days to read the submissions and provide feedback. This gave students only three or four days to incorporate feedback into their Final Design report submission, which is typically due the Wednesday in the week following the end of classes in week 12.

In 2017, an experiment was conducted to utilize a peer-review process to see if the timeliness of the feedback could be improved and if the assessment workload for both the TA and the Instructor could be improved. The peer-review process is a beneficial way for students to consider the quality of a peer's work and provide assessment with both qualitative and quantitative feedback on the strengths, weaknesses and considerations of another student's work[1],[2].

During one of the two-hour long class tutorial time slots in Week 11 paper copies of the Individual Draft submissions were randomly distributed to the students, along with a paper copy of a simple rubric. When handing out the submissions for review no one received a fellow teammate's submission, unless the team had a Non-disclosure Agreement (NDA) in place, in which case teammates assessed each other's submissions. The rubric had three rows for providing assessments of the written communication, "Concision", "Logical Progression of Thoughts", and the overall "Spelling, Grammar, and Formatting" of the submission, and four levels of assessment, "Marginal", "Developing", "High Quality", and "Mastery", corresponding to a "D", "C", "B", and

"A" in letter grades respectively. The rubric also had space for the assessor to add hand-written comments. No instructions, beyond the descriptors within the rubric, were provided to the assessors to identify what level the submission was.

Students were instructed to focus their assessments, comments, and feedback on just the communication elements of the submission and ignore the technical elements.

Our primary question was on the students' ability to take the activity seriously and provide reasonable quality assessments and feedback. The students were given 30 to 45 minutes to read a submission, mark-up the submission with suggested edits, and fill out the rubric.

Once the first round of assessments was completed the paper submissions were redistributed throughout the class and a second copy of the rubric was handed out for a second round of assessments. It is possible that mark-up, with suggested edits, from the first assessor influenced the second, however the second assessor did not have access to the completed rubric from the first assessor. The TA and the instructor took the following day to go through the rubrics and provide a simple subjective four-level assessment of the effort and detail in the feedback provided.

- Level 0 – No comments, just rubric assessment
- Level 1 – Minimal effort and detail
- Level 2 – Reasonable quality of effort and detail
- Level 3 – Exceeded expectations

The paper copies of both the marked-up submissions and the rubrics were returned to the class within a week of the original submission due date.

Results were very promising with 84% of the feedback being at or above Level 2 (45.8% at Level 2 and 38.2% at Level 3). The process was not perfect, and a number of deficiencies were identified. For example:

- Some students arrived late for the session.
- Some students couldn't stay for the entire session.
- Some students completed their reviews more quickly than others.

All of which resulted in a less than ideal redistribution for the second round, which resulted in some submissions only being assessed once and others three times. It was also subsequently determined, through conversations with students, that the rubric used was too simple and did not provide enough levels of assessment.

These deficiencies, and others, were planned to be addressed the following year in 2018 when the course Instructor and the TA were made aware of the Aropä online peer-review system. It was developed and is voluntarily maintained by two academics in the School of Computing Science at the University of Glasgow and has been provided free, worldwide, since 2009. Several features of the Aropä system would directly address many of the deficiencies identified in the paper-based process used in the Fall of 2017. The students would be able to access the system 24-hours a day. This would allow them to schedule their work at their convenience. With the Aropä system a random assignment of submissions to assessors, with the capability to handle submissions covered by NDAs, and customizable assessment rubrics is possible. Based on the available features and relative ease of use, for both the students and the instructional team, it was decided to trial using the Aropä system with the 2018 Individual Draft of the Final Design Report assignment.

All teams work on unique projects and so copying content gleaned from the review process to include in thier own Final Design report submission is not a concern.

## Cognitive Assessment Redesign

The course Instructor and TA were made aware of the Aropä system, by colleagues from the Faculty of Engineering and Applied Science Office, during a meeting to discuss the possibility of applying for a grant from the Cognitive Assessment Redesign (CAR) initiative. CAR is an internal Queen's source of funds. One element of the initiative was to provide funding for improving undergraduate assignments and rubrics that evaluate critical thinking. While critical thinking is a very important part of all Capstone designs, there was no explicit assessment of this in any assignment in the course. If a low-stakes critical thinking assignment could be developed for use in the course, the instructional team would be able to trial the Aropä system and work out any issues with the process before the Individual Draft assignment.

Incorporating an assignment on the topic of Motivation had been a past objective for the course and as funding from the CAR initiative was available, it was decided to design such an assignment that would assess critical thinking and use the Aropä online peer-review system for assessment and feedback. The rubric for assessing this assignment was developed in collaboration with a colleague who was involved with the CAR initiative.

## Motivation Assignment Methodology

In the Motivation assignment the students were asked to read an article, Five Research-Tested Team Motivation Strategies by Richard E. Clark (2005) [3], and to "not only read and think critically about the assigned article, but also to reflect on your personal motivations, attitudes and behaviours. And bring your personal experience working on school teams to the table to compare and contrast the assigned article to your personal motivation experience

while working on school project teams" in a two-page submission.

Students were assigned an identification number, different than their StudentID for reasons of privacy and anonymity, to use in the Aropä system. Submissions were in the form of a PDF file. The Aropä system randomly assigned each student, who made a submission, two of their peer's motivation submissions to review. Students were given one week to complete their reviews. For each review the student was presented with an online version of the rubric and asked to assess three critical thinking elements, "Synopsis of Article", "Analysis, Discussion and Application", and "Conclusions", and five communication elements, "Concision", "Precision", "Clarity", "Logical Progression of Thoughts", and the overall "Structure and Appearance" of the submission. The online rubric also had four comment boxes, one for each of the three critical thinking elements and one for the communications elements. A five-level scale for assessments was used, "Not Demonstrated", "Marginal", Developing", "Expectation", and "Outstanding" corresponding to an "F", "D", "C", "B", and "A" in letter grades. Students did not complete a 'rater practice' activity before using the Aropä system.

The TA reviewed all the high and low assessments (subjectively chosen as those above 90% and below 60%), as well as any submission where the difference between the total weighted score of the two peer-review assessments was greater than 20%, and any submission where only a single review was conducted. Students were awarded marks for this assignment based on the average of the two reviews they received from their peers, as well as marks for each review that they completed. Adjustments were made, if warranted, based on the TA's review process.

An additional random sample of 30 submissions, excluding any submissions that had already been reviewed by the TA, were assessed by the TA.

## Motivation Assignment Results

Of the total 167 student submissions completed in 2018, 31 submissions were required to be reviewed by the TA, with 26 submissions having a difference in total weighted score greater than 20% and 5 submissions were only reviewed by a single peer.

The distributions of scores shown in Figure 1 for peers are based on the average of the two total weighted scores from each of the assessment rubrics for the author of a submission. Numerical scores were created by converting "Outstanding" to 5 out of 5, "Expectation" to 4 out of 5, "Developing" to 3 out of 5, etc. Scores are shown as a percentage out of 100. "Non-Adjusted Peers" is the raw data from all student assessors. Submissions where only a single review was conducted were excluded from the "Non-Adjusted Peers" distribution. "Adjusted Peers" is the data after the adjustments, based on the TA review, have been applied. The "TA Only" distribution only includes their 61 assessments.

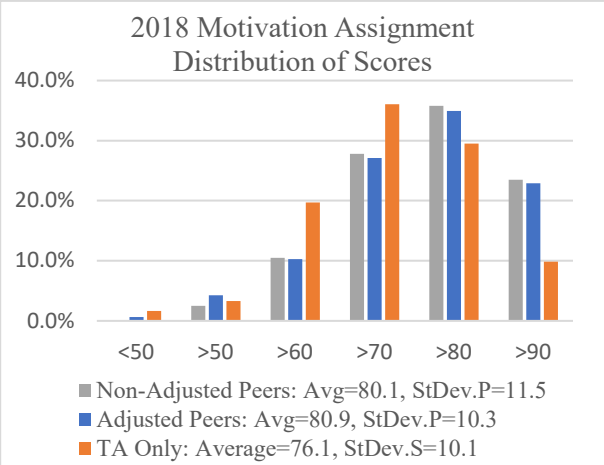


*Figure 1: Comparing distribution of non-adjusted peer-review scores, adjusted peer-review scores and TA review scores for the 2018 Motivation assignment.*

It appeared that students were more generous with their assessment scores on average than the TA. During the review of the assessment scores and comments, by the TA and instructor, it was noted that assessors tended to be more honest, i.e. critical, with the comments provided and slightly biased toward awarding higher scores. This was not completely unexpected, and this bias was deemed to be acceptable for a low-stake assignment in terms of awarding marks. Based on these results it was felt that the majority of the students in the class took their assessor role seriously.

To determine if the quality of the feedback provided was reasonable a subjective review of feedback was conducted by the TA and Instructor, along with an analysis of the total count of words provided in the four comment boxes. As students were asked to provide detailed comments to justify the assessments that they made, simple one or two word comments, such as, "Well Done", or "Very Good", were not consider to be reasonable feedback. We were looking for more substantive feedback to be considered reasonable. Students were not given any guidelines on the minimum or maximum number of words to include in their comments.

Figure 2 shows that over 80% (86.2%) of the class received at least 200 words of comments in total from all the reviews that they received. Over 98% (98.1%) of the class received at least 150 words, with five students receiving over 1,000 words, and everyone received at least 84 words of comments as feedback. The vast majority of the comments that were reviewed were at least at a Level 2, being reasonable in both quality of effort and detail provided.

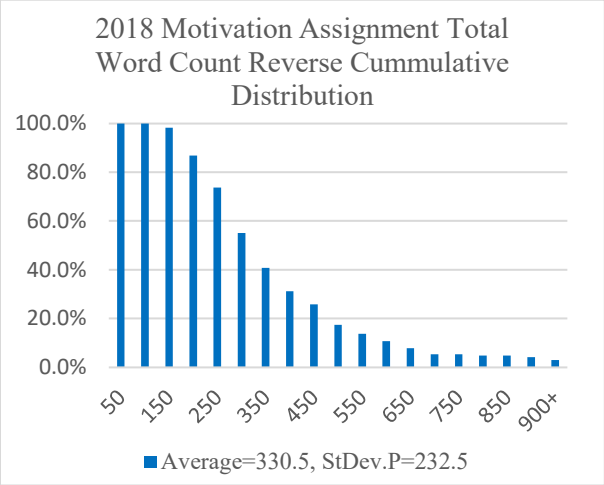


*Figure 2: Distribution of total word count of feedback provided by peer reviewers in the four comment boxes in the online rubric.*

Based on what we deemed to be a positive experience with the Motivation assignment the Aropä online peer-review system was used for the Fall 2018 and Fall 2019 Individual Draft of the Final Design Report assignments.

## FeedbackFruits 2020 and 2021

In the Fall of 2020 the University licensed a suite of tools that was integrated into the campus wide Learning Management System (LMS) from a third-party supplier called FeedbackFruits. The Peer Review tool provides very similar features and functionality compared to the Aropä online system. The primary advantage of using FeedbackFruits over the Aropä online system is the integration with the University's D2L Brightspace LMS. This integration provided a familiar interface for the students to access the tool, upload submissions, provide feedback, and simplified the process of posting the assessment results to the class. Due to the success experienced using a peer review process, in the Fall of 2021 three additional peer reviewed assignments were added to the course.

## Conclusion

Based on the results from the Motivation and Individual Draft assignments a peer-review process has been successful in the Capstone Design course 11 times over the past five years, one paper based, then six times using the Aropä online peer-review system and four times using the FeedbackFruits tool.

The level of effort by reviewers to provide both reasonable quality and quantity of feedback, as well as the distribution of assessment scores has been consistent over all the assignments using the peer-review approach.

Recognizing that total word count does not necessarily represent the quality of the feedback, and that there would be duplication in feedback received from the different reviewers, it still seems that the majority of the class took both aspects of the peer-review process seriously and that authors received reasonable feedback about their submission.

Total workload for both the TA and Instructor is significantly less than what was required to assess and provide feedback on the Individual Drafts assignments prior to 2017, even when considering the administrative time associated with both the Aropä and FeedbackFruits systems, and this also includes the time to review any outliers assessments.

Utilizing a peer review system has allowed the due date for the Individual Drafts assignment to be pushed into week 11 with individual feedback now provided by the middle of week 12. This gives students six to seven days to incorporate feedback into their Final Design report submission.

To improve the quality of the feedback, additional resources, either online or in-class, on best practices when reviewing a peer's work should be considered.

## References

[1] A. Lui and H. Andrade, "Student Peer Assessment," in *Encyclopedia of Science Education*, Dordrecht, Springer Netherlands, 2015, pp. 1003-1005.

[2] F. Rios and R. Alba-Flores, "Successful Capstone Design Projects Using Peer Review," in *Capstone Design Conference* , Rochester, 2018.

[3] Clark, R. E. Research-Tested Team Motivation Strategies. Performance improvement (International Society for Performance Improvement) 2005, 44 (1), 13–16. https://doi.org/10.1002/pfi.4140440107.

[i] While it is commonly referred to as the Capstone Design course, it is two single-term concurrent courses, MECH 460, a full credit course, under which essentially the elements of design are assessed and MECH 464, a half credit course, under which essentially the communications and project management elements are assessed.