

# Peer Assessment Tool to Provide Mid-Course Student Feedback

Lynne Cowe Falls PhD. P.Eng.

Department of Civil Engineering, Schulich School of Engineering, University of Calgary

## ABSTRACT

The Canadian Engineering Accreditation Board (CEAB) as well as the CDIO organization, have adopted graduate attributes/learning outcomes upon which engineering schools can evaluate their progress and/or be accredited against. Five of the CEAB graduate attributes are focused on easily defined and assessed technical skills (knowledge base for engineering, problem analysis, investigation, design and, use of engineering tools) but seven of the attributes are more difficult to define and to assess within the bounds of traditional engineering curricula. One of the more difficult graduate attributes to build curriculum around and to assess is individual and teamwork defined, as “An ability to work effectively as a member and leader in teams, preferably in a multi-disciplinary setting.” Assessing individual work is quite easy, but assessing teamwork can be challenging. One of the tools available to assess team dynamics is peer assessment and/or evaluation surveys. Peer assessment and evaluation is frequently used in group work courses in two ways: a) to distribute the grades amongst the group after the work is completed and, b) to provide students with an opportunity to vent their frustration, daylight group issues that should have been brought forward earlier in the course and/or reflect on the entire experience. Most of the evaluation tools/processes are used at the end of the course, which does not provide any opportunity for students to recover from a bad evaluation and/or adjust their behaviour as a result of negative feedback from their peers.

This paper will describe the process for multiple peer evaluations within the framework of an eight-month capstone design course that provide opportunities for students to reflect on their performance. The method used was a transparent way for marks to be allocated based upon peer evaluation and this paper provides a roadmap for incorporation of this tool into large and/or long group projects.

## Keywords

Standard 11: Learning Assessment; Standard 8: Active Learning.

CDIO syllabus 2.5:professional skills and attitudes; 3.1:multi-disciplinary teamwork; 3.2 communications

## INTRODUCTION

The Canadian Engineering Accreditation Board (CEAB) as well as the CDIO organization, has adopted graduate attributes/learning outcomes upon which engineering schools can evaluate their progress and/or be accredited against. Five of the CEAB graduate attributes are focused on easily defined and assessed technical skills (knowledge base for engineering, problem analysis, investigation, design and, use of engineering tools) but seven of the attributes are more difficult to define and to assess within the bounds of traditional engineering curricula. One of the more difficult graduate attributes to build curriculum around and to assess is individual and teamwork defined, as “An ability to work effectively as a member and leader in teams, preferably in a multi-disciplinary setting.” Assessing individual work is quite easy, but assessing teamwork can be challenging. One of the tools available to assess team dynamics is peer assessment and/or evaluation surveys.

Peer assessment and evaluation is frequently used in group work courses in two ways: a) to distribute the grades amongst the group after the work is completed and, b) to provide students with an opportunity to vent their frustration, daylight group issues that should have been brought forward earlier in the course and/or reflect on the entire experience. Most of the evaluation tools/processes are used at the end of the course, which does not provide any opportunity for students to recover from a bad evaluation and/or adjust their behaviour as a result of negative feedback from their peers.

## BACKGROUND

The Schulich School of Engineering at the University of Calgary, like all Canadian universities requires that students in their final year complete a two-term capstone design project. The projects come from industry and are open-ended real-world problems that should take a minimum of 200 hours per student to complete over two, four-month terms. The groups and projects vary from department to department and group size can be as small as four and as large as twelve students. Each project has a faculty advisor and industry advisor who together provide guidance but not direct technical support. The Department of Civil Engineering groups generally fall into the large category with groups as large as twelve and projects can range from very specific, such as design of a pedestrian bridge to multi-disciplinary as greenfield subdivision development plans. Because of the size and scope of the projects, most groups had four phases: development of the terms of reference, research, conceptual design and detailed design and prototyping was not possible.

Student assessment was comprised of five components:

1. Initial terms of reference (5%)
2. End of fall term report (20%)
3. End of project report and poster (60%)
4. Two presentations (end of fall term and end of winter term) (10%)
5. Personal learning journal. (5%)

One of the first requirements for the course was the initial terms of reference wherein the group defined the scope, goals and objectives of their project and the final deliverables (drawings, conceptual plans, etc.). Because this was a group project course, all but assessment item number five (the personal learning journal), the final grade was very much dependent upon the whole group and like many groups, the potential existed for a lot of the work being done by a

few members and the remaining members benefitting from their efforts – the so-called ‘dead cow’ effect. The personal learning journal was a weekly requirement of the course to report the number of hours spent each week by major task (‘billable hours’) as well as an opportunity to reflect on the group/project progress, knowledge gained and/or difficulties encountered and personal feelings for the week. Three times during the term, the students had to submit a compiled personal learning journal that required a look back at the previous six to eight weeks and reflect on how far they had come. The personal learning journal is described in more detail in Cowe Falls 2015.

To overcome this, previous iterations of the course had included a year-end peer and course evaluation. The purpose of the evaluation was twofold: to provide students with a mechanism for commenting on how the course had progressed (everything from logistics and course organization to comments on the relative difficulty and merit of specific project topics) and secondly, to provide a tool whereby students could evaluate their peer’s performance. The course coordinator adjusted individual grades using the evaluation and there were cases where a student’s mark was adjusted from a pass to a fail on the basis of a poor peer evaluation. As the evaluation took place at the end of the course, the results of the peer evaluation were not transparent and students had no recourse to correct poor performance.

#### **PEER ASSESSMENT PROCESS – 2013/2014**

For the 2013-2014 capstone design course, a different approach was taken, one whereby peer assessment was done three times during the course and the results returned to the students. The peer assessment form comprised of a table with a column for each team member (with the name pre-filled into the top row of the table) and twelve rows, one row for each of eleven questions and, a final row for the total of each column. The students’ names were pre-filled for each team not only for simplicity of data entry, but also to provide some security for the student evaluators as it was impossible to identify the rater from the handwriting.

Each student had to assess themselves, as well as the other team members, on a scale of 0 (the lowest/worst evaluation) to 10 (the highest/best evaluation) for each of the following eleven questions:

1. The effort he/she made to coordinate his/her portion of the work with the group.
2. Quality of the work done.
3. His/her effectiveness in gathering data necessary for the success of the project.
4. Attendance and contributions at group meetings.
5. The extent to which she/he did their share of the project.
6. How well she/he contributed to the team achieving its goals – attitude and use of initiative.
7. Contribution to preparing/submitting documents and presentations.
8. How conscientiously did he/she meet team deadlines.
9. Extent to which he/she contributed ideas to, and participated in, group analysis of the project.
10. How cooperative he/she was during the project.
11. Overall performance as a team member.

As can be seen, the questions pertain not just to quality of the work being done by the student (Q2) but also to their attendance (Q4), initiative (Q 6 & 9), contribution and deliverables (Q7, 8 & 9) and their effort and overall performance (Q1, 3, 5, 10 & 11). Free-form comments were

invited at the end of the forma and while very few students included comments, those that did, generally elaborated on why a group member was being marked down. Any free-form comments were not transmitted to the students to protect team dynamics, but they were used by the course coordinator, in one case, to intervene through an all-group meeting for a group that was in serious jeopardy of disintegrating.

Students were also asked to distribute the grades amongst the group by assigning a total of 100 marks times the total number of group members. This meant that in an eight-person team, there were 800 marks available for distribution and in a three-member team, there were only 300 marks available for distribution. The total number of marks constrained the grade distributions and students could give someone more marks for having taken on more work i.e., 110 / 100, but to do that someone else in the team had to receive fewer grades i.e., 90/100. The grade distribution was a further check on the students' evaluations and in some cases provided more insight to team dynamics than the eleven questions. Two teams decided to give every member 10 out of 10 for every question in the evaluation, but individuals within the team allocated the grades differentially to a few members who were seen to be contributing more to the project.

Within a week of the peer assessment deadline, the scores were reported back to the students by an individual email and each student was given their mean score by question, the group mean score by question and their 'delta' score which was the difference between their mean and the group mean. This was a way to normalize their ratings relative to the overall group rather than to the class as a whole. They were also given an overall mean evaluation, delta mean evaluation and the mean grade distribution.

Table 1 presents raw data from a group of five students, showing the ratings per evaluator by student and question (Q1 to Q11). In the first column, the student being rated is listed with the rater shown in brackets (Joe (Leah) represents the score given to Joe by Leah) In this group, one student (Dan, shown in *italics*) is being rated more severely by the other students in the group with some scores in the sub 5 range. Joe is the notable exception as he has rated Dan as 10 across the board. Dan has also rated himself as below par in some questions (2, 3, 4, 9 & 11). In the final column the mean rating for the student is given and as can be seen Dan has lower evaluations that the others.

Table 1: Individual Ratings by Student, Student Rater and, Question for a Group of Five Students (Joe, Tim, Leah, Dan Sue and Pat) in Peer Evaluation No. 1 (Week 10).

| Student (rater) | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q8 | Q9 | Q10 | Q11 | Mean rating |
|-----------------|----|----|----|----|----|----|----|----|----|-----|-----|-------------|
| Joe (Joe)       | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10  | 10  | 10.0        |
| Joe (Tim)       | 10 | 8  | 8  | 9  | 8  | 9  | 9  | 10 | 8  | 10  | 9   | 8.9         |
| Joe (Leah)      | 6  | 5  | 8  | 7  | 6  | 6  | 7  | 9  | 7  | 8   | 8   | 7.0         |
| Joe (Dan)       | 10 | 6  | 6  | 6  | 8  | 9  | 8  | 9  | 8  | 10  | 7   | 7.9         |
| Joe (Sue)       | 8  | 7  | 8  | 7  | 8  | 8  | 8  | 8  | 7  | 8   | 8   | 7.7         |
| Tim (Joe)       | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10  | 10  | 10.0        |
| Tim (Tim)       | 10 | 9  | 8  | 10 | 8  | 9  | 10 | 10 | 9  | 10  | 9   | 9.3         |
| Tim (Leah)      | 9  | 9  | 7  | 9  | 9  | 9  | 9  | 9  | 8  | 8   | 9   | 8.6         |
| Tim (Dan)       | 10 | 9  | 9  | 9  | 9  | 9  | 9  | 9  | 10 | 10  | 9   | 9.3         |
| Tim (Sue)       | 9  | 8  | 7  | 9  | 8  | 8  | 8  | 8  | 8  | 9   | 9   | 8.3         |
| Leah (Joe)      | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10  | 10  | 10.0        |

|                   |      |      |      |      |      |      |      |      |      |      |      |      |
|-------------------|------|------|------|------|------|------|------|------|------|------|------|------|
| Leah (Tim)        | 10   | 9    | 8    | 10   | 9    | 9    | 10   | 10   | 9    | 10   | 9    | 9.4  |
| Leah(Leah)        | 9    | 8    | 8    | 9    | 10   | 9    | 9    | 9    | 9    | 9    | 9    | 8.9  |
| Leah (Dan)        | 10   | 9    | 8    | 9    | 9    | 9    | 10   | 9    | 10   | 10   | 9    | 9.3  |
| Leah (Sue)        | 9    | 8    | 8    | 10   | 8    | 8    | 8    | 8    | 9    | 8    | 9    | 8.5  |
| Dan (Joe)         | 10   | 10   | 10   | 10   | 10   | 10   | 10   | 10   | 10   | 10   | 10   | 10.0 |
| Dan (Tim)         | 10   | 8    | 8    | 9    | 8    | 9    | 9    | 10   | 8    | 10   | 9    | 8.9  |
| Dan (Leah)        | 6    | 6    | 3    | 6    | 7    | 6    | 7    | 9    | 6    | 6    | 7    | 6.3  |
| Dan (Dan)         | 10   | 6    | 5    | 5    | 9    | 8    | 8    | 9    | 5    | 10   | 6    | 7.4  |
| Dan (Sue)         | 8    | 8    | 7    | 6    | 8    | 8    | 8    | 8    | 5    | 8    | 8    | 7.5  |
| Sue (Joe)         | 10   | 10   | 10   | 10   | 10   | 10   | 10   | 10   | 10   | 10   | 10   | 10.0 |
| Sue (Tim)         | 10   | 9    | 8    | 9    | 8    | 9    | 10   | 10   | 9    | 10   | 9    | 9.1  |
| Sue (Leah)        | 10   | 9    | 10   | 10   | 10   | 9    | 10   | 9    | 9    | 10   | 9    | 9.5  |
| Sue (Dan)         | 10   | 9    | 9    | 10   | 9    | 9    | 10   | 9    | 10   | 10   | 9    | 9.5  |
| Sue (Sue)         | 10   | 9    | 10   | 10   | 9    | 8    | 8    | 8    | 10   | 8    | 10   | 9.0  |
| Q <sub>Mean</sub> | 9.23 | 8.40 | 8.20 | 8.66 | 8.66 | 8.89 | 9.11 | 9.43 | 8.51 | 9.14 | 8.89 | 8.83 |

Table 2: Mean Evaluation by Student and Question

| Student | Q1  | Q2  | Q3  | Q4  | Q5  | Q6  | Q7  | Q8  | Q9  | Q10 | Q11 | Mean rating |
|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-------------|
| Joe     | 8.7 | 7.6 | 8.0 | 8.0 | 8.1 | 8.6 | 8.6 | 9.4 | 8.0 | 9.0 | 8.4 | 8.40        |
| Tim     | 9.4 | 8.9 | 8.3 | 9.1 | 8.7 | 9.1 | 9.3 | 9.4 | 8.9 | 9.3 | 9.1 | 9.05        |
| Leah    | 9.4 | 8.7 | 8.6 | 9.3 | 9.0 | 9.1 | 9.4 | 9.4 | 9.3 | 9.3 | 9.3 | 9.17        |
| Dan     | 8.7 | 7.9 | 7.0 | 7.3 | 8.4 | 8.4 | 8.7 | 9.4 | 7.0 | 8.7 | 8.1 | 8.16        |
| Sue     | 9.9 | 9.0 | 9.1 | 9.6 | 9.0 | 9.1 | 9.6 | 9.4 | 9.4 | 9.4 | 9.4 | 9.36        |

The raw data was processed using a pivot table to calculate a mean value for each question and an overall mean evaluation score. To normalize the group, the mean score for each question was calculated ( $Q_{\text{Mean}}$ ) as shown in Table 2 and the student delta relative to each question was calculated as shown in Table 3.

Table 3: Individual Student Delta by Question Peer Evaluation No. 1 (Week 10)

| Student | Q1    | Q2    | Q3    | Q4    | Q5    | Q6    | Q7    | Q8   | Q9    | Q10   | Q11   | Delta Evaln. |
|---------|-------|-------|-------|-------|-------|-------|-------|------|-------|-------|-------|--------------|
| Joe     | -0.51 | -0.83 | -0.20 | -0.66 | -0.51 | -0.31 | -0.54 | 0.00 | -0.51 | -0.14 | -0.46 | <b>-0.43</b> |
| Tim     | 0.20  | 0.46  | 0.09  | 0.49  | 0.06  | 0.26  | 0.17  | 0.00 | 0.34  | 0.14  | 0.26  | 0.22         |
| Leah    | 0.20  | 0.31  | 0.37  | 0.63  | 0.34  | 0.26  | 0.31  | 0.00 | 0.77  | 0.14  | 0.40  | 0.34         |
| Dan     | -0.51 | -0.54 | -1.20 | -1.37 | -0.23 | -0.46 | -0.40 | 0.00 | -1.51 | -0.43 | -0.74 | <b>-0.67</b> |
| Sue     | 0.63  | 0.60  | 0.94  | 0.91  | 0.34  | 0.26  | 0.46  | 0.00 | 0.91  | 0.29  | 0.54  | 0.54         |

Comparing the student's rating against the mean for each question provides the student with an understanding of how they are perceived relative to the group mean. If a score is negative, then they are underperforming relative to the group and if the score is positive, they are over performing relative to the group. In Table 3, Joe and Dan are evaluated as under performers by the group with deltas of -0.43 and -0.67, respectively while Sue is evaluated as the highest performer relative to her peers with a delta of 0.54.

The evaluations were repeated three times: at week ten, week 18 and week 26 (the final week) which provided the students with opportunities to improve on poor ratings before they had a major impact on the final grade. After each evaluation, the students received individual emails providing them with their mean score by question, the group mean score by question and their delta values for each question as well as, their overall delta (the final column in Table 2). This allowed them to see where they were considered weak and where they were evaluated as strong by their peers. Any student with large negative deltas were contacted by the instructor via email and asked if they would like to discuss their evaluations and this provided an opportunity for one-on-one coaching.

As mentioned above, part of the course assessment included a personal learning journal and some students commented on their peer assessment in their journal entries:

*“My goal for second peer evaluation is to get much higher mark than what I received before. In order to make this happen, I will have to speak more actively with members in the meeting.”*

*“My peer evaluations came in and I was extremely happy about these! I was scored below average but am fine with this as I do think it reflects where I am in the group.”*

Table 4 presents the final calculation table for the group with the three peer evaluation deltas (negative values are shown in bold), the cumulative delta (a summation of the PE 1-3 values), the mean delta (the average of PE1-3) and, the minimum/maximum bonus or penalty. The final grade for each student was adjusted using the mean of the three delta evaluations and no adjustment was greater than plus or minus 10% of the group grade. In this example, students received either the maximum 10% bonus or penalty, but in cases where the mean delta value was less than 10% (delta 0.05 or -0.03) the actual value was used. Because the maximum penalty/bonus was 10% mean delta values greater than 0.10 were fixed at 0.10. For example, Leah has a mean delta of 0.19, but her bonus is 0.10. The final report grade of 82 was adjusted using four scenarios (PE3 delta, cumulative delta, mean delta and min/max delta). The final mark was calculated using the final grade (col G) times the min/max value (col F). In order to provide transparency to the students, the final grade adjustment was communicated by email in advance of the official grade postings. This allowed for any appeals, but none were submitted.

Table 4: Final Individual Student Delta by Question Peer Evaluation and Grade Adjustment

| Student | Delta PE 1 (A) | Delta PE (B) | Delta P3 (C) | Cum. Delta (D) | Mean Delta (E) | Min/max set (F) | Final report mark (G) | Adj. Delta 3 (G*C) | Adj cum (G*D) | Adj Mean (G*E) | Final Mark (G*F) |
|---------|----------------|--------------|--------------|----------------|----------------|-----------------|-----------------------|--------------------|---------------|----------------|------------------|
| Joe     | <b>-0.43</b>   | <b>-0.23</b> | <b>-0.04</b> | -0.70          | -0.23          | -0.10           | <b>82</b>             | 79                 | 25            | 63             | 74               |
| Tim     | 0.22           | 0.30         | <b>-0.22</b> | 0.30           | 0.10           | 0.10            | <b>82</b>             | 64                 | 107           | 90             | 90               |
| Leah    | 0.34           | 0.19         | 0.03         | 0.56           | 0.19           | 0.10            | <b>82</b>             | 84                 | 128           | 97             | 90               |
| Dan     | <b>-0.67</b>   | <b>-1.05</b> | <b>-0.11</b> | -1.83          | -0.61          | <b>-0.10</b>    | <b>82</b>             | 73                 | -68           | 32             | 74               |
| Sue     | 0.54           | 0.79         | 0.33         | 1.66           | 0.55           | <b>0.10</b>     | <b>82</b>             | 109                | 218           | 127            | 90               |

The final grades were consistent with the three peer evaluations. As can be seen in Table 4, Joe and Dan were consistently ranked below the group mean in each evaluation and their final marks are reduced. Dan was the weakest member of the group as ranked by his peers, but he did improve in the final peer evaluation from -1.05 to -0.11. Tim’s third peer evaluation sank in

the third evaluation period possibly because of the rotation of project management duties which put him in the captain's seat at a critical time (final design and report production) and the group rated him more harshly in this period. Leah was a consistent high performer, but the star of the group was Sue who perhaps should receive more than the maximum 10% because of her high ranking throughout the project, similarly, it could be argued that Dan should receive more than the maximum 10% penalty for consistent weakness.

One group in the course decided not to evaluate anyone harshly and all members rated each other at 10. After the first peer evaluation it was pointed out the group that there was a danger of coat-tail riding (or 'dead cow syndrome') whereby someone will benefit as a result of the efforts of others, however, they were adamant in their plan to evaluate everyone with a 10.

## **Conclusion**

Learning to work effectively as a member and leader in teams is one of the graduate attributes of the Canadian Engineering Accreditation Board's and it is, by definition, experiential in nature. Peer assessment on a recurring and/or repeating basis as described in this paper can be a way for students to learn how to fairly and constructively evaluate their peer's work and hence, learn to work in teams. As students transition into the work place, performance measurement is generally done by their supervisor and it may be several years before a young engineer is required to conduct a performance review on a peer and/or direct report. Learning how to assess teamwork constructively can be a challenging part of job and in the work place non-performing team members can have a cost. Teaching performance or peer assessment is an important skill to develop as a leader to enable future success in multi-disciplinary teams.

## **REFERENCES**

Cowe Falls, Lynne 2015, "Personal Learning Journals in a Capstone Design Course as a Tool to Encourage Life-Long Learning", CDIO 11<sup>th</sup> International Conference, Chengdu, China

## **BIOGRAPHICAL INFORMATION**

**Lynne Cowe Falls**, PhD, P. Eng., FCAE, FCSCE, is an Associate Professor in Civil Engineering at the Schulich School of Engineering, the University of Calgary. She is a co-author of over 30 technical papers and several books in the area of pavement and infrastructure management and most recently of *Current Pavement Management*. With over 20 years in industry prior to joining the University of Calgary, she is a Vice-President and Board Member of the Transportation Association of Canada.

## **Corresponding author**

Dr. Lynne Cowe Falls  
Schulich School of Engineering  
University of Calgary  
Calgary, Alberta, Canada T2N 1NW  
1-403-220-5505  
[lcowefal@ucalgary.ca](mailto:lcowefal@ucalgary.ca)



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License](https://creativecommons.org/licenses/by-nc-nd/3.0/).