Higher Education Quality Council of Ontario

An agency of the Government of Ontario

Teaching Team-effectiveness in Large Classes

Patricia K. Sheridan, Greg J. Evans,
Doug Reeve, University of Toronto

Published by

The Higher Education Quality Council of Ontario

1 Yonge Street, Suite 2402
Toronto, ON Canada, M5E 1E5

Phone: (416) 212-3893
Fax: (416) 212-3899
Web: www.heqco.ca
E-mail: info@heqco.ca

Cite this publication in the following format:

Sheridan, P. K., Evans, G. J., & Reeve, D. (2015). *Teaching Team-effectiveness in Large Classes.* Toronto: Higher Education Quality Council of Ontario.

Executive Summary

Instruction of team skills is quickly emerging as an important and missing dimension of engineering education. This project evaluated a new framework for guiding students in providing self- and peer assessments of their effectiveness in teamwork. This framework is the foundation for a new web-based tool that offers students structured feedback from teammates, along with personalized exercises and actionable strategies that guide targeted learning in the areas thereby identified. Specifically, the study documented in this report investigated whether the feedback framework, when used for intra-team self and peer feedback, increased students' abilities to learn about and improve their team-effectiveness in executing design projects.

The framework consisted of 27 competencies across three aspects of team-effectiveness: organizational, relational and communication competencies. The framework was tested in a randomized controlled experiment in a first-year engineering design class of 280 students against an unstructured feedback prompt. Students were asked to provide feedback at the mid-point of the course and to provide their thoughts on the utility of the feedback they received in an end-of-term survey. Student assessments were also compared to teaching assistant assessments.

Students using the framework found the feedback they received to be more actionable than unstructured feedback, and found that it motivated them more to improve their performance than did the students receiving unstructured feedback. Students in the unstructured group received feedback on fewer and less diverse team-effectiveness competencies than those in the framework group. Students in the unstructured group received feedback on approximately 10 competencies on average, and approximately half of that feedback was on organizational competencies; students in the unstructured group received little feedback on relational competencies. The unstructured feedback also primarily identified the students' strengths, which were phrased as praise of their performance, with minimal discussion of the students' weaknesses.

Students in the framework group were able to peer-assess their team members accurately when using the framework but were less able to self-assess accurately. When compared to the assessments of their teaching assistants, students' peer assessments correlated significantly with teaching assistants' assessments across all three aspects of team-effectiveness. However, only students' self-assessments along the organizational aspect correlated significantly with teaching assistants' assessments.

While students in the unstructured group did not receive this more comprehensive feedback, they did comment that the textual feedback made them feel more committed to their team, as it demonstrated that their team members had an interest in them and in the team as a whole. Based on this benefit and on students in the feedback group's requests for examples and comments in their feedback, we believe that a hybrid of the feedback framework along with some textual feedback would be the best method for providing feedback moving forward.

Table of Contents

List of Figures

List of Tables

Acknowledgments

The authors would like to thank Jason Foster and Rob Irish for allowing this research to occur in their course, Jennie Phillips for her assistance in facilitating and transcribing the participant focus groups, Lobna El Gammal for her assistance with the data analysis, Estelle Oliva-Fisher for her assistance with question development, Cindy Rottman for her guidance with the research methodology, Jason Foster for his technical assistance with the surveying system, as well as Susan McCahan, Jason Bazylak and Peter Weiss for their constructive input in the design and execution of the study.

The authors would like to especially thank all of the students and teaching assistants who made this study possible.

1. Introduction

Intentional instruction of team skills is quickly emerging as an important and missing dimension of engineering education. This project investigated and evaluated a new intervention that supports personalized learning of team effectiveness in large undergraduate courses. Specifically, we evaluated the effectiveness of a new web-based tool to provide students with structured feedback from teammates, along with personalized exercises and actionable strategies that guide targeted learning in the areas thereby identified. This intervention aimed to provide students with a safe, virtual environment in which they could: i) learn about their team-effectiveness and team issues, and ii) identify methods to improve their areas of weakness before trying them with their team members. A fundamental component of this intervention is the feedback framework that forms the basis of the tool. The feedback framework grounds the assessments students provide and the feedback they receive along specific competencies, guiding students to recognize and reflect on necessary teamwork skills. This study aimed to investigate whether the feedback framework, when used for intra-team self and peer feedback, would increase a student's ability to learn about and improve their team-effectiveness in their project teams.

2. Background and Motivation

Engineering is a team-based profession that requires students to be both technically proficient and effective at teamwork. The Canadian Engineering Accreditation Board mandates that all students who graduate from engineering programs be capable of working effectively alone as well as in teams (Canadian Engineering Accreditation Board, 2010). Upon graduation, when students transition to working in industry, they will work on large-scale, complex problems that require multiple individuals working effectively in teams to design appropriate solutions. In endeavouring to prepare students for this environment, most undergraduate engineering programs have students work on team-based projects that model what the students will experience in industry. The projects endeavour to instill in students the skills necessary to work effectively in these teams as well as in their future careers. However, students have traditionally been expected to develop teamwork skills implicitly simply by participating in the team project.

At the University of Toronto, engineering students are introduced to teamwork in their first-year engineering design courses. First-year classes, in which students learn the foundational skills they need for the remainder of their degrees and for their careers, are traditionally the largest. For example, class size in the engineering program at the University of Toronto ranges from 100 to 1,000 students, with the two first-year engineering design classes being the largest at approximately 280 and 1,000 students. In these courses, students have been grouped into teams to simulate industry working conditions. However, the sizes of these teams (ranging from three to seven students) are mainly constrained by facilities and instructor/teaching assistant time, especially grading time. In these large classes, providing feedback on student performance in a timely manner is always a challenge. As a result, these students have traditionally received feedback on the quality of their team's deliverables (e.g., design reports, presentations and prototypes) from their teaching assistants, but rarely receive feedback on how they as individuals are performing as team

members. Courses have traditionally informed students that the teaching assistants are available to support but not assess their team-effectiveness.

The team-based project environment in first-year engineering design courses is a high-stakes environment. The projects are too complex for one person to address on their own and require some level of collaboration between team members. The students do not always work effectively together but are still graded together – the whole team receives whatever grade the deliverable receives, regardless of how effectively they worked together. When students in the project teams do not pull their weight, the remainder of the students in the team learn not to like teamwork and not to trust their team members. They see teamwork as an undesirable experience being forced on them that jeopardizes the thing they value: their grades.

Our belief is that many of these issues could be mitigated by providing students with personalized support in the development of their teamwork skills. The first part of this support is providing students with an awareness of what is happening in their teams. In small classes, an instructor can do this thanks to the personal nature of teacher-student interactions resulting from a significantly decreased assessment load. However, even here, the instructor may only see part of what is occurring within a team. Thus, due to the way in which student teams work, we believe that the best people to provide this feedback, both in large and small classes, are the students themselves. They see their team members for the entire duration of the project, in particular in team meetings outside of class or tutorial time when instructors are not present. Our objective is to use self and peer assessment within the student project teams to provide personalized feedback to each of the team members on their effectiveness as a team member during the project.

Approaches to conducting self and peer assessments of teamwork skills in engineers have been a popular focus of research over the last five to 10 years. Web-based tools have been developed in the US, most notably the Comprehensive Assessment of Team-member Effectiveness (Loughry, Ohland & Moore, 2007), and in the UK, the WebPA (Loughborough University, 2009). These tools originated to ensure that instructors have sufficient information to know if they need to modify individual students' grades based on their relative contributions to a team deliverable. As these tools have expanded to become web-based resources to facilitate the learning of team-effectiveness, they have approached teamwork from a strong task-focused orientation. Students in engineering are already traditionally more oriented towards the task side of teamwork than to the people side, so there exists a greater need to facilitate their development of the relational and communication skills of teamwork. As a result, rather than adopting one of the existing frameworks, we developed our own that pushes students to think about relational skills in addition to task skills when providing and receiving feedback. We sought with this study to assess the effectiveness of this self and peer assessment feedback framework.

This report documents the assessment of our team member effectiveness feedback framework as a foundation to facilitate student development of their competency as team members. Feedback originating from the use of the feedback framework was compared to feedback developed from an unstructured feedback prompt to assess the usefulness and effectiveness of the feedback framework. This report first presents the framework in Section 3, outlines the study design and research questions in Section 4,

responds to these research questions individually in Sections 5-8, and finally concludes with our findings around the utility of the feedback framework. A discussion of planned future work is outlined in Section 9.

3. Feedback Framework

The team-effectiveness feedback framework that was used in this intervention is presented in **Table 3-1**.

Table 3-1: The 27 competencies of the feedback framework divided into three aspects of team member effectiveness. The competency numbers preceding each competency will be used to reference the competencies in the analysis.

Organizational Aspects	Relational Aspects	Communication Aspects
O1. Support team rules	R11. Build the trust of teammates	C20. Exchange information in a timely manner
O2. Attend team meetings prepared	R12. Motivate others on the team to do their best	C21. Introduce new ideas
O3. Contribute to making meetings effective	R13. Raise contentious issues in a constructive way	C22. Openly express opinions
O4. Do their fair share of the work	R14. Solicit input before proceeding	C23. Promote constructive brainstorming
O5. Deliver their work on time	R15. Adopt suggestions from other members	C24. Actively listen to teammates
O6. Produce high quality work	R16. Accept feedback about strengths and weaknesses	C25. Provide constructive feedback
O7. Help to plan, set goals, and organize work	R17. Show respect for other teammates	C26. Make sure that teammates understand important information and instructions
O8. Track team progress vs. your timeline	R18. Demonstrate accountability	C27. Help the team build consensus
O9. Encourage progress to meet goals and deadlines	R19. Collaborate effectively	
O10. Display dedication and determination		

Note: This framework has been analyzed in other studies as well. See Sheridan et al. (2013) and (2014).

This framework was developed from four existing inventories and designed to include a greater focus on non-task-related competencies (Bushe & Coetzer, 1995; Lingard, 2010; Moore, Diefes-Dux & Imbrie, 2006; Maxwell, 2011). A synthesis of these inventories was developed and redundancy between behaviours was eliminated. Competencies were categorized according to three aspects of team member effectiveness (**Table 3-1**): organizational competencies, which focus on the project management and task completion aspect of effective teamwork; relational competencies, which focus on how students build effective means of collaborating together and trust in each other; and communication competencies, which focus on how students leverage their interactions to exchange information and promote productive discussions.

For each competency, there is a behaviorally anchored rating scale that describes a student's level of engagement in utilizing that competency to improve team performance. Students assess themselves and their teammates according to a 7-point descriptive Likert scale, where an assessment of 1-2 represents an *unengaged* team member, 3-5 a *self-focused* team member, and 6-7 a *team-focused* team member. To obtain a team-focused assessment along each competency, the student has to demonstrate an ability to exhibit and promote the behaviours of the competency in their team members. The behaviourally anchored rating scale for all competencies is available in Appendix A. By structuring the assessments using a behaviourally anchored rating scale for each competency, students should: i) be able to assess using a common scale, increasing the consistency in assessments received by a student, and ii) be able to provide feedback that is specific enough that a teammate can identify their performance level and how to improve it.

4. Study Design and Research Methods

The objective of the study was to understand how students perceive, interpret and use our feedback framework to promote learning about team member effectiveness behaviours in their project teams through the use of self and peer assessment. Our goal was to design an effective framework to guide students to provide feedback in order to facilitate the learning of teamwork within student project teams in large classes, with minimal need for additional contact-hour resources. Our hypothesis was that:

> *Students can be guided to provide useful personalized feedback on team-effectiveness to their teammates using our team-effectiveness framework.*

A study to assess the utility of the framework in facilitating feedback was tested in the Winter 2012 term of Praxis II, a 280-student cornerstone design course in first-year engineering at the University of Toronto. This course previously had students provide some unstructured comments about their team members to the teaching team as a means of expressing concern in cases where team members did not complete equitable shares of the project work. As providing unstructured comments about each team member was already a mode of reflection about team-effectiveness in the course, we wanted to compare our framework to the pre-existing unstructured feedback method to determine whether our framework could be more useful in guiding students to improve their team-effectiveness behaviours. To assess this, we used the following four research questions to guide our inquiry:

1. Does the framework guide students to provide a greater breadth, quantity or accuracy of feedback?
2. Do students perceive the peer feedback from the framework to be more useful than unstructured feedback?
3. Can students provide feedback similar to that of a trained observer (course teaching assistant) when using the framework?
4. Is the framework accessible (e.g., jargon, descriptions, levels of competency) to students and trained observers?

The theoretical orientation taken to learning in this study is that of social constructivism. This is because the team acts as a learning community in which to develop a greater understanding of one's own team member effectiveness behaviours, as well as a greater understanding of team member effectiveness in general. Self and peer assessments of team member effectiveness act as scaffolding to support students in observing and thereby developing a better understanding of team member effectiveness behaviours. Improvement in behaviour is believed to be motivated both intrinsically and extrinsically through the desire to be a better team member and to create a better team, respectively.

4.1. Course Context

Praxis II is a first-year engineering design course in the first-year Engineering Science program at the University of Toronto. Praxis II is a 280-student course that takes place in the second term of the school year and builds on the introductory design, communication and teamwork principles covered in their first-term course, Praxis I. The course is built on the pedagogical models of Kolb, Perry and Vygotsky in the design and execution of its course lectures, tutorials and assignments. Specifically related to this study, the class (students and teaching assistants) attended three half-hour lectures on team-effectiveness, where they were introduced to Tuckman's (1965) and Lencioni's (2002) models of effective teamwork, the aspects and competencies of our feedback framework (**Table 3-1**), and how these behaviours manifest in highly effective, high-performance teams.

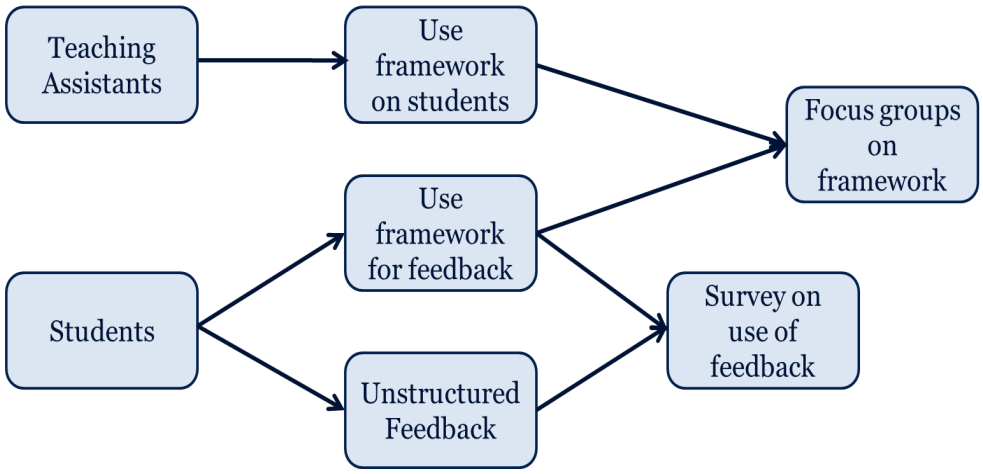
The course is designed as a service-, project- and team-based learning course in which students spend the first half of the 13-week term engaging with communities around the City of Toronto to identify engineering design opportunities they can address, and the second half of the term addressing a subset of opportunities selected by the teaching team and based on those identified in the first half of the term. The students work in self-selected teams of three to four students for the entire 13 weeks of the course. Students can select their team from the 22 to 27 students in their tutorial section; students may not be able to select to work with their friends if they are not in the same tutorial section. As a result of working with the same team for the 13 weeks of the course, students have sufficient time to determine how they and their team members work in teams, as well as improve their teamwork skills and behaviours based on feedback during the course.

The course has a teaching team of nine members: two course instructors, one of whom is a design expert and the other a communication expert, and seven teaching assistants, whose areas of expertise straddle engineering, the arts and the humanities. Students are grouped into tutorials of 22 to 27 students and are paired with two teaching assistants of complementary backgrounds. The role of the teaching assistants is to challenge the ideas of the students, instill good design and communication practices, model constructive and collegial disagreement, and demonstrate effective collaboration in their teaching style.

4.2. Study Design

The study looked at the effectiveness of the feedback framework by comparing it to unstructured feedback in a randomized controlled experiment within the same class. An overview of the study demonstrating which participants completed which components is provided in Figure 4-1.

Figure 4-1: Overview of the study design demonstrating which participants complete which assessments, surveys and attend focus groups



4.3. Student Participation

The entire class' teams were divided randomly and approximately equally into the unstructured and framework feedback groups. Individual students were then able to consent to letting us use their feedback for research purposes. Student teams were used to separate students into the experiment groups in order to ensure that all students within a team were using one type of feedback mechanism, and that students only received feedback of one type.

Immediately after the student teams submitted their first major deliverable in week 7, students were asked to provide self- and peer assessments of their effectiveness as team members. Students provided peer assessments within their project teams only for their team members.

Forty-eight percent of the class was grouped into the framework group and completed this assessment using our feedback framework (Appendix A). These students completed their assessments by selecting the descriptor that matched their team member's behaviour for each question. The remaining 52% were grouped into the unstructured group and completed this assessment by providing unstructured feedback that responded to the prompt: "Please provide feedback to yourself and your team members based on your/their team-effectiveness over the course of this project."

Students in both groups were asked to provide feedback for their team members online on sequential pages of a survey. Students completed these surveys on LimeSurvey, an open-source surveying software that was implemented on a secure server on campus. Students were provided a link to complete their survey, and depending on whether they were in the framework or unstructured feedback group, they saw and completed a different survey.

One week after completing the assessment, students received their self and peer feedback online for review. Students in the framework group received numerical feedback from their team members that corresponded to the descriptors the students used to provide their assessments (Appendix A). Feedback in the framework group identified for each student their lowest-ranked competencies as weaknesses (by highlighting their feedback along these competencies in red) and their highest-ranked competencies as their strengths (by highlighting their feedback along these competencies in green). The ranking of these competencies was determined according to a sorting algorithm that took into account the values of the self-assessment, peer assessments, and the difference between the two.

The algorithm worked to identify three strengths and three weaknesses for the students. However, in the case of competencies with tied rankings, the minimum number of ranks that resulted in three or more competencies being highlighted were coloured as strengths or weaknesses, respectively. For example, if four competencies all had the same ranking that was higher than all other competencies, only that rank (those four competencies) would be highlighted. Including all students who had tied rankings, the maximum number of competencies that was ever highlighted as a student's strengths or weaknesses based on the sorting algorithm was five.

Students in the unstructured group received unedited textual feedback, with no strengths and weaknesses intentionally highlighted.

Students in both groups also received a list of tools and techniques that they could use to enhance their effectiveness for each of the competencies outlined in the framework. An example of framework feedback is provided in Appendix B, and an example of unstructured feedback is provided in Appendix C.

At the end of the course, students in both groups completed the same end-of-term survey on the usefulness of the feedback they received (Appendix D). Questions in the survey analyzed the quantity, depth and breadth of the feedback, as well as its ability to motivate students to improve their effectiveness in teams based on the feedback.

Focus groups with students who used the framework were conducted after the end-of-term survey to assess the framework's utility, capture data on how participants perceived and used it, and their perceptions of the utility of the feedback it generated. The focus groups were conducted by an experienced focus group facilitator from outside of engineering and asked the questions outlined in Appendix E.

4.4. Teaching Assistant Participation

Teaching assistants (TAs) participated in the study by using the framework to provide assessments for all students in their tutorial sections. Teaching assistants were used as trained observers to assess the concurrent validity of the framework. TAs are typically the individuals who have identified and addressed team dysfunction situations and who students in the course approach when they have questions or issues with their teamwork. Additionally, with a class of 280 students, these were the only individuals who had sufficient one-on-one contact with the students in their teams, and were already scheduled to be supporting, and therefore observing, them in tutorials. The demographic information of the seven TAs is shown in Table 4-1. These TAs supported nine tutorial sections in pairs, with each TA staffing at least two tutorials and some staffing three. The pairing of TAs in tutorials provided a fortuitous opportunity to acquire two observations of each student with which to compare the students' self- and peer assessments. This allowed us to assess whether students provided similar feedback to their tutorial TAs, as well as the ability of the TAs as trained observers.

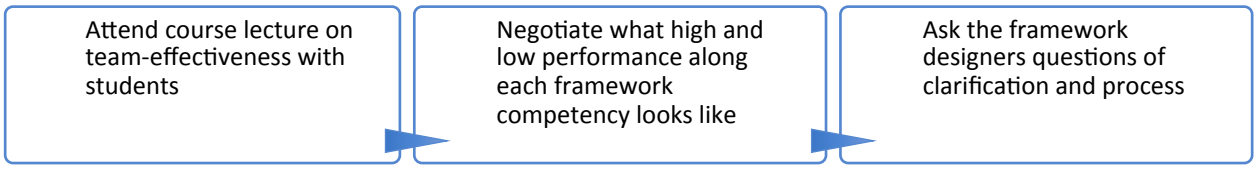
Teaching assistants were trained before providing feedback about their students according to the process shown in Figure 4-2. After attending the same lectures on team-effectiveness as the students, the TAs met with the designers of the framework for 1.5 hours to discuss how they would be using the framework to assess their students. As almost all TAs had previously been TAs in the course, the group elected to use the previous year's teams that they and the framework designer were familiar with as the case studies to calibrate their assessments. These exemplar teams were described to the two new TAs in front of the framework designer by the other five TAs, so that they could all assess the teams based on the provided descriptions. Using these case studies, the TAs negotiated what high and low performance looked like for each framework competency until their assessments of performance were in agreement. Finally, the TAs asked the framework designer any questions of clarification about what different behaviours in the framework could look like outside of the teams discussed.

Table 4-1: Demographic information of the teaching assistants (TAs) who participated in the study, their backgrounds and their respective previous affiliations with the course

	TA1	TA2	TA3	TA4	TA5	TA6	TA7
Background	Eng	Eng	Eng	Eng	Eng	Arts	Humanities
Gender	Female	Male	Male	Male	Female	Female	Female
Previous experience with the course	Student	Student	Student and TA	Student and TA	TA	TA	TA
Native English Speaker	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Note: For background, Eng represents engineering

Figure 4-2: Training sequence in which the teaching assistants participated prior to providing any assessments of student performance



TAs were asked to complete their assessments differently than were the students in two respects. First, teaching assistants completed the assessments two weeks after the students. As students provided their feedback immediately after submitting their first major team deliverable, TAs were grading at that time. Approximately two weeks after the students provided feedback, the TAs were asked to independently provide feedback for every student in their tutorial sections using the framework. As this meant that each TA had between 54 and 81 students to assess, they were given more time to complete the framework than the student teams who only had to assess three or four team members. The TAs were given two weeks to complete these assessments, meaning that some TA assessments were completed up to a month from the time when the students completed them. Second, at the training session TAs requested a 'do not know' option for each of the competencies. The TAs commented that they would prefer not to be forced to provide an incorrect assessment for a student when they felt that they were not able to assess the student along that competency.

A focus group with the TAs was conducted after the end of term to assess the framework's utility, capture data on how they perceived and used it, and their perceptions of the utility of the feedback it generated. The focus group was conducted by an experienced focus group facilitator from outside of engineering and asked the same questions as the student focus group, outlined in Appendix E.

4.5. Participation Rates

The rate of students' consent to participate in the research study was high at 77.5%, totalling 218 students from the class – 112 from the unstructured group and 106 from the framework group. Since the students had to complete the assessments of their team members for course purposes, as part of a 2% completed/not completed deliverable, there was no additional effort on the part of the students to participate in the study. Students were incentivized to allow us to use their data in the study through the opportunity to win one $100 gift card.

Of the 112 students in the unstructured group who agreed to participate in the study, the data from 105 students were used in this analysis. One team of three students was completely dysfunctional and as a result, they used their unstructured feedback as a place to vent their anger rather than provide feedback. These data were removed, as we felt that they had the potential to significantly skew the results of the depth and breadth of the unstructured feedback content. Four other students were removed, as each of them had only one other team member who consented to participate in the study, therefore eliminating the

ability to complete any feedback comparisons between peers. All 105 students provided feedback and responded to the end-of-term survey.

Of the 106 students in the framework group who agreed to participate in the study, the data from 105 students were used. One student's feedback was removed, as they were the only person in their team to consent, providing no peer feedback for this individual. Only 95 of these students responded to the end-of-term survey.

All teaching assistants in the course agreed to provide feedback on their students, as well as participate in a focus group on their use of the feedback framework. As this was additional work on top of the teaching assistants' normal course duties, teaching assistants were compensated at the university's teaching assistant pay rate for the additional hours required to complete the framework for their students. While data were obtained from the teaching assistants for 100% of the students, only the data for the 77.5% who consented to participate were used for the analysis.

Either due to the lack of grade-based incentives or to the timing after exams, student participation in the focus groups was significantly lower than expected at three students, making all findings from the student focus group informative but not significant.

4.6. Research Methods

The study was conducted as a control-condition experiment within the same class, thus allowing us to compare the effects of both modes of providing feedback within the same context. The study was approved by the University of Toronto's Research Ethics Board for Social Sciences, Humanities, and Education. This study followed a mixed methods approach that used both qualitative and quantitative inquiry. All qualitative analysis was completed in NVivo, with quantitative analyses completed in SPSS and Microsoft Excel.

Feedback that students received in response to the unstructured feedback prompt was coded according to the feedback framework (Table 3-1) to identify which competencies were discussed in common between the unstructured feedback and the framework feedback. Codes were set up for the 27 competencies, as well as for four other competency themes that emerged in the unstructured feedback but did not map directly onto the framework's competencies. A discussion of these themes is included in Appendix F. The number of occurrences of each of the competency codes was then calculated for comparison with the framework feedback and overlaps between codes investigated. Additionally, the type of feedback provided to a student (individual or team-level) as well as the content of the feedback (design related, teamwork related, etc.) was also coded.

Feedback that students received from the framework was analyzed using intra-class correlations to understand how students were using the framework for self- and peer assessments. Trends in assessment patterns were determined and compared to those of the unstructured feedback. As TAs also used the framework, TA assessments were analyzed similarly.

The end-of-term survey contained a combination of both qualitative and quantitative questions and was completed by students in both the unstructured and framework feedback groups. Qualitative questions were coded thematically and a comparison of comments across the two groups was performed. Coding themes focused on motivation (if students tried to improve, why did they do it), action (what students did to improve) and impact (what response did the feedback evoke in the students in terms of belief or commitment). Quantitative questions measured students' level of agreement with different statements regarding the feedback and were compared to identify statistically significant differences between the two groups' perceptions of their feedback.

The focus groups were audio recorded and transcribed. The transcriptions were then coded thematically to identify similarities and differences in student and TA responses, and were analyzed to determine similarities and differences.

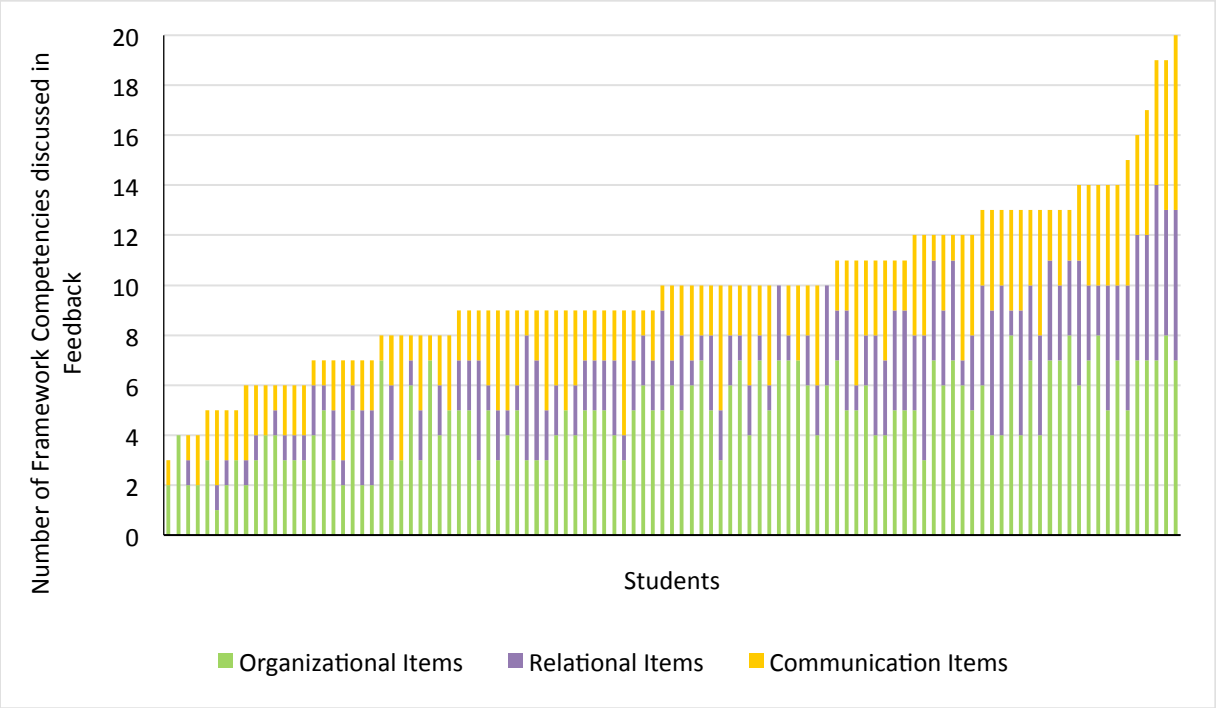
5. Differences in the Quantity, Breadth and Accuracy of Student Feedback

This section responds to the first research question – *does the framework guide students to provide a greater quantity, breadth or accuracy of feedback than the unstructured prompt* – by comparing the types of feedback received by students in the two groups.

5.1. Quantity of Team Member Effectiveness Content

Quantity was measured by the number of different team member effectiveness competencies a student received in their feedback that matched those in our feedback framework. All students in the unstructured group received feedback on at least three competencies, with a few students receiving feedback on up to 20 competencies (Figure 5-1) discussed in our framework.

Figure 5-1: Number of framework competencies discussed in feedback received by students in the unstructured feedback group in response to the prompt "Please provide feedback to yourself and your team members based on your/their team-effectiveness over the course of this project"



On average, students in the unstructured group received feedback on nine to 10 competencies. All students received feedback on at least one organizational competency, 97% of students received feedback on a communication competency, but only 88% of students received feedback on a relational competency.

Four students in the unstructured group, including three from the same team, chose to structure their feedback based on the framework's competencies presented to students during the three half-hour lectures on team-effectiveness. Even though these students chose to use the framework's competency-language to provide feedback, they did not provide feedback on all 27 competencies but primarily used the terminology of the competencies to frame their feedback.

By comparison, all students in the framework group received feedback on 27 competencies, as students were required to provide feedback on all competencies included in the framework.

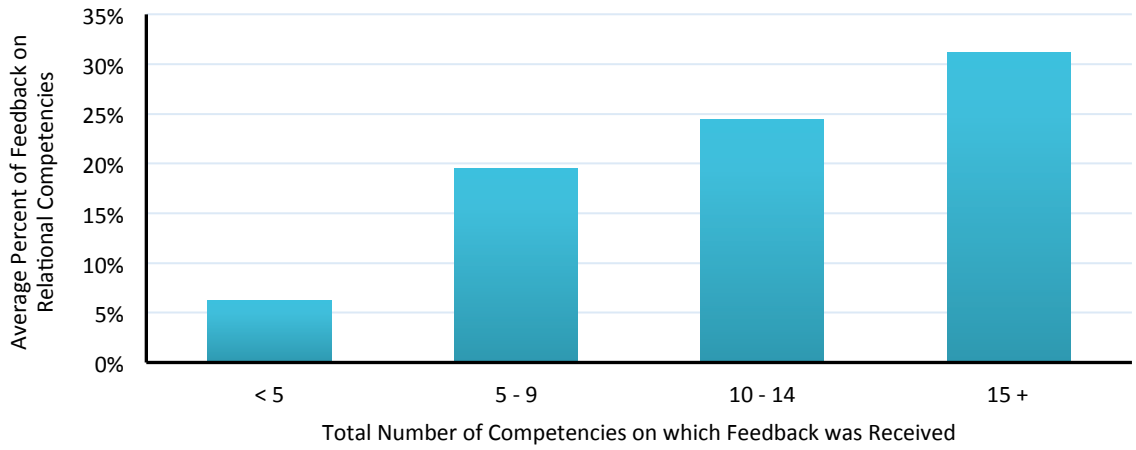
Therefore, students in the framework group received a greater quantity of team member effectiveness feedback.

5.2. Breadth of Team Member Effectiveness Content

Breadth was evaluated based on whether students received feedback covering a range of competencies or whether the comments provided feedback on few similar competencies. Competencies discussed in the feedback were grouped into the three aspects of team member effectiveness in the framework to assess the breadth of the feedback a student received.

Students in the unstructured group generally received significantly more feedback on their organizational competencies than on the other two aspects' competencies. Students in the unstructured group on average received feedback on five organizational behaviours (*SD*=1.7), two relational behaviours (*SD*=1.6) and three communication behaviours (*SD*=1.4). This corresponds to approximately 50% of a student's feedback being on organizational competencies, 22% on relational, and 28% on communication competencies. While the amount of organizational and communication content by percentage was fairly stable regardless of the number of competencies on which a student received feedback, students who received feedback on few competencies received little to no feedback on relational competencies (see Figure 5-2). Competencies that were discussed the most in feedback (over 60% of students) included *produces high quality work*, *does their fair share of the work*, *delivers their work on time* and *introduces new ideas*. Competencies that were discussed the least in feedback (less than 10% of students) included *support team rules* and *build the trust of teammates*.

Figure 5-2: Percent of relational content in an unstructured group's student's feedback based on the total amount of feedback received



Students in the framework group received feedback on 10 organizational behaviours, nine relational behaviours and eight communication behaviours. On average, students in the unstructured group received feedback on less than half as many competencies as students in the framework group.

A two-way contingency table analysis was conducted to evaluate whether there were differences between the unstructured and the framework group's perceptions of the breadth of their feedback in response to the same question on the end-of-term survey, "*I received feedback on: a broad range of topics*" (Table 5-1). The two variables used in each of these analyses were the experimental condition with two levels (unstructured, framework), and student response to the question on a Likert scale (strongly disagree, disagree, slightly disagree, neutral, slightly agree, agree and strongly agree).

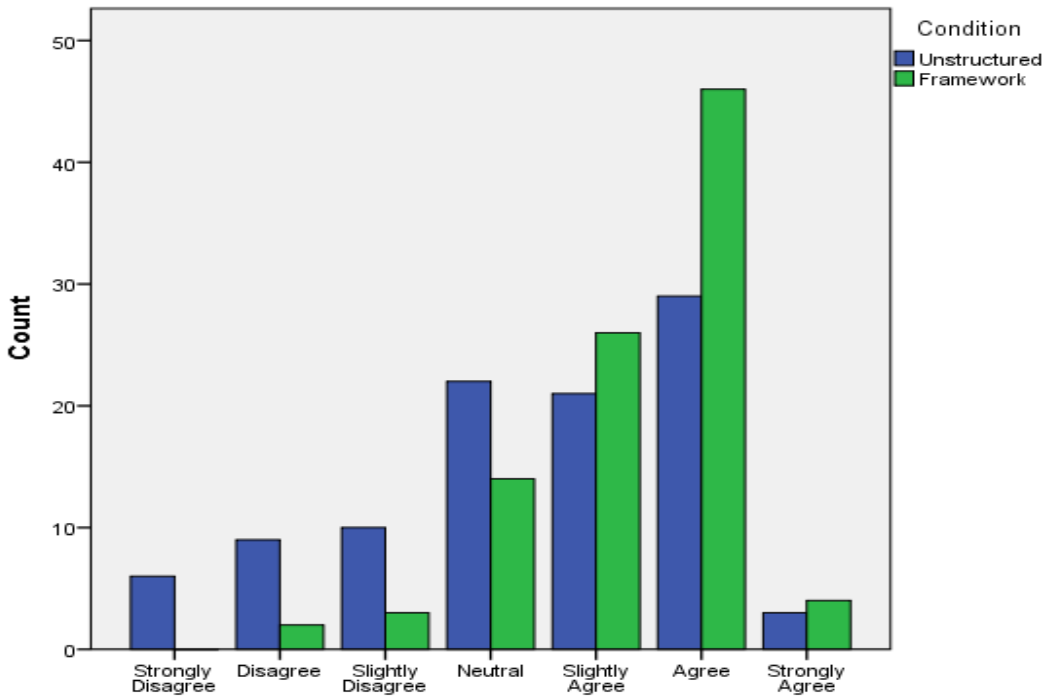
Table 5-1: Contingency table analysis comparing group agreement on receiving a broad range of topics in their feedback

Question	Pearson's χ^2	Cramer's V
I received feedback on: a broad range of topics	20.4	.32**

Note: ** $p \leq .01$

As can be seen from Table 5-1, there was a significant difference in students' perceptions of the breadth of their feedback across the two groups. Looking at the graph of the student responses to this question, it can be seen that students in the unstructured feedback group perceived that they did not receive feedback on as broad a range of topics as the framework group (Figure 5-3).

Figure 5-3: Distribution of responses to the statement "I received feedback on: a broad range of topics" by group



5.3. Other Types of Information in the Feedback

In addition to the feedback discussed above on students' team member effectiveness, a number of other items were presented in the unstructured feedback that did not support student awareness of their team member effectiveness. Two types of feedback fit this description: unrelated feedback and team-level feedback.

5.3.1. Unrelated Feedback

Some students in the unstructured group received feedback on their course-related knowledge/skills (17 students) or were provided with a description of what work their teammates perceived they contributed to the project (34 students, of which 11 received no team member effectiveness-related feedback). Non-feedback statements, generic statements (e.g., "good job") that provided no information about the students' team-effectiveness competencies or course skills, were received by nine students and provided them with no information that they could use as either a point of reference or guidance about their performance.

No students in the framework group received unrelated feedback due to the nature of the feedback framework.

5.3.2. Team-level feedback

Self-assessments in particular were not used effectively by students in the unstructured group. When students were asked to provide feedback on themselves, they often reflected on their whole team's effectiveness rather than on their own individual effectiveness. Twenty-four percent (24%) of students in the group did this, providing no specific reflection on their own performance as a team member. No students provided team-level assessments as peer feedback without discussing how the receiver contributed to the team's effectiveness – a key component that was missing from the self-assessments.

No students in the framework group received team-level feedback due to the nature of the feedback framework. As a result, no students in the framework group missed out on this opportunity to compare their self- and peer assessments.

5.4. Accuracy – Agreement between Self- and Peer Assessments

Accuracy was measured by examining the amount of unconsidered assessments provided by each group, the agreement between self- and peer assessments, as well as through student perceptions of the accuracy of their feedback as reported in their end-of-term surveys. Assessments that demonstrated no considered reflection on the part of the provider were unlikely to provide accurate assessments of the receiver's performance. Agreement in the unstructured group was measured in terms of the number of competencies discussed in common across multiple team members. Agreement in the framework group looked at

correlation between self- and peer assessments using intra-class correlation and Spearman's rank correlation.

5.4.1. Unconsidered Assessments

Unconsidered assessments for the unstructured group consisted of non-teamwork-related comments, non-feedback statements such as generic performance statements, and descriptions (rather than assessments) of a student's contribution to team deliverables. Forty-seven individual assessments were judged to be unconsidered, of which the majority were provided by eight students to each of their team members. This resulted in 45% of students in the unstructured group receiving at least one unconsidered assessment, out of a total of three or four assessments from their whole team.

Unconsidered assessments in the framework group were those in which a student gave themselves and/or their peers the same numeric assessment across all 27 competencies. Twelve assessments met these criteria, with seven of them coming from two individuals. One student provided the same assessment for all team members across all competencies, and a second student provided their peers with the same assessment across all competencies but demonstrated consideration in their own assessment. It was unclear whether the other five assessments were considered, as their assessment value did not differ substantially from the considered assessments of their other team members. Thus, seven out of 94 individuals in the framework group received unconsidered assessments (approximately 7.5%), substantially less than the 45% in the unstructured group.

The unconsidered assessments – 12 from the unstructured group and seven from the framework group – were removed before completing the analyses discussed below.

5.4.2. Agreement in the Unstructured Feedback

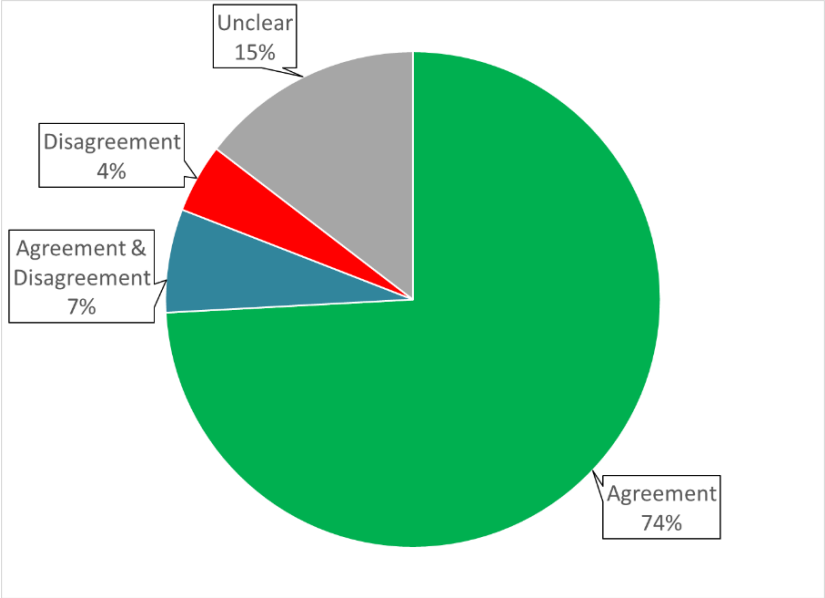
Students in the unstructured group received limited feedback that demonstrated agreement or disagreement on their performance, as their feedback contained few competencies that were discussed in common by more than one assessor. Thus, the ability to assess agreement is limited to the cases in which either the student's peers discussed at least one competency in common, or in which a student and at least one of their peers discussed at least one competency in common.

Using the 27 competencies that were coded based on the feedback framework, 14% of all competency-coded feedback in the unstructured group addressed competencies that were discussed in common by two or more peers. This resulted in 83 students receiving feedback that discussed at least one competency in common, with an agreement on the student's performance on that competency. On average, 24% of a student's feedback discussed competencies in common. The maximum percentage of feedback discussing competencies in common was 63%.

Competencies discussed in common were grouped according to whether the peers agreed on the performance level of the student, whether the peers disagreed on the performance level, or whether it was

unclear whether there was agreement or disagreement between the peer assessments. The distributions of these performance-level agreements on a per student basis can be seen below in Figure 5-4. Across the 83 students who received feedback in common, 60 students received feedback that, on average, agreed on the performance level of the student, 4 received feedback that disagreed on the performance level of the student, and 6 students received feedback that agreed on the performance level of at least one competency and disagreed on at least one other. For the remaining 13 students' competencies discussed in common, it was unclear whether there was agreement or disagreement between assessments.

Figure 5-4: Distribution of performance agreement in student feedback for competencies that were discussed in common by peer assessors



5.4.3. Agreement in the Framework Feedback

To assess the agreement of the feedback a student received in the framework group, the inter-rater reliability between peer assessors and the correlation between students' self- and peer assessments were determined. Inter-rater reliability was determined using the intra-class correlation coefficients (ICCs) for peer assessments, using a two-way random effects model for consistency. Ninety-one of the 95 students had at least two peer assessors and were considered in this analysis. Peer ICCs were then grouped according to their level of agreement between the peer assessors (Table 5-2). Thirty-one percent of the feedback showed no agreement between the peer assessors, and 8.8% showed substantial agreement between the peer assessors. The majority of assessments showed limited agreement, ranging from slight to moderate agreement. Therefore, it is likely that the feedback showed differing levels of agreement along the different aspects of the framework.

As a result of the limited agreement among peers, the level of agreement between self- and peer assessments was determined. Looking specifically into our hypothesis that the limited agreement may be aspect-specific, accuracy was determined by comparing the average self-assessments to the peer assessments across the three aspects of the framework. Using a Spearman's rank correlation (Table 5-3), we see a significant correlation between the students' self- and peer assessments in the organizational aspect. However, there is no significant correlation between students' self- and peer assessments in the relational and communication aspects. This strong correlation in self- and peer assessments along only the organizational aspect may explain the variance in agreement seen in the peer agreement discussed above.

Table 5-2: Distribution of agreement between peer assessments of a student as measured using ICCs in a two-way random effects model for consistency

Level of Agreement	Number of Students	ICC value range
No agreement	28	< 0
Slight agreement	18	0 - 0.2
Fair agreement	20	0.21 - 0.4
Moderate agreement	17	0.41 - 0.6
Substantial agreement	8	0.61 - 0.8
Perfect agreement	0	0.81 - 1

Table 5-3: Spearman's rank correlation between students' self- and peer assessments for each aspect of the framework

Aspect	Correlation
Organizational	.45**
Relational	.19
Communication	.16

Note: Students' self-assessments (n = number of competencies) and peer assessments (n = number of competencies x number of peers) were averaged separately to determine each student's self-assessed and peer-assessed level of competency for the aspect. These averages for each aspect were then correlated over all the students. ** - $p \leq .01$.

5.4.4. Perceived Accuracy

Using a two-way contingency table analysis, there was no significant difference across the two groups in their agreement with the statement "*The feedback I received described me exactly how I perceived myself*" from their end-of-term surveys.

This similarity in terms of how students perceived the accuracy of the feedback extended into students responses to the open-ended questions on the end-of-term survey. Four students from the unstructured feedback group and five students from the framework feedback group commented that they felt their feedback was fake and that it was not accurate. As this accounted for the same percentage of students in either group (5%), neither group on average perceived the feedback they received to be more or less fake than the other group. Additionally, an equal number of students (8) from each group challenged their feedback from their peers, stating that they felt that it did not reflect how they behaved in their team, or that their assessments were based on one incident and not their average behaviour.

One difference between the groups, however, was in terms of the perceived consistency of the feedback. Seven students in the framework group claimed that their feedback presented two differing perspectives on their performance and that it thus became hard to know how they were actually performing in their team. No such comments were provided by students in the unstructured feedback group.

5.5. Discussion

In addition to the differences in feedback around quantity, breadth and accuracy of team member effectiveness competencies, three other notable trends emerged in the analysis. First, students, primarily those in the unstructured feedback group, did not always receive enough feedback to provide them with a clear assessment of their effectiveness as a team member, limiting the usefulness of the feedback they received. Second, students in both groups strongly privileged organizational competencies in their feedback. Third, students in the unstructured feedback group did not always use the self-assessment prompt correctly.

5.5.1. Limited Quantity of Unstructured Feedback

Within the unstructured feedback group, students' limited ability to find agreement among their peer assessments prevented them from getting a clear picture of how they were performing as a team member. Few students had competencies discussed in common, which was limited further by the amount of unrelated feedback students generated due to the unstructured nature of the prompt. This limited quantity affected the ability of the feedback to provide the receiver with a consistent message about their performance. While there was a fair amount of agreement within the peers about students' performance when they discussed competencies in common (66% of all feedback discussed in common), there was not sufficient total feedback in common (14% of total feedback) to provide students with a sense of consistency in peer feedback. This resulted in only 9.24% of the unstructured feedback discussing competencies in common with agreement about performance. Essentially, an unstructured prompt (and how students choose to interpret it) will most likely produce peer feedback that raises few aspects in common. Students may thus perceive this feedback as random bits of information without a consistent message, making it easier to dismiss as not accurately describing them.

5.5.2. Student Privileging of Organizational Competencies

Student feedback in both the unstructured and framework groups demonstrated a strong identification with and privileging of organizational competencies. This is not surprising, as anecdotally students in engineering have a tendency to privilege work-contribution-related competencies.

This privileging was evident in the distribution of competencies discussed in the unstructured group's feedback through the strong prevalence of organizational items in the unstructured feedback (approx. 50% of all feedback was organizational). Additionally, students who received little feedback received almost all of their feedback on organizational competencies. As a result, unstructured feedback may perpetuate student privileging of organizational competencies and limit their ability to grow as effective team members by not broadening student understanding of team member effectiveness.

When looking at which items were discussed in the unstructured feedback, we find the same organizational predominance, with the exception of two competencies. The top five most discussed competencies in the unstructured feedback were predominantly organizational, with one exception. Given the privileging by engineering students of work-contribution-related competencies, the high level of response relating to the *introduce new ideas* competency (third most discussed) is likely a product of the team's project work. Students who came up with the idea that was selected to address the design project were lauded extensively for doing so in their feedback. One organizational competency that was very minimally discussed was *support team rules.* This is not alarming as in the course context, team rules were discussed as norms of practice but were not codified. As a result, students in the course would not be predisposed to reflecting on a student's role with respect to team rules.

This privileging can also be seen in the framework feedback in the correlation between student self- and peer assessments. The only significantly correlated assessments were those along the organizational aspect. While this does not affect the feedback a student receives, it does imply that students may perceive the feedback provided along the organizational aspect to be more accurate than the feedback along the other two aspects as they agree more closely with their self-assessments.

5.5.3. Limited Utility of Self-assessments in the Unstructured Group

Student self-assessments in the unstructured group were not used effectively, as many students used it as a space to reflect on their team's effectiveness rather than their effectiveness as a team member. The self-assessment was meant to act as a point of comparison for the student when receiving feedback to see how differently they, as compared to their peer team members, perceived their effectiveness. Without an individual-level self-assessment, this comparison is less effective and the feedback therefore less useful. Additionally, it does not provide the receiver with any guidance on how to improve their behaviour, as what is needed to improve the team as a whole might be different from what they need to improve as a team member. Thus, not having any self-assessment of team member effectiveness limited the usefulness of the feedback in creating a greater awareness of how the students were performing for this 23% of the unstructured group.

5.6. Conclusion and Response to Research Question

Comparing holistically the feedback received between the two groups, students in the unstructured group received feedback on substantially fewer teamwork competencies than did those in the framework group, and received their feedback primarily on organizational competencies. Students in the unstructured feedback group received significantly less feedback on relational competencies than did those in the framework group – an area known to be underprivileged in engineering students. Students in this group also received a substantial amount of their feedback on non-teamwork-related issues/skills. In terms of the accuracy of the feedback received by students in either group, the framework group received less feedback that appeared to have no considered reflection on the part of the provider than did the unstructured group. Consistency in the feedback provided by students was greater in the framework group, in part due to the set competencies on which all students had to provide feedback, but it did not appear to limit the ability to receive a diversity of feedback, as seen in students' comments on receiving different assessments from their peers.

Therefore, in response to the research question, the framework provides greater breadth, quantity and accuracy of feedback than the unstructured feedback prompt.

6. Differences in Student Perceptions of the Usefulness of their Feedback

This section responds to the second research question – *do students perceive the feedback from the framework to be more useful than unstructured feedback?* – by comparing students' received feedback and their comments on the end-of-term surveys. As students in both groups answered the same questions on the end-of-term survey, comparative analyses are possible. Three notable differences surfaced in the analysis: i) identification of strengths and weaknesses in the feedback, ii) perceived amount of actionable feedback, and iii) improvement based on feedback.

6.1. Identification of a Student's Strengths and Weaknesses

Feedback with clear strengths and weaknesses is more useful to students, as it identifies what they are good at and what they need to improve. This allows students to focus their efforts on improving certain key areas. Feedback from both groups was compared to determine which more readily identified strengths and weaknesses in feedback.

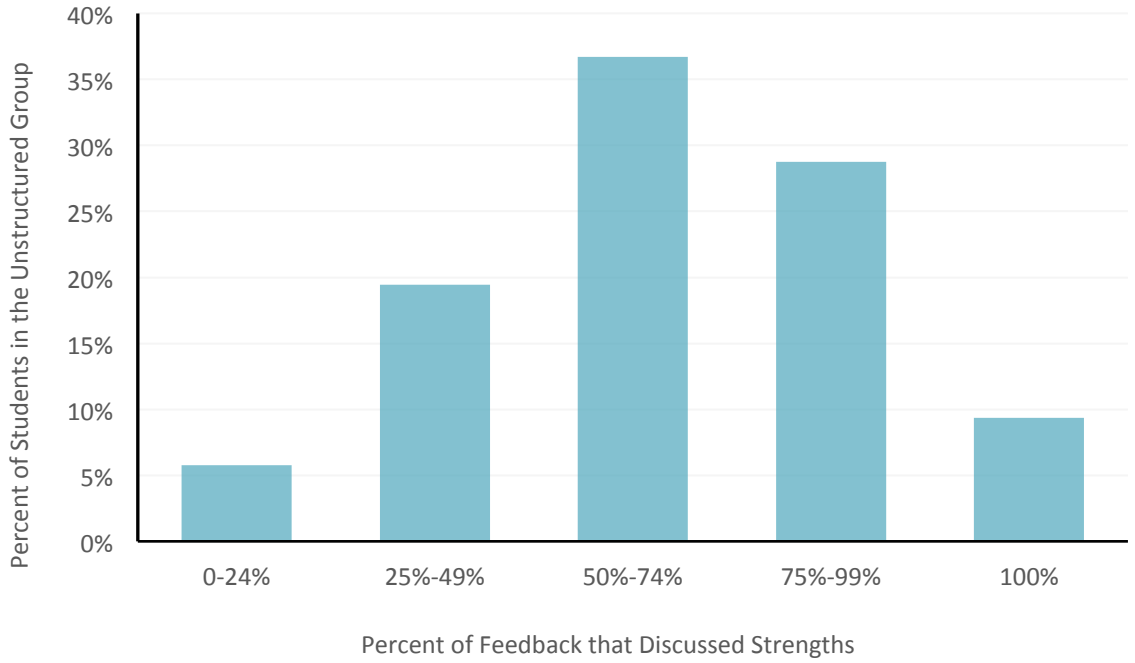
6.1.1. Unstructured Feedback

The feedback from the unstructured group was significantly less granular in terms of performance than the framework feedback. Students were either lauded on what they did well, critiqued on what they did poorly, or informed of what they did or did not do. As a result the unstructured feedback, which had already been

coded to identify team member effectiveness competencies, was coded again according to these three types of feedback.

The unstructured feedback received by students discussed the strengths of the students more commonly than it did weaknesses. Twenty-five students (24%) in the unstructured group did not receive any feedback on weaknesses or critique of their behaviour, and six students (6%) received no praise or identified strengths in their evaluations. Students within both of these groups also received neutral feedback, which simply stated what they contributed to the project. To determine the prevalence of strengths in the feedback received by students, the feedback was binned into five categories according to the percentage of the feedback that discussed strengths. As can be seen in Figure 6-1 below, 75% of the students in the unstructured group received feedback that primarily discussed their strengths as opposed to their weaknesses (opportunities for improvement).

Figure 6-1: Percent of peer-feedback received by students in the unstructured feedback group that discussed their strengths.



6.1.2. Framework Feedback

All students in the framework group had a similar number of identified strengths and weaknesses. This was a result of the sorting algorithm ranking student framework feedback, such that students' strengths and weaknesses were determined and highlighted for them. All students in the framework group had three to five competencies identified in their feedback as strengths, and three to five competencies identified as

areas for improvement. This corresponded to students having at least 11% of their feedback identifying strengths and at least another 11% identifying their areas for improvement.

To ensure that the strengths and weaknesses were not just arbitrarily identified but actually flagged competencies with which the receiver had a distinct difference in their level of performance, a Related-Samples Wilcoxon signed rank test was conducted on the difference between the average ratings of the competencies identified as strengths and the average ratings of the competencies identified as weaknesses. There was a significant difference (p = .01) in the averages of the strengths and weaknesses identified for the students, meaning that on average students were receiving feedback that demonstrated distinct differences in their performance level across the competencies.

6.1.3. Student Perception of the Strengths and Weaknesses Identified

There was a significant difference in the perception of the identification of strengths and weaknesses across the unstructured and feedback groups. While there was no significant difference between the groups as to their perception of their strengths, there was a significant difference across the two groups in terms of the identification of their weaknesses. A two-way contingency table analysis was conducted to evaluate whether there were differences between the unstructured and the framework group's perceptions of the content of their feedback in response to the same question on their end-of-term survey (Table 6-1). The two variables used in each of these analyses were the experimental condition with two levels (unstructured, framework), and student response to the question on a Likert scale (strongly disagree, disagree, slightly disagree, neutral, slightly agree, agree and strongly agree).

Table 6-1: Contingency table analysis comparing group agreement on receiving identified weaknesses in their feedback

Question	Pearson's χ^2	Cramer's V
The feedback I received was structured in such a way that I found out what my weaknesses are	12.4	.25 *

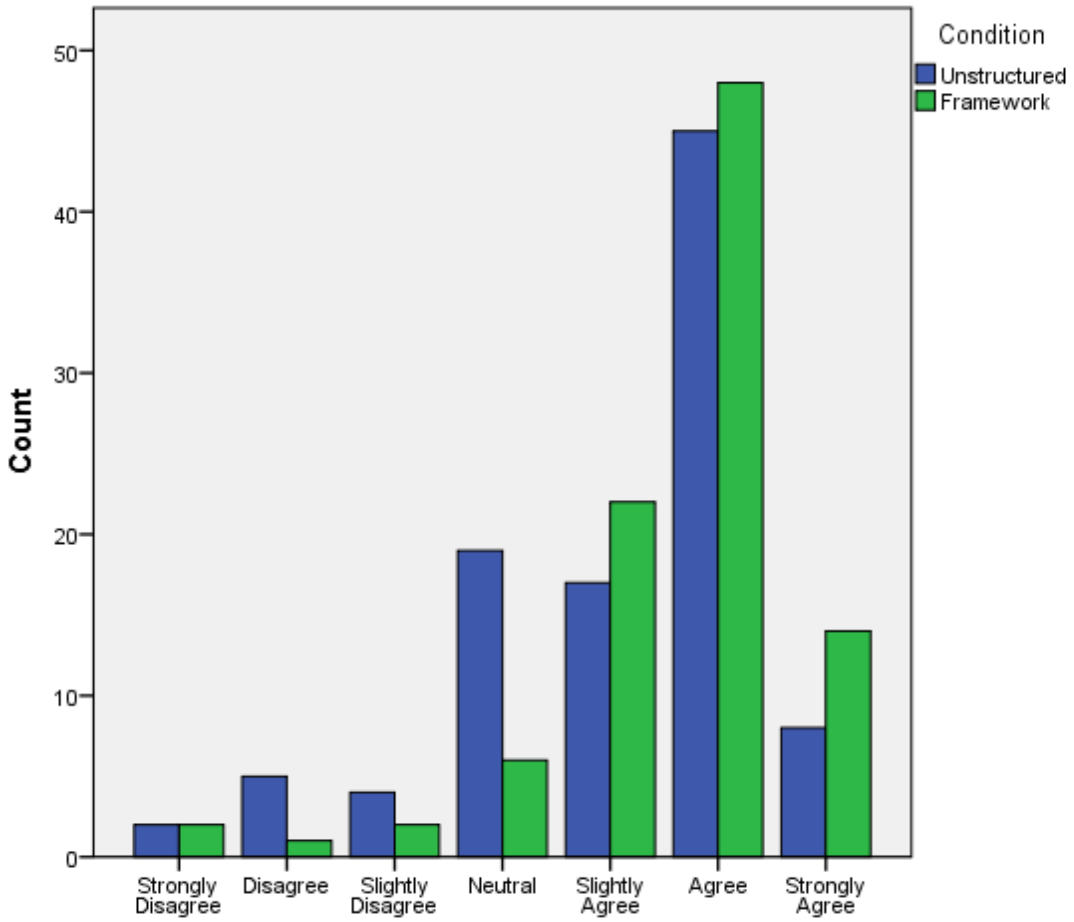
Note: * $p \leq .05$

As can be seen from the response distributions in Figure 6-2, the framework group demonstrated more overall agreement with the statement than the unstructured group, indicating that students in the unstructured group noticed the distinct lack of weaknesses provided in their feedback.

The students in the unstructured group remarked on this lack of critique of their behaviour in the open-ended questions on their end-of-term surveys. The strongest theme (greatest number of students commenting on it) among responses from students who received unstructured feedback was that they did not receive feedback on areas in which they could improve. Thirty-one students (30%) in the unstructured group commented that they wanted more critique in their feedback, compared with seven students (7%) from the framework group. The strongest theme among the students who received framework feedback was that they wanted more justification, comments or examples for the assessments they received (23%).

Therefore, students in the framework group received and perceived that they received more identifiable weaknesses on which to improve, while both groups perceived a similar number of identified strengths.

Figure 6-2: Distribution of responses to the statement "the feedback I received was structured in such a way that I found out what my weaknesses are" by group



6.2. Improvement based on Feedback

Improvement based on feedback was different across the two groups. Motivation to improve performance appeared stronger for the unstructured feedback group, due to the tone of their feedback. However, self-reported improvement along specific weaknesses identified in the feedback was higher for the framework feedback group.

The tone of the feedback appeared to have a large influence on students' motivation to improve and develop a sense of team cohesion. A two-way contingency table analysis was conducted to evaluate

whether there were differences between the unstructured and the framework group's perceptions of the tone of their feedback in response to the same question on their end-of-term survey (Table 6-2). The two variables used in the analysis were the experimental condition (unstructured, framework) and student response to the question on a three-point scale (Positive, Neutral, Negative).

Table 6-2: Contingency table analysis comparing the tone of the feedback received by the two groups

Question	Pearson's χ^2	Cramer's V
Feedback to me was phrased in a {positive, neutral, negative} tone	8.5	.21**

Note: ** $p \leq .01$

There was a significant difference in the perception of the tone of the feedback. Students in the unstructured group felt that their feedback was phrased more positively, whereas students in the framework group were mostly split as to whether their feedback was presented in a neutral or a positive tone.

Twenty-three students in the unstructured group, compared to 12 students in the framework group, commented that the feedback they received increased their commitment to their team or made them feel that their team members valued their contributions. Comments from students in the framework group seemed to be looking for this personal connection in their feedback. Students in the framework group commented that they wanted more justification, comments or examples for the assessments they received (23 students) (i.e., some textual feedback from their peers).

One difference, however, between the two groups was in the behaviours they indicated that they worked on improving during the period between receiving their initial feedback and completing the end-of-term survey. Students from the framework group cited more weaknesses that were identified in their feedback than did students in the unstructured feedback group (Figure 6-3). Both groups cited working on other improvements that were not indicated in their feedback but that they thought were equally important.

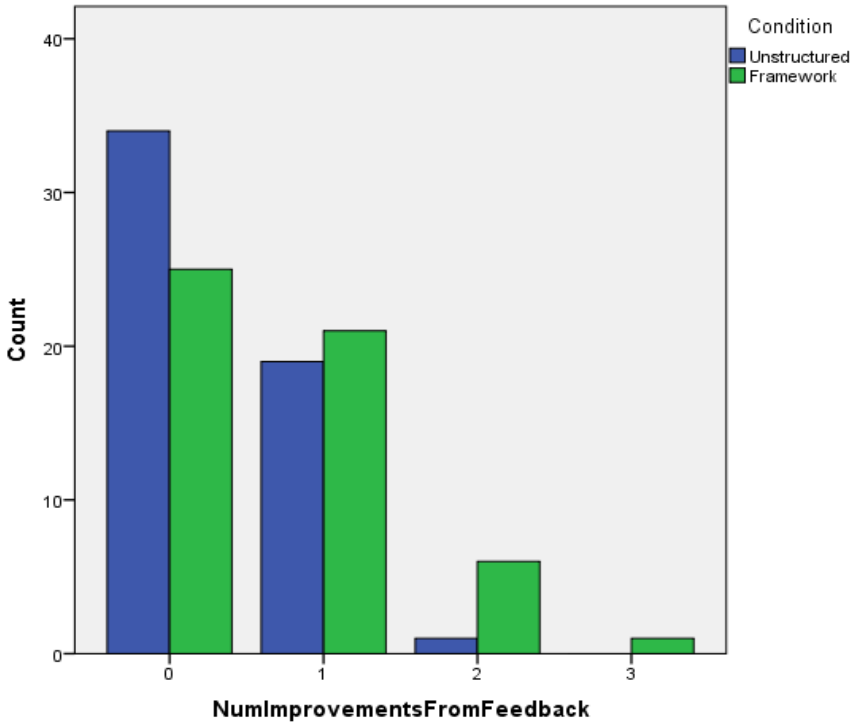
6.3. Discussion

While the tone of the textual feedback provided by students in the unstructured prompt increased student commitment to improving, the lack of identified weaknesses limited the usefulness of the feedback as a means to promote improvement and learning. Since the unstructured feedback group received feedback that was phrased in the words of their team members and mostly focused on strengths instead of weaknesses, it makes sense that students would perceive this feedback as more positive than the framework feedback received by that group. However, in a few cases where students were provided constructive criticism, the context for the feedback was provided to explain the assessment. This aspect of the unstructured feedback strengthened the utility of the feedback for these few students.

This lack of identified weaknesses may also explain why the number of weaknesses on which students identified improving was the same across both groups, but the feedback-identified weaknesses on which

students improved was less for the unstructured group. Having fewer identified weaknesses could inhibit students' ability to recall and/or consciously improve. Additionally, since the amount of competency-specific feedback was minimal (14%) for the unstructured feedback group, it is likely that even when a weakness was identified, it was written off as an anomaly if it was not also discussed by other peers.

Figure 6-3: Number of weaknesses identified in students' end-of-term surveys that they commented as improving on based on their feedback, by group



6.4. Conclusion and Response to Research Question

In response to the research question – *do students perceive the peer feedback from the framework to be more useful than unstructured feedback?* – the findings above indicate that students do find the framework feedback to be more useful.

From these analyses, it was clear that the framework guided students to identify strengths and weaknesses more clearly than the unstructured feedback prompt, as students had a prescribed set of competencies to assess in the framework. Students in the unstructured group had significantly less agreement with the statement that their feedback identified their weaknesses, and also vocalised this issue in their open-ended responses in the end-of-term survey. As a result, students receiving feedback from the framework receive a greater amount of critique and therefore were able to identify their weaknesses so that they could improve on them. This was confirmed when students in the framework group cited that they worked on improving

aspects of their team-effectiveness that were identified as weaknesses in their feedback more frequently than did students in the unstructured group.

Therefore, because the framework feedback indicates the potential to create more of the desired behavioural change articulated by team members in their peer feedback, it makes the framework feedback more useful. However, based on the impact of the unstructured feedback's tone on team cohesion, a hybrid of the framework and unstructured feedback should be investigated in the future to allow for the inclusion of textual comments in the framework.

7. Difference between Student and TA Assessments

This section responds to the third research question – *can students provide feedback similar to that of a trained observer (course teaching assistant) when using the framework?* – by comparing the assessments provided by the teaching assistants (using the framework) with the students assessments in both the unstructured and framework groups. Teaching assistants were used as trained observers because of the quantity and quality of their interaction with the students in weekly tutorial sessions. In endeavouring to assess the concurrent validity of the framework, it was not enough for students to be able to provide similar ratings to their peers using the framework. In a more absolute sense, they needed to provide ratings that were in fact a representation of the assessed student's team member effectiveness.

An investigation of teaching assistants' ability to provide feedback using the framework was provided by Sheridan et al. (2014). The key findings around how TAs used the framework were:

1. TAs provided assessments that were on average lower than those of the students.
2. The average number of competencies on which a TA provided feedback was 16 (minimum 2, maximum 27), due to their use of the 'do not know' option.
3. TAs provided a similar amount of feedback across the three aspects.

Two notable differences surfaced in this comparison of teaching assistant to student assessments: i) similarity of strengths and weaknesses identification, and ii) correlation between teaching assistant, and self- and peer assessments.

7.1. Similarity in Identification of a Student's Strengths and Weaknesses

The same algorithm used on students' framework assessments was applied to the TAs' assessments to sort out the three to five top ranked strengths and three to five bottom ranked weaknesses for each student. These TA-identified strengths and weaknesses were then compared to the strengths and weaknesses identified by students in the framework feedback group, and to the competency-coded strengths and weaknesses discussed in feedback for the unstructured feedback group.

When looking at the percentage of students in each group, the agreement between student and TA feedback in terms of identifying at least one strength or weakness in common was higher for the framework

group (86%) than for the unstructured group (69%) (Table 7-1). Additionally, while TAs and students rarely agreed on both a strength and a weakness for any given student in both groups, this occurred more frequently in the framework group (18%) than in the unstructured group (7%).

Table 7-1: Frequency of agreement between student and TA identified the same strengths and weaknesses in the unstructured and framework feedback groups

	Percent of Students in Feedback Group	
Type of Agreement	Unstructured	Framework
1 strength OR 1 weakness	69%	86%
1 strength AND 1 weakness	7%	18%

7.2. Correlation between TA and Student Framework Assessments

To determine whether the students provided similar feedback to the TAs when given the framework, the correlations between student self- and peer assessments and TA assessments were computed. As the number of competencies was substantial, and given that a number of the within-aspect competencies are highly correlated, correlation between student and TA assessments to determine the similarity of their assessments was done at the aspect-level of the framework. Correlations were computed across students' and TAs' average organizational, relational and communication assessments. These self averages and peer averages for each aspect were then compared to the TAs' average assessment of the student's competency by correlating over all the students.

A Spearman's rank correlation was conducted between the aspect-level TA and peer assessments (Table 7-2) and TA and self-assessments (Table 7-3) for all students in the framework group. When using the framework, the peer assessments and TA assessments of individual students were significantly correlated across all aspects. However, the self-assessments were only significantly correlated with the TA assessments for the organizational aspect – the aspect that students in both the framework and unstructured group were more comfortable assessing. This is the same trend that we saw in Table 5-3, where student self- and peer assessments were only significantly correlated along the organizational aspect as well.

Table 7-2: Spearman's rank correlation between students' peer assessments and TAs' assessments of each student along the three aspects of the framework

Aspect	Spearman's Rank Correlation
Organizational	.35**
Relational	.27**
Communication	.24*

Note: Students' peer assessments (n = number of competencies x number of peers) were averaged separately to determine each student's peer-assessed level of competency for the aspect. * $p \leq .05$, ** $p \leq .01$

Table 7-3: Spearman's rank correlation between students' self-assessments and TAs' assessments of each student along the three aspects of the framework

Aspect	Spearman's Rank Correlation
Organizational	.26**
Relational	.00
Communication	.16

Note: Students' self-assessments (n = number of competencies) were averaged separately to determine each student's self-assessed level of competency for the aspect. ** $p \leq .01$

7.3. Discussion

In addition to the differences in teaching assistant and student assessments, three other notable trends emerged in the analysis. First, teaching assistants were hesitant about the assessment context. Second, the quantity of unstructured feedback limits the applicability of the findings around agreement. Third, student self-assessments by aspect demonstrate a different privileging than does peer assessments.

7.3.1. Teaching Assistant Hesitancy around the Assessment Context

The objective of this part of the study was initially to use TAs as the gold standard against which to compare students' feedback to determine accuracy. While the results of the assessments provided corroboration that students can provide peer assessments that are similar to those of their TAs, the teaching assistants were hesitant to be considered a gold standard. TAs are the front-line support for team issues and guidance, which requires them to be attuned to how the teams are functioning. However, they commented that they were more tuned in to the teams than to the individuals within the teams themselves. As a result, they were more comfortable and more confident assessing the teams than the individual students.

Part of the TAs' hesitation may be related to the size of the feedback framework (27 competencies), but based on TA comments from the focus groups we hypothesize that most of the discomfort is due to the cognitive load on the teaching assistants during tutorials. Since the TAs need to provide different scaffolding to each of the teams to support the development of their design and communication skills, their primary mode of observation is during these interactions. In these conversations, the TAs need to assess the current state of the team's design work and what scaffolding (in the form of advice, resources or coaching) is necessary to move the team's design work forward. As a result, this substantially limits their ability simply to observe the students as they work in their teams and biases their assessments against those students with whom they do not frequently interact in the teams. However, given the significant correlations between peer and TA assessments at the aspect level, reducing the TA feedback framework to substantially fewer competencies or to an aspect-level framework may strengthen these correlations, as TAs would then be able to provide more confident and hopefully more accurate assessments of their students.

7.3.2. Limitations on the Agreement for Unstructured Assessments

As discussed in Table 7-1, strengths and weaknesses identified by TAs agreed strongly with students assessments from the framework group. However, it must be noted that this finding is limited by the quantity of unstructured team member effectiveness feedback. As the quantity of unstructured feedback was substantially less than that of the framework feedback, there was a lower likelihood of a student having chosen to discuss a strength or weakness identified by the TA. Since all students in the framework group commented on all framework competencies, there was a greater likelihood of agreement between this group and the TAs. Therefore, while the agreement of the TAs' assessments with the framework group is greater, there are not sufficient data from the unstructured group to claim any difference in agreement between the two groups.

7.3.3. Aspect-level Discrepancies in Assessment

As discussed in Table 7-2, TA-peer assessments were significantly correlated along all aspects of the framework, while self-assessments were only correlated with TA assessments along the organizational aspect. This finding supports the discussion in Section 0 that students in this class privilege organizational competencies and are able to self-assess them more accurately than those in the other aspects.

TA-self-assessment correlation for the relational aspect was .00, which demonstrated no correlation between TA and self-assessments at all. This may be a product of the teaching assistants' lack of comfort with the relational competencies as well. The teaching assistants commented in their focus group that the relational aspect was the hardest to assess. However, there were no clear trends across the three aspects with respect to the number of competencies on which they provided feedback (due to the 'do not know' option). This may also be a product of the previously discussed privileging, as students in engineering are not traditionally comfortable with relational competencies. This was seen also in the unstructured feedback, in which less than 20% of the competencies discussed were relational.

Therefore, while the framework does provide a substantial basis for students to give feedback to their team members that is similar to that of their TAs, students had difficulty self-assessing their abilities in regards to relational and communication competencies. This inability to self-assess accurately speaks further to the value of using peer feedback as a teaching tool to help students develop their self-awareness of and competency in team member effectiveness.

7.4. Conclusion and Response to Research Question

Comparing the feedback received between the two groups and the teaching assistants, we can claim that student peer assessors can be used for intra-team evaluations more effectively than teaching assistants. Agreement between teaching assistant and student feedback in terms of the usefulness of the feedback (identified strengths and weaknesses) is greater when students use the framework than when they use an unstructured prompt. Additionally, student peer assessments are significantly correlated with teaching assistant assessments across all three aspects of the framework. Student peers can provide feedback similar

to that of trained TA observers. The same discrepancy between self- and peer assessments was seen with self- and TA assessments, further corroborating that peer assessments provide similar information with which students can broaden their awareness of their teamwork skills.

Therefore, in response to the research question, the framework guides students (as peers) to provide feedback similar to that of trained observers (teaching assistants).

8. Student and TA Feedback on the Usability of the Framework

This section responds to the fourth research question – *is the framework accessible (e.g., jargon, descriptions, levels of competency) to students and trained observers?* – by discussing four prominent themes around the usability of the framework, as discussed by students and TAs in their end-of-term focus groups. These four themes are: i) accessibility of the competencies, ii) accessibility of the rating scale, iii) presentation of the feedback, and iv) missing content.

8.1. Holistic Assessment of Framework Accessibility

The language of the feedback framework received no complaints or comments from either the students or the teaching assistants. Students were particularly happy that there was no "engineering jargon" in the framework, which they felt enhanced its accessibility. Students in the focus group thought that the framework was "thorough." They described the framework as "pertinent" and "relevant," and said that it "touched on more detailed elements of teams than [surveys] usually touch on." Students also commented that the framework was an effective learning tool in and of itself, and remarked that completing the framework at the mid-point of their team projects allowed them to reflect on where they were or were not doing well and reminded them of what competencies effective teamwork required.

Both students and TAs liked that the assessments were grouped into the three aspects of team-effectiveness. Both groups commented that they found this helped them get into a certain headspace and evaluate a specific aspect of the experience. This simplified thinking and helped to clarify any sources of confusion in the questions. One student commented that they would have preferred the competencies to have been presented as questions rather than statements, as it would "allow [the student] to focus more on what [they are] grading [their team member] on." They did not elaborate on why they thought this would be the case.

Students commented that instead of rating each team member along the competencies sequentially, they wanted to provide feedback for each team member along the competencies simultaneously – the students wanted to complete relative rankings of their team members. The students felt that if they were rating all their team members at once, they would be able to provide more accurate assessments, and as a result of having fewer (perceived) questions to complete they would provide more accurate assessments. Students commented that completing the entire framework for each team member sequentially resulted in fatigue by the end of the survey, such that they were not fully reading the competency or rating scale and may not have been as accurate for the last team member assessed.

8.2. Accessibility of the Framework's Competencies

Students in the focus group did not comment specifically on the list of competencies, but the TAs did feel that some competencies could not be assessed at an individual level. For competencies like *produce high quality work*, the TAs would only be familiar with team performance and could not rate team members individually. TAs also felt that the competencies *display dedication and determination, track team progress vs. project timeline, demonstrate accountability, exchange information in a timely manner, raise contentious issues in a constructive way* and *solicit input before proceeding* were all poorly worded, as they were not able to assess these competencies through observation during the tutorials. The TAs saw these competencies as the product of a series or sequence of behaviours that combine together. Given that they were not participating in the teams or observing them all the time, they were only able to see some of the behaviours, not all of the sequence, that combine to create the competency. They commented that they would have preferred to mark off specific individual behaviours they saw rather than trying to deduce the students' performance along the competency.

8.3 Accessibility of the Competencies' Behaviourally Anchored Rating Scales

Comments regarding the accessibility of the behaviourally anchored rating scales (BARS) for each competency centred around three different aspects of the BARS: their descriptions, their sequencing and the number of rating points.

8.3.1. BARS Descriptions

Students found the BARS descriptions long. Given that they had to provide feedback to each team member individually, they mentioned that they would stop reading them after a while. The students proposed having a longer description of the competency and a few words briefly describing each of the different rating points. However, this would not provide enough information to distinguish clearly the different levels of performance for the students.

TAs felt that the descriptions for the BARS did not outline behaviours that they could clearly observe in tutorials without being a member of the team. TAs felt that in their assessment context, the BARS required deduction rather than observation, which made them uncomfortable when answering. In particular, they felt that the descriptions were not representative of the personality types of their students. TAs perceived that the descriptions privileged extroversion to the detriment of introverted students, especially with respect to the communication aspect, as the only way they could assess students *introducing new ideas* or *openly expressing opinions* in tutorials was verbally. Additionally, TAs felt that when they completed the framework for teams they perceived as strong, they found them only to be performing satisfactorily based on the BARS descriptions. TAs found in these cases that the BARS descriptions did not map to their perceptions of the team's functioning.

8.3.2. Sequencing of the BARS

Students did not express any concerns around the sequencing of the descriptions along the BARS, but the TAs were unsure of the continuums along which the BARS were developed for each competency. For competency R12 – "*motivate others on the team to do their best*" – TAs found that the examples that were provided in the descriptions of the BARS in place of the word "motivate" sometimes caused there to appear as though there were two continuums along which the competency was being assessed. TAs found this dual continuum to occur most prominently in the relational aspect, where they would agree with one half of a description but not the other half. In this situation, TAs were unsure of how to rate the student and would most frequently pick the lower rating option. For the relational aspect, the TAs commented that they found the in-between options to be largely irrelevant, as the scale was more nominal than ordinal.

TAs felt that the low end of the BARS was not sequenced appropriately, as they felt that it was better to do something wrong than to avoid doing anything at all. In particular, with respect to competency R13 – "*raise contentious issues in a constructive way*" – they felt that the descriptions for levels 1 and 3 on the scale should be flipped, so that "avoided contentious issues" would be the description for level 1 and "raised contentious issues in a destructive manner" would be the description for level 3, as this organization would show a progression of engagement with the team that better maps to the conceptual framework of the BARS. On the low end of the scale as well, the TAs commented that they felt that the language used often implied that the students would be acting maliciously or demonstrating a "self-interest in the destruction of the team," which did not map on to their understanding of the students. This was not the intention of the framework designers, who meant to communicate lack of engagement on the lowest end of the scale and self-centredness at the mid-range of the scale. TAs found that the language around doing something to "serve one's own purposes" appeared malicious, whereas the framework designers did not intend it that way.

Students found the high end of the BARS to have vague and over-arching competency wording, instead of the collaborative orientation intended by the framework designers. The students commented that language such as that in R19 – "*Collaborated with others in a manner which promoted openness and understanding among team members*" – was open to interpretation and therefore would be assessed differently by everyone. The students in the focus group commented that a description that "narrows the answer more" and specifies a manner of promoting openness would ensure that the framework had a single interpretation.

8.3.3. Number of Points along the BARS

Two students in the focus group and the TAs believed that there were too many rating options for each competency. Having descriptions on only four of the seven points on the scale created additional confusion, as they were not sure how to interpret the middle options. The majority of focus group participants recommended eliminating these undescribed rating points as a way of reducing the number of points along the scale and making the rating scale clearer.

8.4. Presentation of Feedback

Students in the focus group were asked about the presentation of the feedback (see example provided in Appendix B) to determine if there were ways to make the feedback easier to read and understand. No markedly negative comments were provided about the presentation of the feedback, nor any comments suggesting that students were unable to interpret the feedback or identify their strengths and weaknesses from it. Students in the focus group commented that they liked the colour coding, which allowed them to know "on instant glance" what their strengths and weaknesses were. However, while having the strengths and weaknesses visually separable from the other competencies was a plus, a different set of colours was recommended by a colour-blind student, who commented that it was difficult to distinguish between the green and yellow that demarcated strengths from average competencies.

8.5. Missing Content

Finally, to ensure that the framework addressed students' feedback needs, students were asked if there was anything they perceived as missing from their feedback. Comments from the focus group participants were combined with students' responses to question 9 of the end-of-term survey. The strongest theme among the responses was that of a desire for comments or examples from the feedback providers – something that could guide the receiver to know exactly how or why they received the numeric assessments on the framework and how their team members wish them to improve their performance. Students in the focus group saw this as a way to justify high or low assessments, which was also a theme discussed by the TAs. The TAs were more inclined to have an optional justification box after each question in order to justify their assessment and/or their confidence in their assessment when they felt it was necessary.

Students, on the other hand, simply commented that they wanted some justification but indicated no preference between it being per competency or per student. This idea was supported by all students at the focus group as well as by 23% of the students in the framework group, who commented on this in their end-of-term survey. They wanted to know where they excelled, with specific examples, and to be able to learn by knowing why they were assessed the way they were.

8.6. Discussion

The most common theme that surfaced unintentionally in the focus groups was how the feedback could be leveraged differently in future classes to better facilitate student improvement. Students suggested some other ways to represent the feedback, including as an average of team members' assessments, by showing their performance relative to their team's or their class' average performance, or by integrating the performance level descriptions into the feedback instead of providing it as a separate handout with the feedback. Given engineering students' competitive nature, this might appeal to their desire to be the best in the class.

By comparison, TAs provided more feedback about the accessibility of the framework as an assessment instrument. This is likely due to their familiarity with grading rubrics as a result of their teaching work.

One recommendation for improving the effectiveness of the feedback from the focus groups, and something that was commented on by some students on their end-of-term surveys, was to add a debrief on the feedback during tutorial time with the team as a whole. This student recommended the inclusion of a structured discussion with the team about their weaknesses to encourage everyone to work more effectively together. This could be designed as a walk-through or scripted exercise that the students run themselves, with the TAs available for support as needed. This would allow each team to receive its own personalized feedback within the structure of the large class. This activity could guide the students through specific exercises or activities for the team to explore areas where they all are weak, or to look at new ways in which they can leverage the strengths of the team members. Having the different resources listed separately from the feedback online did not appear to be a detriment, though it was not commented on positively in the survey or focus groups either. A more intentional way to integrate the feedback and methods of improvement – either in person or online – needs to be conceived.

8.7. Conclusion and Response to Research Question

Combining student feedback about the accessibility of the framework from their end-of-term survey, and student and TA feedback from the focus groups at the end of the term, we conclude that the framework is accessible as a feedback guide for intra-team use. For the TAs, the framework is too long and difficult to complete given their interactions with the students and the cognitive load of their teaching responsibilities. Both groups identified a number of ways in which the framework could be made more relevant and intuitive.

9. Conclusions and Future Work

This project evaluated the effectiveness of a feedback framework as a foundation for a new web-based tool that provides students with structured feedback from teammates, along with personalized exercises and actionable strategies that guide targeted learning in the areas thereby identified. Specifically, the study investigated whether the feedback framework, when used for intra-team self and peer feedback, would increase a student's ability to learn about and improve their team-effectiveness.

When feedback from the framework is compared to feedback gathered using an unstructured prompt, we find that feedback from the framework has a greater ability to increase student self-awareness and provide them with information that can guide them in developing team-effectiveness competencies. Students who used the framework had significantly different strengths and weaknesses reliably identified for them, and they received consistent assessments of their performance from their team members. Students in the unstructured feedback group had less consistent feedback that more often discussed strengths alone than both strengths and weaknesses. Additionally, as students in the unstructured feedback group discussed whichever competencies they felt were important, the feedback is less targeted and focuses on how the team perceived the student's performance.

Students in the feedback group had significantly greater agreement with the statements that they received feedback on a broader range of topics and on their weaknesses than did students in the unstructured group.

While students in the unstructured group did not have these same benefits, they did comment that the textual feedback made them feel more committed to their team, as it demonstrated that their team members had an interest in their development and in that of the team as a whole. Based on this benefit and on requests from students in the feedback group for examples and comments in their feedback, we believe that a hybrid combining the feedback framework and some textual feedback would be ideal.

This document addresses an important component of a larger, ongoing project looking at the effectiveness of our feedback framework to facilitate feedback that guides students to learn about and develop their competence as effective team members. We have already incorporated much of the learning achieved through this study into our web-based tool. For example, the number of competencies being addressed has been substantially reduced and an option for holistic freeform feedback incorporated into the framework. Further, as students often work in diverse teams, we are in the process of investigating how diversity of personality characteristics in teams affects the feedback a student receives from their team. This investigation is currently in analysis and will be published at a later date.

References

Bushe, G. R., & Coetzer, G. (1995). Appreciative Inquiry as a Team-Development Intervention: A Controlled Experiment. *Journal of Applied Behavioral Science, 31*(1), 13-30.

Canadian Engineering Accreditation Board (2010). *Accreditation Criteria and Procedures.* Retrieved from http://engineerscanada.ca/e/files/Accreditation_Criteria_Procedures_2010.pdf

Lencioni, P. M. (2002). *The Five Dysfunctions of a Team: A Leadership Fable.* San Francisco, CA: Jossey-Bass.

Lingard, R. W. (2010). Teaching and Assessing Teamwork Skills in Engineering and Computer Science. *Journal of Systemics, Cybernetics and Informatics, 8*(1), 34-37.

Loughborough University (2009). Retrieved from http://www.webpa.ac.uk/

Loughry, M. L., Ohland, M. W., & Moore, D. D. (2007). Development of a Theory-based Assessment of Team member Effectiveness. *Educational and Psychological Measurement, 67*(3), 505-524.

Maxwell, J. (2011, July). Team Effectiveness Questionaire. Retrieved from http://www.firststepstraining.com/resources/teq/index.htm

Moore, T. J., Diefes-Dux, H., & Imbrie, P. K. (2006). Assessment of Team Effectiveness During Complex Mathematical Modeling Tasks. 36th ASEE/IEEE Frontiers in Education Conference. San Diego, CA.

Sheridan, P. K., Phillips, J., El Gammal, L., Evans, G., & Reeve, D. A. (2013). Team Effectiveness Inventory for Guided Reflection and Feedback. American Society of Engineering Education Annual Conference and Exposition. Atlanta, Georgia.

Sheridan, P. K., Reeve, D., & Evans, G. (2014). Understanding Teaching Assistants' Assessment of Individual Teamwork Performance. American Society for Engineering Education Annual Conference and Exhibition. Indianapolis, IN.

Tuckman, B. W. (1965). Developmental sequence in small groups. *Psychological Bulletin, 63*(6), 384-399.

Higher Education Quality Council of Ontario

An agency of the Government of Ontario