

**MEASURING THE INFLUENCE OF TEAM FUNCTIONING ON DESIGN PROJECT OUTCOMES**

Peter M. Ostafichuk, Carol Naylor, and Markus Fengler
University of British Columbia
ostafichuk@mech.ubc.ca

**Abstract** The influence of team function on design project outcomes was examined in this study. Team function was considered across six key dimensions, including unity, communication, distribution of responsibility, problem solving, conflict management, and team self-evaluation. Three different methods were used to quantify team function: a survey in which students self-rated their team’s function, a comparison of performance on quizzes first performed individually and then as a team (as a measure of the degree of communication, problem solving, and unity), and an analysis of the differentiation in inter-team peer evaluation scores (as a measure of distribution of responsibility, conflict management, and unity). Design project outcomes were measured as a composite of grades from competition prototypes, written reports, oral and poster presentations, and several other deliverables. These scores were normalized to remove year-to-year variability. Statistically significant relationships between each measure of team function and design project outcomes were observed.

For each dimension of team function, teams with high average self-rating on the survey also had 4% to 6% higher normalized design project scores compared to those with low self-ratings. On the quizzes, teams that were more likely to answer a question incorrectly when one or more members knew the correct answer (suggesting a lack of communication, unequal input to problem solving, or reduced team unity) also received lower normalized design project scores by as much as 4%. The full relationship between this metric and project outcome was more complicated though, as the teams least likely to answer incorrectly when some members had the correct answer performed below average on the projects. Lastly, a trend of decreasing composite project score was correlated with increasing inter-team differences in peer evaluation scores (suggesting unequal distribution of responsibility, increased conflict, or reduced team unity). Interestingly, teams that did not differentiate peer evaluation scores at all (i.e. each team member received the same peer evaluation score ‘no matter what’) had project scores 7% lower on average than teams with a small non-zero differentiation in peer evaluation scores. Taken together, the results of this study support the hypothesis that team function plays an important role in project outcomes, contributing better than half a letter grade difference.

**Keywords:** Team dynamics, measuring team function, peer evaluation, Team-Based Learning, TBL, design project outcomes.

1. INTRODUCTION

Team work is an essential element of engineering and is increasingly expected in higher education by students, employers, and accrediting bodies [1],[3]. In the course considered in this study, extensive effort has gone into promoting effective team function, including through careful team formation by the instructors; required (and tested) readings on team development, personality type, and conflict management; team dynamics workshops run by trained facilitators; ongoing teaching assistant oversight of team function; and regular peer evaluation. This raises the following question: does improved team function translate into improved design project outcomes? And if so, to what degree?

This study examines team function on six general criteria: unity, communication, distribution of responsibility, problem solving, conflict management, and self-evaluation of the team’s performance. Three approaches are used in an attempt to quantify the criteria for teams in a second year mechanical engineering design project course. In total, cohorts from nine academic years, 2005W to 2013W, with over 1000 students and 180 project teams are considered in this study. Design project deliverables – including graded performance in course design competitions, written reports, oral and poster presentations, and other deliverables – are used as a composite measure of project outcomes.

The objective of this study was to determine if a relationship between team function and design project outcomes could be observed. It was expected that improved team function would correlate with improved design project outcomes. This builds in part on previous studies that examined student personality type influences on team function [13] and peer evaluation [12].
The sections below first outline the course context for the study, with details on course format, team formation, peer evaluation, and team dynamics instruction. The methodology for the study is then presented, outlining the assessment of team function by three different methods and the relationship to design project outcomes. The results and discussion section presents the data collected and their interpretation, and is followed by conclusions that summarize the work highlight implications for teams in design engineering education.

2. COURSE CONTEXT

This study was conducted in the Department of Mechanical Engineering at the University of British Columbia (UBC) in a second-year mechanical design course (MECH 223). The course is part of the integrated Mech 2 Program introduced in 2004 [8]. The typical course enrollment is 115-125 students and nine cohorts from 2005W to 2013W were considered. The course is delivered using the Team-Based Learning (TBL) approach [7] (details on the course-specific TBL implementation can be found in [5],[6], and [9]). All students attend a common lecture section (i.e. there are approximately 120 students in the classroom at one time) and they are split into four sections for other activities, such as tutorials, team meetings with a teaching assistant, computer labs, and so on. The MECH 223 course is atypical in several respects: first, it is a full-time course that runs for seven weeks (students do not take other courses at the same time); second, the course is split into two parts (four weeks in January and three weeks in April, each with a separate major design project); and, third, the course is large in scope at seven credits (a typical course at UBC is three credits).

Following recommended practice, teams of six to seven students (20 teams in the course) were instructor-formed [2] in order to maximize diversity [4],[15] and to minimize previously established subgroups [7]. Prior to the course, students completed an abbreviated version of the Myers-Briggs Type Indicator (MBTI) online through the TypeFocus tool (http://www.typefocus.com). A mandatory course intake questionnaire then collected each student’s MBTI preferences as well as self-reported ability with hand skills, software skills, communication skills, and team skills. The above information was combined with GPA from previous courses to form teams that were heterogeneous across all personality, skill, and GPA criteria. Each team possessed at least two members with each MBTI preference with the exception of the Feeling preference; in most years there were not enough students reporting a preference for Feeling to uniformly distribute them across all teams in light of other team formation constraints. The same project teams were maintained for the course duration, including the January and April sessions.

As part of course requirements, students completed six mandatory peer evaluations using the iPeer online software tool (http://ipeer.ctl.ubc.ca). Three evaluations were completed using a point-based evaluation in which each student distributed an average of 100 points per other team member (students could reward above average performance with a score above 100, but that required lowering the evaluation scores, on average, for other team members). The other three evaluations were completed using a rubric-based approach in which students evaluated their teammates on four criteria (communication, initiative, responsibility, and professional behaviour) across four levels of mastery (below expectations, marginal, meets expectations, and exceeds expectations). The raw numerical scores for the rubric-based evaluation were normalized on an evaluator-by-evaluator basis to an average of 100 to match the point-based evaluations. Both forms of evaluation required the students to provide comments to justify the scores they assigned. A comparison of each of the two methods was previously reported in [12]. Half of the teams used the point-based evaluation for the first project and rubric-based for the second, and the other half of the teams used the same evaluations in reverse order. After each evaluation, students received aggregate scores and anonymous, randomly ordered comments from their teammates. The average peer evaluation score for each student at the end of the term was multiplied against their team’s net grade in order to determine the individual portion of the team grade recorded for that student.

Lastly, as part of the design course, students were introduced to team dynamics in multiple ways. Students completed a short textbook reading (8 pages from [11]) at the start of the course. The reading covered the stages of team development, the MBTI and its implications for team work, and strategies for conflict management. At the start of each project they also participated in a team dynamics workshop (one in January and one in April). The first workshop focused on understanding the MBTI and developing strategies for working as a team based on the distribution of types on each team. The second workshop focused on debriefing the experiences from the first project and then developing strategies for giving and receiving feedback and managing conflict. Further information on the team dynamics workshops is available in [10].

3. METHODOLOGY

Team function was measured in three different ways. First, a compulsory project exit survey that asked students to rate their own and their team’s performance in six criteria was administered each year. Second, responses on identical quizzes completed first individually and then teams were compared to measure the degree to which team decision making was based on the knowledge
possessed by the individual team members. Third, average deviations in peer evaluation score across the team were considered as a measure of the teams’ tendencies to acknowledge and reward/penalize differences in contribution. All three measures of team function were compared against a composite grade of deliverables for two projects. The graded deliverables changed slightly over the nine years of the course considered in this study, but nominally were as shown in Table 1 for each project (i.e. two of each item per year). The percentages shown at the end of each description represent the approximate grade weighting towards the composite project grade. The composite project grade represented 35% of the total course grade.

Table 1: Graded Design Project Deliverables

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Written reports</td>
<td>Formal written recommendation reports outlining the design process and decision making for the competition prototypes (37%)</td>
</tr>
<tr>
<td>Competition prototypes</td>
<td>Comprehensive physical prototypes used for course-wide design competition (29%)</td>
</tr>
<tr>
<td>Oral presentations</td>
<td>Formal oral presentations outlining the detailed process and decision making for the competition prototypes (17%)</td>
</tr>
<tr>
<td>Poster presentations</td>
<td>Formal poster presentations outlining team strategy and general process for the competition prototypes (11%)</td>
</tr>
<tr>
<td>Prototype demonstrations</td>
<td>Mid-project focused, physical prototypes demonstrating progress in a particular aspect of the project (3%)</td>
</tr>
<tr>
<td>Logbooks</td>
<td>Team logbooks documenting work through project (3%)</td>
</tr>
</tbody>
</table>

The methodology to measure team function in each of the three approaches (exit survey, quiz performance, and peer evaluations) are described further in the next section.

3.1. Exit Survey to Measure Team Function

At the end of the first project, in January, students were given a compulsory project exit survey to be completed online. The survey start date was roughly one week after the project completion, and the end date was roughly two weeks later. Completion of the survey carried the grade weight of one tutorial exercise (approximately 0.5% of the course grade). The survey results formed the basis of the second team dynamics workshop in April, at the start of the second half of the course and second project. Survey results were anonymous, and only aggregate results and representative, anonymous comments were shared with the class. Among other things, the survey asked students to rate their team’s performance in six criteria using a 10-point rating scale (1 = very poor to 10 = very good). The six criteria were adapted from [14] and are included below, as worded in the survey:

- **Unity**: Unity exists when the team creates a strong common purpose, sticks together, and people support one another.
- **Communication**: Effective communication is open and two-way, with discussions that are in-depth with careful listening.
- **Distribution of responsibility**: Effective distribution of responsibility occurs when all members have equal say and tasks are fairly distributed and carried out by all members.
- **Problem solving**: For effective problem solving, everyone’s ideas are used to solve problems and contribute to the project’s final plan.
- **Conflict management**: In effective conflict management, differences in opinion are acknowledged and a sound and fair resolution is reached through discussion.
- **Group self-evaluation**: Effective group self-evaluation occurs when the group periodically takes time to critique, discuss, and improve its performance.

It was hypothesized that teams with a higher average team rating would produce superior results and would receive higher grades on the project deliverables.

3.2. Quiz Performance to Measure Team Function

As part of the TBL teaching approach, students complete a set of assigned readings for each course module and then take an in-class quiz on those readings. The reading quiz is completed prior to any classes, tutorials, or other teaching activities on the reading topics. The reading quiz consists usually of 20 questions, all multiple choice. Of particular importance in the current context, students first complete the reading quiz individually by marking their responses on a computer score card, and then they complete the quiz a second time in their teams. The team quiz questions are identical to those on the individual test, but responses are recorded using an Immediate Feedback Assessment Technique (IF-AT) scratch card, as shown in Fig. 1.

On the IF-AT card questions are numbered as rows down the left side and answer choices are labelled as columns across the top of the card. Teams discuss the question and scratch the box corresponding to their response. If they are correct they reveal a star symbol (★), such as for Question 1, C, in the figure. If teams are incorrect, no star is revealed (such as Question 2, B, in the figure), and then teams discuss the question further and try again until they reveal the correct answer (Question 2, A). The score per question is based on the number of
scratches to reveal the star. (In MECH 223 1 try = 4 points; 2 tries = 2 points; 3 tries = 1 point; and 4 or 5 tries = 0 points.)

Having both individual and team responses to the exact same questions on tests taken immediately after one another allows comparison of how team decision making relates to decision making of the individual team members. Specifically, the IF-AT card reveals cases where the team is incorrect (such as Question 2, answer B, in Fig. 1) and the computer-scored individual tests, taken prior to the team test, reveal how each person responded to the same question without the influence of the other team members.

In total, over 11,000 question events were analyzed from the 2008 to 2012 offerings of the course. These were based on 30 quizzes, 10-20 questions per quiz, and 20 teams per year. A total of 1,143 incorrect team responses to questions were identified.

The metric used to quantify team function in this case was the number of instances per team where individual members had a question correct in the individual test but then the team picked a different answer in the team test. It is suggested a high value of this metric in general corresponds to lower team function in communication, problem solving, and unity, as described in the previous section (e.g. teams that base decisions on ‘majority rule’ or the ‘loudest voice’). It is further hypothesized that teams with a low value of this metric will tend to have improved project outcomes.

3.3. Peer Evaluation Scores Received

The peer evaluation scores received were examined in terms of the distribution of scores in each team. As described in Section 2, each student completed a total of six peer evaluations in which they evaluated each other member on their team. The metric of team function considered was the degree of differentiation between peer evaluation scores measured in terms of the standard deviation of scores received across all team members for all six evaluations (nominally 36 individual scores per team).

It is suggested that teams with less variation between peer evaluation score will tend to have improved distribution of responsibility, conflict management, and unity function, as defined in Section 3.1. Furthermore, it is hypothesized that improvement in these areas of team function will also tend to result in better project outcomes.

4. RESULTS AND DISCUSSION

The sections below present the results and discussion for the comparison of project outcomes to team function, measured using the three methods in Section 3: student self-rating of team function, comparison of individual and team quiz performance, and variation in peer evaluation scores.

4.1. How Measured Team Function Relates to Student Ratings of Team Function

The comparison of student self-rating of team performance with composite project score is shown in Fig. 2. Independently for each of the six team function criteria, the 25% of the class that gave their team the highest rating in the survey for that criterion (‘higher team function score’) are compared to the 25% that gave their team the lowest rating (‘lower team function score’). The project scores for each year have been normalized to an average of 1.00 to remove year-to-year variability in project difficulty, grading practices, and cohorts. Error bars represent 95% confidence intervals based on the measured data.

![Fig. 2. Normalized project score versus student self-rating of team function](image-url)
In each measure of team function, the students who rated their team’s function as high had a higher composite project score on average compared to those who rated their team’s function as low. The gaps between the two groups were statistically significant (p < 0.001) in every case. The difference in project score ranged from approximately 4% (communication and conflict management) to approximately 6% (unity and distribution of responsibility). Overall, the results suggest there is a very strong correlation between student perception of team function and the design outcomes the team produces. Put another way, it appears student assessment of team performance is valid.

4.2. How Team Function Relates to Decision Making on Team Quizzes

The results comparing normalized project score to decision making on team quizzes is shown in Fig. 3. Several items in this chart warrant description. First, the four quartiles are based on the number of instances where an individual had the correct answer on the individual quiz but the team chose a different (incorrect) answer on the team quiz. The quartiles were determined across combined data from the 2008W to 2012W cohorts and were based on average instances per team per quiz:

- Q1 = 0 to 5.7 instances
- Q2 = 5.7 to 9.2
- Q3 = 9.2 to 12.5
- Q4 = 12.5 to 22.0

Second, the p-values shown between the bars in the chart are results of the two-tail T-test for difference in means between the two adjacent quartiles (i.e. the difference in project scores for the Q1 and Q2 groups are statistically significant with a p-value of 0.008).

The general trend observed above from Q2 to Q4 followed expectations: the more likely teams were to choose an answer on a team quiz that was incorrect in spite of the fact that one of the team members actually knew the correct answer was correlated with decreasing score on team design project outcomes. What was not expected was that the students in Q1, where there were the fewest cases of teams going against correct team members in quizzes, actually showed a reduction in team design project outcomes. The low p-values associated with the comparison of Q1 and Q2 suggest this result is statistically significant, as is the comparison of Q3 and Q4. The scale of impact on team project outcomes was 4% between best and worst performing quartiles. It is important to note that the team quiz scores followed expectations: Q1 was the top performer (1.015), followed by Q2 and Q3 (both at 1.002), and Q4 was the bottom performer (0.981).

Overall, these results suggest that not listening to and not considering all team members’ input is undesirable (i.e. Q2-Q4 above). In considering the results for Q1, this may be suggesting that, in the quiz setting, the team can rely on the ‘one smart individual’ who has the bandwidth to know all the quiz answers, but lacks the bandwidth to maintain such performance on the project. Another possible explanation for the Q1 result is that a high degree of focus on equal and democratic input on all decisions (even trivial ones) might benefit quiz performance but slow a team down on a large, open-ended design project.

4.3. How Team Function Relates to Peer Evaluation Score Differentiation

The relationship between normalized project score and peer evaluation score differentiation is shown in Fig. 4. The peer evaluation differentiation was quantified for each team by, first, computing the standard deviation of peer evaluation scores received by every teammate on their six evaluations, and, second, averaging these standard deviations across all members on a team. The normalized project scores were then averaged based on peer evaluation scores binned in 5% intervals, with the special case of no differentiation (0%) grouped separately.
Teams with the smallest non-zero degree of differentiation tended to perform well on the projects, as was expected, and there was a general decrease in project performance as the peer evaluation score differentiation increased. The error bars in the charts represent 95% confidence intervals. Teams with a very large degree of differentiation (the >15% category) also had the highest project scores; however, there were only 5 of 180 teams in this category so it would be premature to suggest this finding represents a broader trend. It has been coloured differently in the chart for this reason.

A second interesting and arguably more important finding was that teams with no peer evaluation differentiation – teams where each member always received the same score as each other member, no matter what – had the lowest average project performance. Anecdotally, we have been aware that teams who enter into these types of agreements tend to have disproportionately more conflict and infighting when there no longer is a meaningful recourse to individuals who stop pulling their weight or begin to dominate team decisions. If nothing else, these results show that not utilizing the peer evaluation process as intended is a marker for poor project success.

Overall, the results suggest that teams with a high degree of unity, distribution of responsibility, and conflict management, as measured by low peer evaluation differentiation, tend to produce better design project outcomes. The scale of the effect, not including the five teams at the upper end of the scale of differentiation, was approximately 7% in terms of normalized project score.

5. CONCLUSIONS

The influence of team function on design project outcomes was examined in this study. Team function was interpreted in terms of six key dimensions: team unity, team communication, distribution of responsibility, problem solving, conflict management, and team self-evaluation. Three different methods were used to assess team function:

- A survey in which students self-assessed their team’s performance in each of the above dimensions of team function
- A comparison of individual and team performance on identical quizzes as a pseudo-measure for communication, problem solving, and unity (based on how often teams chose incorrect answers when one of the team members knew the correct answer)
- A measurement of the differentiation in peer evaluation scores as a pseudo-measure for distribution of responsibility, conflict management, and unity

Design project outcomes were measured using a composite of several distinct, graded second year design project deliverables, including comprehensive prototypes for design competitions, formal written reports, oral presentations, poster presentations, focused prototype demonstrations, and logbooks.

With all three measures of team function, statistically significant relationships to design project outcomes were observed. Teams that self-rated as performing well in the different dimensions of team function scored 4 to 6% higher on normalized project scores compared to those that self-rated as performing less well. There was a general trend that teams with a greater tendency to answer team quizzes incorrectly when some members had the correct answer (i.e. not listening to the wisdom within the team) were also likely to receive lower composite design project scores by as much as 4%. The teams least likely to answer incorrectly when some members had the correct answer performed below average, perhaps due to a lack of firm leadership or an overreliance on democratic decision making approaches during the projects. Lastly, there was a general decrease in composite project scores as the differences between peer evaluation scores increased on a team. Similar peer evaluation scores between members of a team is suggestive of equal distribution of responsibility, less conflict, and strong unity for that team. The exception was teams that did not differentiate peer evaluation scores at all had average normalized project scores 7% lower than teams with a small non-zero differentiation in peer evaluations.

References


