Abstract - As a result of an expanding global marketplace and multidisciplinary technological advances, the successful, modern engineer must work effectively in teams. Further emphasis on the importance of preparing students for teamwork was promulgated by ABET’s Engineering Criteria 2000. While the concept of team-based learning is valued among many educators, the quantitative evaluation of individual student performance associated with a team-based deliverable can be challenging. The focus of this research was to evaluate the ability of engineering students enrolled in a multidisciplinary, team-based capstone design course to evaluate their personal, peer and team performance. A team-based peer/self evaluation form, developed by the Synthesis Coalition and modified by the instructors to fit the needs of the course, was utilized to collect the assessment data. Data were collected for a two year period by a single faculty member that served as instructor. Data analyses were performed to show the relationships between student self-evaluation and faculty measures. Factors such as student gender and ethnicity were considered.

Keywords: senior design, self-assessment, problem-based learning, interdisciplinary teams

INTRODUCTION

Literature has validated the concept of team-oriented, problem-based learning (Northwood et al., 2003; Woods et al., 1997 and 2000). The importance of incorporating teamwork skills development into engineering curricula was highlighted by the Accreditation Board for Engineering and Technology (ABET) in the EC 2000 recommendations Criterion 3, letter d that states “Engineering programs must demonstrate that their students attain an ability to function on multi-disciplinary teams.” With the inclusion of this personal performance skill into accreditation, more academic work is being conducted by student teams and faculty are now tasked with grading work produced by teams and then assigning appropriate grades to individual team members. As a result, the student’s ability to provide meaningful self and peer-assessment of team performance is important so that individual performance can be properly assessed and feedback can be provided to the instructor letting them know if course learning objectives, including team work skills, are being met.

Although the validity of self and peer ratings are often questioned, a variety of assessment tools have been developed that account for individual performance by team members working within a group. For example a peer rating system was developed and implemented within the Departments of Communication & Electrical Engineering and Computer System Engineering at the Royal Melbourne Institute of Technology University, which accounted for individual performances on team projects (Brown, 1995). Within this system, team members confidentially rate their peers and through this process a weighting factor is developed for each team member that can be used by the instructor to manipulate the student’s final project grade. Additional examples of peer and self rating systems for

---

1 Mercer University School of Engineering, Department of Environmental Engineering, 1400 Coleman Ave, Macon, Georgia 31207, lackey_l@mercer.edu

2 Mercer University School of Engineering, Program Director, Department of Environmental Engineering, 1400 Coleman Ave, Macon, Georgia 31207, mines_ro@mercer.edu

3 Mercer University School of Engineering, Department of Mechanical Engineering, 1400 Coleman Ave, Macon, Georgia 31207, jenkins_he@mercer.edu
the assessment of individual contributions to team work include The Team Developer (McGourty and DeMeuse, 2000; McAnear et al., 2000), SPARK (self and peer assessment resource kit) developed by Freeman and McKenzi (2002) and the Self/Peer Team Assessment Survey (Van Duzer and McMartin, 1999).

The importance of the validity and reliability of self and peer assessment tools has been highlighted in the work conducted by a variety of researchers (Van Duzer and McMartin, 1999; Mehta and Danielson, 1999; Kaufman and Felder, 2000). Furthermore, a quantitative study that compared student-generated and instructor grades was performed by Boud and Falchikov (1989). Within this study, the researchers questioned if students (1) under- or over-rate themselves compared to the instructor, (2) of different academic abilities have similar tendencies, (3) in different kinds or levels of courses tend to over- or under-predict performance, and (4) improve their ability to objectively grade themselves with practice. Additional research explored whether students change their self and peer assessment practices if their remarks are used by the instructor to develop course grades.

The goal of this project was to compare student to faculty generated scores on two teaming related questions from the Self/Peer Team Assessment Survey developed by Van Duzer and McMartin, 1999. Question 1 requires individual students to rank their team productivity while the second question ranks individual student performance as a team member. Student responses are compared to associated faculty scores. The survey was administered within the Mercer University School of Engineering (MUSE) two-semester senior design sequence (EGR 487-488) at the end of each term.

METHODS

Course Description

The Mercer University School of Engineering requires all engineering students as well as students majoring in Industrial Management to complete a sequential, two-semester design experience during the senior year. Students are encouraged to form multidisciplinary teams and to focus on industrial sponsored projects. Teams most typically consist of three students studying the appropriate disciplines for successful completion of their selected project. Communication between student teams, client, technical advisors (discipline specific faculty member) and course instructor is forced through periodic meetings and written progress reports. A review of appropriate refereed literature is required of each student, student peer review of final documents is encouraged and students participate in a “just in time” lecture series (topics include team building, brainstorming, design process review, engineering ethics, intellectual property, communication skills, future of engineering design, project management) during the first semester.

Within the capstone design sequence, the focus of the first semester is project development where students write a proposal and then conduct a preliminary design review (PDR). The PDR is presented in both written and oral forms. During the second semester, students build and test as prescribed in their PDR. The process is culminated with the written and oral deliveries of their respective critical design reviews (CDR) where the build and test phase is detailed.

Study Participants

Over a two year period, a total of 76 students, managed by one instructor, completed the senior design sequence. Demographic information is provided in Table 1.

Assessment Instruments

At the conclusion of both the first and second semester of the senior design course, each student is required to complete a Self/Peer Team Assessment form (Van Duzer, E. & McMartin, 1999). This tool was modified by MUSE faculty in an effort to assess student progress associated with a variety of learning outcomes. The entire form is provided for reference in Appendix 1. For this study, student responses for two questions on the assessment form, as outlined in Table 2, were evaluated.

As a course requirement, each student must have a faculty member within their specific discipline serve as a technical advisor to their team. As a result, each design team may be working with one to three faculty members (technical advisors) in addition to their instructor. Technical advisors were also asked to complete a survey that was very similar to the student survey detailed in Appendix 1. Table 2 also presents the two questions utilized in this study from the technical advisor survey. Although the completion of the forms is required by students and recommended for the faculty technical advisors, some data are missing and is appropriately noted within discussions.
and tables. Within the table, the number in parenthesis associated with a specific response identifies the numerical value used to represent that response. Faculty assigned grades were assessed using a four-point, A – F system.

Table 1. Summation of a two-year period of senior design student demographics (n = 76). All students in this study had the same instructor for the course.

<table>
<thead>
<tr>
<th>Student</th>
<th>n</th>
<th>Percent of total population (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male Total</td>
<td>46</td>
<td>61</td>
</tr>
<tr>
<td>Minority male</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>Female Total</td>
<td>30</td>
<td>39</td>
</tr>
<tr>
<td>Minority Female</td>
<td>13</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td>76</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Assessment questions utilized in this project. Notice the similarity between questions 1 and 2 for student team members and technical advisors. The number in parenthesis associated with a specific response identifies the numerical value used to represent that response.

<table>
<thead>
<tr>
<th>Student item</th>
<th>1. How productive was the group overall?^a</th>
<th>1. From your perspective, how productive was the team overall?^a</th>
<th>2. Overall, how would you rate your own ability to perform effectively on this multidisciplinary team?^b</th>
<th>2. Overall, how effective would you say this team has been at working together, based on your experience with other project teams?^b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student item</td>
<td>Accomplished some but not all of the project’s requirements (1)</td>
<td>Met the project requirements but could have done much better (2)</td>
<td>Efficiently accomplished goals that we set for ourselves (3)</td>
<td>Went way beyond what we had to do exceeding even our own goals (4)</td>
</tr>
<tr>
<td>Faculty item</td>
<td>Accomplished some but not all of the project’s requirements (1)</td>
<td>Met the project requirements but could have done much better (2)</td>
<td>Efficiently accomplished goals that we set for ourselves (3)</td>
<td>Went way beyond what we had to do exceeding even our own goals (4)</td>
</tr>
<tr>
<td>Student item</td>
<td>2. Overall, how would you rate your own ability to perform effectively on this multidisciplinary team?^b</td>
<td>Poor (1)</td>
<td>Improvement needed (2)</td>
<td>Good (3)</td>
</tr>
<tr>
<td>Faculty item</td>
<td>2. Overall, how effective would you say this team has been at working together, based on your experience with other project teams?^b</td>
<td>Very ineffective (1)</td>
<td>Ineffective (2)</td>
<td>Effective (3)</td>
</tr>
</tbody>
</table>

^a This question parallels to item 1c of the assessment instrument detailed in Appendix 1.
^b This question parallels to item 5 of the assessment instrument detailed in Appendix 1.

The two-sample t-test was used to derive all levels of significance reported (unless otherwise directly noted in the text). Average student ratings were compared to technical advisor and instructor rating using linear regression analysis. Correlation coefficients were derived using Pearson’s formula.

RESULTS AND DISCUSSION

Group productivity was assessed using question 1 in Table 2 by both individual team members and faculty technical advisors. For each team (n=33), student and technical advisor rankings were averaged and the relationship between scores is presented in Figure 1. Each point on the graph represents averages for one team and results from the first (PDR) and second (CDR) semester are delineated. It is apparent that as the technical advisor score increases, the average student score typically increases even though there is appreciable scatter in the data.

Two lines are shown in Figure 1. One is the 45° line going through the origin and point (4,4). For a data point to fall on this line, both students and faculty would have to rank team productivity the same. The second line on the graph represents the linear regression line through all the data. A moderate to strong positive correlation, r = 0.65, is
noted between student and faculty responses. Results indicate that weaker teams tended to over predict group productivity while overachievers tended to under predict their success. For example, when the technical advisor response is above approximately 3, more points are located below the 45° line than above. Further evidence of this trend is noted by the slope of the linear regression line being less than one.

Figure 1. Correlation between student and technical advisor assessment of team productivity.

Figure 2 shows a similar plot for assessment question 2 (refer to Table 2) that deals with student ability to function on multidisciplinary teams. All of the observations made above for question 1 are also applicable to question 2. That is, the same trends and tendencies were observed. The regression coefficient, $r = 0.203$, indicates a very weak positive correlation between student and technical advisor average scores.

Figure 2. Correlation between student and technical advisor assessment of student ability to work on a multidisciplinary team.

Average group productivity ratings are shown in Table 3 for both the first and second semester. The Wilcoxon (Wilcoxon, 1945) signed rank test showed that the observed averages for group productivity for student and instructor (measured by course grade) were significantly different. When considering all students, peer ratings
correlated positively with faculty assigned course grade showing a Pearson’s correlation constant, \( r = 0.53 \) with \( p<0.0001 \).

Although not statistically significant, for all groups of students listed in Table 3, the average student ranking quantifying group productivity was lower in the second semester (\( p=0.28 \)). This indicated that students perceived their performance during the building and testing phase of the project inferior to their performance during the planning phase. During the first semester, the average rank of all students regarding their team’s group productivity was 2.84±0.5 while the second semester average was 2.71±0.71. Possible reasons why faculty course grade rankings are higher than student ranked productivity scores are discussed. Course grades include a variety of variables including credit for completing and submitting the Self/Peer Assessment form and several team and individual progress reports that are due at intervals throughout the semester. These types of activities tended to inflate the overall student-assigned course grade.

Although not shown to be significant (\( p=0.31 \)), during the first semester, male students ranked their performance (2.86±0.65) associated with group productivity higher than female students (2.70±0.70). Similarly, minority students ranked their group productivity performance lower than the average observed for all non-minority male and female students (\( p=0.01 \)). Second semester scores showed no significant difference (\( p=0.23 \)) between the average rankings from all non-minority 2.80±0.68 and minority students 2.50±0.76.

The student’s perception of their ability to perform on multidisciplinary teams is also addressed in Table 3. In contrast to student perception of team performance between the first and second semesters, student perception of their ability to function as a team member increased from the first to second semester. Although not shown to be statistically significant (\( p=0.24 \)), the average score for all students after the completion of the first and second semester activities increased from 3.27±0.57 and 3.55±0.18, respectively. No significant difference was noted between scoring from any two groups during activities from either semester.

Table 3. Average student team scores describing group productivity and student ability to work on multidisciplinary teams. Average instructor grades are also provided.

<table>
<thead>
<tr>
<th>Population</th>
<th>Group Productivity (Refers to question 1 in Table 2)</th>
<th>Ability to Perform on Multidisciplinary Team (Refers to question 2 in Table 2)</th>
<th>Instructor Assigned Grade (Course grade = 0-4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st Semester (PDR)</td>
<td>2nd Semester (CDR)</td>
<td>1st Semester (PDR)</td>
</tr>
<tr>
<td>All students (SD)</td>
<td>2.84(0.50)</td>
<td>2.71(0.71)</td>
<td>3.27(0.57)</td>
</tr>
<tr>
<td>Male (SD)</td>
<td>2.86(0.65)</td>
<td>2.77(0.71)</td>
<td>3.21(0.53)</td>
</tr>
<tr>
<td>Female (SD)</td>
<td>2.70(0.70)</td>
<td>2.61(0.72)</td>
<td>3.39(0.63)</td>
</tr>
<tr>
<td>Minority (SD)</td>
<td>2.70(0.86)</td>
<td>2.50(0.76)</td>
<td>3.50(0.50)</td>
</tr>
<tr>
<td>Male Minority (SD)</td>
<td>2.75(0.89)</td>
<td>2.50(0.76)</td>
<td>3.33(0.58)</td>
</tr>
<tr>
<td>Female Minority (SD)</td>
<td>2.67(0.89)</td>
<td>2.50(0.80)</td>
<td>3.56(0.50)</td>
</tr>
</tbody>
</table>

SD = one standard deviation

Table 4 shows the distribution of student scores after completion of second semester activities relative to their technical advisor score for survey question 1 (refer to Table 2). Surprising, inflated ratings (overestimating) by students assessing their team’s overall performance were less common than those team members that either ranked themselves lower or the same as the technical advisor score. Forty eight percent of female students underestimated their team productivity compared to 35% of their male peers.

Table 5 shows the distribution of student scores after completion of second semester activities relative to their technical advisor score for survey question 2 (refer to Table 2). When comparing second semester scores for individual student ability to work on interdisciplinary teams, 55% of the students over-predicted their competence. Female students were more confident in their teaming abilities (65% overestimated) as compared to their male peers (48% overestimated).
Table 4. Distribution of senior design students (n=66) according to gender based on questions 1 (refer to Table 2) from the self-assessment survey conducted during the second semester of the senior design sequence. Overestimating refers to the student score being significantly ($p<0.001$) higher than the average reported by the technical advisors. Underestimating refers to the student score being significantly ($p<0.001$) less than the average reported by the technical advisors.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Overestimate</th>
<th>Underestimate</th>
<th>No Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n %</td>
<td>n %</td>
<td>n %</td>
</tr>
<tr>
<td>Male</td>
<td>43 (65)</td>
<td>12 (28)</td>
<td>15 (35)</td>
</tr>
<tr>
<td>Female</td>
<td>23 (35)</td>
<td>7 (11)</td>
<td>11 (17)</td>
</tr>
</tbody>
</table>

Table 5. Distribution of senior design students (n=66) according to gender based on questions 2 (refer to Table 2) from the self-assessment survey conducted during the second semester of the senior design sequence. Overestimating refers to the student score being significantly ($p<0.001$) higher than the average reported by the technical advisors. Underestimating refers to the student score being significantly ($p<0.001$) less than the average reported by the technical advisors.

<table>
<thead>
<tr>
<th>Gender</th>
<th>Overestimate</th>
<th>Underestimate</th>
<th>No Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n %</td>
<td>n %</td>
<td>n %</td>
</tr>
<tr>
<td>Male</td>
<td>27 (61)</td>
<td>13 (48)</td>
<td>6 (22)</td>
</tr>
<tr>
<td>Female</td>
<td>17 (39)</td>
<td>11 (24)</td>
<td>2 (12)</td>
</tr>
</tbody>
</table>

Conclusions

In the MUSE senior design course sequence (a two-semester experience), students work in small multidisciplinary project teams giving them practical experience in teamwork and communication while also developing their leadership, management and people skills. The interaction with industrial project sponsors and exposure to real world problems in advance of their entering the workforce helps equip them to practice engineering as a career. Within the course, a Self/Peer Assessment form was administered at the end of each semester in an effort to assess individual student as well as team performance. A similar assessment tool was used by senior design team technical advisors. This manuscript focused on data from two questions on the assessment form. Question 1 dealt with how an individual student and their associated technical advisor perceived their overall team performance, while question 2 asked students to rate their individual performance on multidisciplinary teams. Similarly, faculty technical advisors were asked to rate student teaming skills. The following inferences were observed:

1. When student and technical advisor rankings were averaged for question 1 that assessed group productivity, a moderate to strong positive correlation ($r=0.65$) was observed between student and faculty responses. Results indicate that weaker teams tended to over predict group productivity while overachievers tended to under predict their success.

2. When student and technical advisor rankings were averaged for question 2 that assessed student ability to function on a multidisciplinary team, a weak positive correlation ($r=0.20$) was observed between student and faculty responses. Results indicate that weaker students tended to over predict their teaming skills while overachievers tended to under predict their abilities.

3. Student scores relating to group productivity correlated positively with faculty assigned course grades ($r=0.53$ with $p<0.0001$).
4. Results indicated that students perceived their performance during the building and testing phase of the project inferior to their performance during the planning phase.

5. During first semester activities, male students tended to rank their ability to perform on multidisciplinary teams higher than their female peers.

6. Student perception of their ability to function as a team member increased from the first to second semester.

7. Inflated ratings by students assessing their team’s overall performance were less common than those team members that either ranked themselves lower or at the same level as the technical advisor.

8. Second semester assessment activities indicated that 55% of students over-predicted their teaming abilities. Female students were more confident in their teaming skills (65% overestimated) as compared to their male peers (48% overestimated).

REFERENCES


BIOGRAPHICAL INFORMATION

Laura W. Lackey

Laura W. Lackey is an Associate Professor in the Department of Biomedical and Environmental Engineering at the Mercer University School of Engineering. She earned B.S., M.S., and Ph.D. degrees in Chemical Engineering from the University of Tennessee. The terminal degree was awarded in 1992. She has six years of industrial experience at the Tennessee Valley Authority where she conducted research with emphasis on the mitigation of organic wastes through bioremediation. In the seven years since Dr. Lackey began her career at Mercer, she has taught 15 different
courses, ranging from a freshman-level Introduction to Problem Solving course to a senior-level Process Chemistry course, which she developed. Address: Mercer University School of Engineering, 1400 Coleman Ave, Macon, GA, 31207; telephone: 478-301-4106; fax: 478-301-2166; e-mail: lackey_l@mercer.edu.

Richard O. Mines, Jr.

Dr. Mines is Professor and Program Director of Environmental Engineering at Mercer University in Macon, Georgia. He has eighteen years of teaching experience and has been a professor at the Virginia Military Institute, University of South Florida, and Mercer University. Dr. Mines has approximately 7 years of consulting experience with Matotan & Associates, CH2M Hill, and Black & Veatch. He has authored or co-authored over 70 articles on technical and educational related topics. Dr. Mines has conducted numerous studies on the activated sludge process. He is a registered professional engineer in Florida, New Mexico, and Virginia. He is a member of the Water Environment Federation, American Society for Engineering Education, and the American Society of Civil Engineers.

Hodge E. Jenkins

Dr. Hodge Jenkins is an Assistant Professor of Mechanical Engineering in the Department of Mechanical and Industrial Engineering at Mercer University in Macon, Georgia. Prior to coming to Mercer in 2002, Dr. Jenkins was engaged in optical fiber product development with Bell Laboratories of Lucent Technologies. He is a registered professional engineer, and with over 20 years of design and development experience in high-precision design, dynamic structural analysis, process automation, control, and robotics. Dr. Jenkins holds a Ph.D. in Mechanical Engineering from Georgia Institute of Technology in (1996), as well as BSME (1981) and MSME (1985) degrees from the University of Pittsburgh. His professional affiliations include ASME and ASEE.
**APPENDIX 1**

## Self/Peer Team Assessment

Name_________________________ Date______________________________

1. Please circle the rating that best describes your team for each of the three questions below.

   a. Did all members of the group share in the team's responsibilities?
      
      | Some members did no work at all. | A few members did most of the work. | The work was generally shared by all members. | Everyone did an equal share of the work. |

   b. Which of the following best describes the level of conflict at group meetings?
      
      | Open warfare: still unresolved | Disagreements were resolved with considerable difficulty | There were disagreements, but they were easily resolved | No conflict; everyone seemed to agree on what to do |

   c. How productive was the group overall?
      
      | Accomplished some but not all of the project's requirements | Met the project requirements but could have done much better | Efficiently accomplished goals that we set for ourselves | Went way beyond what we had to do exceeding even our own goals |

2. Please rate yourself and each team member on how well the following phrases describe your team's work.

<table>
<thead>
<tr>
<th>Disagree</th>
<th>Tend to disagree</th>
<th>Tend to agree</th>
<th>Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

   Team Member's Names
   
   Self:  

   a. Failed to do an equal share of the work
   b. Kept an open mind/was willing to consider others’ ideas
   c. Was fully engaged in discussions during meetings
   d. Took a leadership role in some aspects of the project
   e. Helped group overcome differences to reach effective solutions
   f. Often tried to excessively dominate group discussions
   g. Contributed useful ideas that help the group succeed
   h. Encouraged group to complete the project on a timely basis
   i. Delivered work when promised/needed
   j. Had difficulty negotiating issues with members of the group
   k. Communicated ideas clearly/effectively

\[ \text{OVER} \]
3. Write a brief description of the problems you encountered in working with this group and how they were resolved.

4. Please distribute 100 points among the members of your team, based on each member's contribution to the team's efforts. (Don't forget to include yourself.) Use integers only. No two people should receive the same number of points.

<table>
<thead>
<tr>
<th>Name:</th>
<th># of Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Self)</td>
<td></td>
</tr>
</tbody>
</table>

Total 100

5. Over all, how would you rate your own ability to perform effectively on this multidisciplinary team?

<table>
<thead>
<tr>
<th>Poor</th>
<th>Improvement needed</th>
<th>Good</th>
<th>Excellent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>