AC 2011-220: GENDER DIFFERENCES IN INDIVIDUAL AND TEAMMATE PERFORMANCE EVALUATIONS BY STUDENTS ON ENGINEERING DESIGN TEAMS

Natalie CT Van Tyne, Colorado School of Mines

Natalie Van Tyne is the director of the Design Engineering Introductory Course Sequence (Design EPICS) Program at Colorado School of Mines, in which she has taught since 2002. The program provides required first- and second-year students with a foundation in the fundamental skills of engineering design: problem solving in teams, decision analysis, graphical representation, oral and written communications, and numerical analysis. Ms. Van Tyne holds bachelors degrees in chemical engineering and Russian language, masters degrees in chemical engineering, environmental engineering and business adminstration, and is a registered professional engineer in Colorado. She is also a proud alumna of Rutgers University, Lehigh University and Colorado School of Mines.

Dr. Chester J. Van Tyne, Colorado School of Mines
Kathryne Van Tyne, University of Chicago

©American Society for Engineering Education, 2011

Gender Differences in Individual and Teammate Performance Evaluations by Students on Engineering Design Teams

Abstract

Evaluations of team participation and of individual students' value to the team were conducted in a first-year engineering design course at a public university in the western United States. The quantitative measures included a self-evaluation, as well as an evaluation of the performance of other team members. This study analyzes data from 866 students (232 women) in 22 classes during the period from spring 2002 through spring 2009. Parameters included: gender, self-evaluation score, evaluation score from peers and final numeric course grade. A Wilcoxon rank sum test was used to compare the differences between self- and team evaluation scores, as well as to compare average peer evaluation with final course grade. We did not find significant differences between male or female students in either self-evaluation or actual performance. It was also found that, irrespective of gender, students who had a 5% higher self-evaluation score than their team evaluation score actually received lower course grades than those who under-valued their contribution to their team. These results indicate that small, engineering-focused institutions may provide a learning environment and underlying support system for women that result in greater self-efficacy; or they may indicate that this type of institution attracts women students who already have a strong commitment to the study of engineering and the necessary tenacity to succeed in this field.

Introduction

Teamwork in engineering education provides students with important experiences that are representative of the modern engineering workplace.[1, 2] While working on a team, students need to consider how their performance meets both personal expectations and those of their team members, similar to the way a modern engineer must consider their own job performance from the perspective of their co-workers. As more women enter the field of engineering, more teamwork experiences, both in universities and in the workplace, will involve persons of both genders.

Effective performance on a team is not as easily measured[3] as other essential skills, such as technical, analytical, or communication abilities. For example, the evaluation of student performance on a team often depends on both personal and peer evaluations. Since both male and female students can serve on the same teams, gender differences in perspective may influence how students evaluate themselves and others. The present paper explores the role of gender in self-evaluation and peer evaluation, with regard to both teamwork and final overall performance, as measured by final course grade in a first-year engineering design course.

Individual Factors for Successful Evaluations

Since part of our study relies upon a self-evaluation of the students' performance on a team, it is important to understand how students approach such evaluations. Self-evaluation of one's teamwork skills and contributions to the success of a design team is influenced, in part, by an attribute known as self-efficacy. Self-efficacy is a person's belief in their ability to do something

(e.g. perform a task, or accomplish a goal).[4,5] Self-efficacy has been linked to positive outcomes in the pursuit of careers in fields that are traditionally less populated by women than by men. A student's belief in their ability to succeed in a field such as engineering is a major factor in the probability of their success, because it has been demonstrated that students need to possess the will, as well as the skills, to succeed in engineering.[6]

Self-efficacy in engineering courses has been measured in several ways by using all-female or mixed-gender groups.[4,6,7] Chemers et al.[8] present evidence to support a theoretical model of the direct and indirect effects of self-efficacy and optimism on academic performance; as a result, students with higher self-efficacy and optimism were found to be more successful in their academic pursuits. In addition, they were better able to cope with both life and school stressors.[8] Coping and personal adjustment abilities are vital to success in a college environment, especially among first-year students encountering numerous situations different from their previous experiences.

Gender Differences in Team Performance

In addition to self-efficacy, variations in team member performance may be influenced by other factors relating to the respective genders of the team members. These include any team members' innate abilities related to engineering work, as well as a variety of social and cultural influences.

Analysts of the innate aptitudes of men and women for the skills related to engineering work emphasize that, unlike 19th and early 20th century beliefs, there are no inherent gender differences in general intelligence.[9] Results in the 1930s from the Primary Mental Abilities (PMA) test led to a belief in significant gender differences in verbal, mathematical, and spatial abilities. More recently, through a combination of meta-analysis and the division of these three major types of abilities into more specific skills, researchers have found greater differences between skill levels among people of the same gender than those of differing gender.[8] Any perceived gender differences in engineering-related abilities may influence the behavior of both male and female students when working on engineering design teams. A number of studies and explanations describe these differences, and offer possible causes for these differences in perception.[2,7,10-14]

Attitudes about gender differences in engineering may also affect team performance. A study of male and female engineering students during the early 1990s showed that the female students faced obstacles to success that prevented them from competing with the male students on an equal basis, even if their qualifications were better than those of the men.[7] Such women earned lower grades, exhibited lower confidence levels, and held lower expectations of personal performance than their male counterparts as they progressed through their studies.[7] The combination of lower confidence levels and lower expectations can affect individual and team performance if the team members view an individual's contributions as inadequate for the team's success.

A student who believes that he or she cannot effectively contribute to a team's success will often be reluctant to try. Female students, in particular, may lack an understanding of what it means to

be successful[13] and may therefore be disappointed in their own team performance, even when their abilities and accomplishments indicate otherwise.[6]

Rudman[12], as well as Babcock and Laschever[13], described the difference in responses among mixed-gender group participants toward the practice of self-promotion as a part of impression management, wherein a student of either gender represents themselves as highly capable, skilled and confident in their abilities to perform well on a team. Experimental results showed that women failed to reward a self-promoting woman. Female self-promotion may be less acceptable to other women because independence and self-focus are traditionally less accepted traits in women.[12] In fact, women who view self-promoting women negatively may actually be responsible for perpetuating gender stereotypes that impede their own socio-economic progress and reduce their own self-efficacy.

For mixed audiences, male response to a woman's self-promotion depended on if the two individuals were outcome-dependent (i.e., did his success depend on her capabilities), such as on an engineering design team. It has been found that women who self-promote may enhance others' impressions of their qualifications at the expense of their own social acceptance.[12] Social acceptance also contributes to "feelings of inclusion," which is identified as a factor in self-efficacy.[4,6]

Oral communication among engineering design team members has also been studied for additional manifestations of social and cultural pressures, which may cause male and female team members to act or respond in distinctly different ways.[2,15] An important conclusion to draw from these studies is that if a balance could be reached between female engineering students eliminating self-effacing speech acts and male engineering students engaging in fewer self-promotional and more modest speech acts, then a more collaborative environment could be developed within the team,[11,15-17] which may raise overall peer evaluation scores.

Rosser[16-17] promotes a transformation of the engineering/science curriculum that may lead to an interdisciplinary approach that uses both qualitative and quantitative methods, removes biases based on the views of one gender or social class, and places solutions in view of the impact on society. These would lead to greater value placed on a design project's solution, which provides positive reinforcement to the dynamics of the successful design team.

Team Performance Based on Other Types of Differences

The experience of Schultz[18] in organizing student teams for a semester-long engineering project reveals a number of factors that affect team performance: abilities, attitudes, and personalities. For example, a team of high-performing, serious students tended to be individualistic and did not work well together. Conversely, one highly successful team in Schultz's course consisted of average students who never argued and showed a strong willingness to work diligently on the project. It seems that a balance of personalities, interest in the project, and attitudes toward working as a team can result in positive team dynamics.

Objectives of Study

The present study was designed and conducted to address the following two research questions:

1. Is there a difference between male and female students regarding self-evaluated performance on a design team, as compared to their evaluation by fellow team members?
2. Is there a relationship between the difference in self-evaluation and evaluation by fellow team members and the final grade in the course (i.e., do low-grade students overrate themselves to a greater extent that high-grade students)?

Methods

In order to evaluate these two research questions, we used data from a first-year engineering design course, where the students performed both a self-evaluation and an evaluation of their teammates. The data were subjected to statistical analysis to determine if there were differences in these evaluations with regard to gender, and with regard to the students' self-evaluation *vis-a-vis* their final course grade.

Data Collection

At a public university in the western United States, we measured effective team participation and value to the team by means of peer evaluations in a first-year design course. On two days during the semester, the students on each team of (nominally) five members were given an exercise where they provided a qualitative evaluation of the performance of each member of the team, as well as themselves. They were asked to consider the following characteristics in this evaluation: reliability, attitude, technical competence, leadership, cooperation, and the support of teammates. Prior to each exercise, there was a discussion of professional evaluations and the need to focus on behaviors, not on personal characteristics. The students were required to list two positive traits and one aspect of team performance that could be improved for each member of the team. At the end of each semester in a final peer evaluation exercise, each team member conducted this evaluation process for a third time, assessing their own performance on the team, as well as that of other team members. For this third evaluation, each person on the team assigned everyone, including themselves, a score from 0.0 to 10.0. This score comprised 5% of their final grade in the course.

For the present study, data were analyzed from students in 22 sections, spanning semesters from spring 2002 through spring 2009. The data consisted of gender, self-evaluation score, evaluation score by peers (i.e. the other team members during the course) and final grade in numeric form. The ranges for the evaluations were from 0.0 to 10.0, while the range for the final grade was from 0.0 to 100.0.

Data Analytics

Data were screened for teams who had evaluation data from less than four team members, and these were excluded from analyses. From an initial 922 participants (248 women, 674 men), this procedure trimmed the study sample to 866 participants (232 women, 634 men). Following this screening, we ran descriptive statistics for all study variables and concluded that self-evaluation

score did not follow a normal distribution based on unacceptably large values for skewness and kurtosis. As a result, a non-parametric test was used to assess differences between gender on self-evaluation scores and final class grade. The specific test used was the Wilcoxon rank sum test. It assigned a rank to every student's "difference" score (i.e. the difference between self and peer evaluation) from lowest to highest, with ranks from 1 to 866 (the total number of people in the sample). The sample was then split by gender and the means of the ranks were compared. A significant difference in these means would indicate that two groups are significantly different from one another.

Results

A Wilcoxon rank sum test was used to compare the differences between self-evaluation scores and team evaluation scores. This test was conducted on our sample set ($n = 232$ for women, $n = 634$ for men, where n is the number of participants). The mean ranks from this test (415 for women and 440 for men) were not found to differ significantly, since the Z-statistic[1] was –1.29 with a significance level of $p = 0.20$. Significant differences would be attributable for $p < 0.05$. We also tested for gender differences in the final class grades, and again, the mean ranks (460 for women and 424 for men) were not found to significantly differ from each other—although these effects were marginal, since the Z-statistic was –1.91 and $p = 0.06$ in this case.

Post-hoc analyses

We decided to examine the final class grades of participants whose self-evaluation scores differed from their team evaluation scores by a margin of 5% (i.e. self-evaluation score was either 5% greater than team evaluation score or 5% below team evaluation score). Once again, we used a Wilcoxon rank sum test to compare the differences in final class grades between these two groups ($n = 155$ for those in the “5% greater” category, $n = 165$ for those in the “5% below” category). The mean ranks (135.6 for “5% greater” and 183.9 for “5% below”) were found to significantly differ ($Z = –4.67$, $p < 0.001$). We also tested for gender differences in final class grade, and again the mean ranks (460 for women, 424 for men) were not found to differ significantly, although these effects were marginal ($Z = –1.91$, $p = 0.06$). The strength of this relationship was measured using the Glass rank biserial correlation coefficient[17]; results indicated a moderate relationship ($r_g = 0.30$[2]) with those in the “5% below” group receiving higher final grades than those in the “5% greater” group.

Discussion

We were surprised by the results of our statistical tests with regard to self-evaluation as a function of gender. We had hypothesized that the female students would undervalue their contribution to the team, while the male students would overvalue their contribution. This valuation would manifest itself in the self-evaluation score being significantly different from the average score by the other team members. The statistical analyses indicate that there was little to

[1] A Z-statistic refers to the number of standard deviations that data, on average, fall above or below the mean on a normal distribution.
[2] Values for $r_g < 0.20$ tend to indicate a weak relationship, whereas values for $r_g > 0.60$ tend to indicate a strong relationship.

no difference between the male and female students in our study with regard to their self-evaluation. As a group, the female students did not, on average, undervalue their contribution to the team; and as a group, the male students, on average, did not overvalue their contribution to the team.

One of the reasons that we did not see a marked gender difference could be that the students at the institution where this study was conducted may exhibit less diversity in their career interests than at most other universities. It is a smaller school that focuses on engineering and science, and students in these first-year engineering design classes do not have many interactions with students who are not engineering or science majors. With few exceptions, comparable studies involving gender differences, as reported in the current engineering education literature, occurred at multi-purpose universities where the student body pursued majors in addition to engineering and science. We speculate that the "typical" student at the institution where this study took place, whether male or female, is more dedicated to the study of engineering. Because of the less diverse nature of those in attendance, these students may also possess a more uniform sense of themselves and their engineering abilities, regardless of their own gender.

The 5% higher group had a higher opinion of their performance on the team as compared to the evaluation of their teammates. This group also received significantly lower final grades than those students in the 5% lower group (i.e., those who had a lower self-evaluation than peer evaluation). It is possible that the members of the 5% higher group also maintain a higher opinion of their overall performance in the class, since the course grade measures more than just performance on a team. A large portion of the final grade is determined by communications activities, both written and graphic, and by technical creativity or innovation of a design solution. While the characteristics necessary to succeed in these activities could be indicative of or correlated with how well an individual contributes to a team, lower team performance (as perceived by fellow team members) could also be related to lower overall performance in the course. Therefore, students who perceive that they are doing better on a team as compared to the evaluation given by other members are the same ones that perform more poorly in the class as a whole.

It should be noted that this study has several limitations, and any resulting generalizations should be made with caution. Although the sample size was reasonable, all of the students were from a single institution and may not be representative of all college students majoring in engineering. Furthermore, the sample sizes of male and female students were unequal, and although this reflects the nature of engineering as a discipline (in both university and professional settings), it is more difficult to accurately compare the self-evaluations and actual performances of these students. Finally, there were a number of teams that were eliminated from the study, due to absences when the peer evaluation was conducted.

Despite these limitations, the present study contributes several important findings to the current literature on gender differences in engineering education. We did not find significant differences between male and female students in either self-evaluation or actual performance. While this finding is contrary to much of the current literature on this topic[3, 4, 6, 7, 8, 12, 15], we hypothesize that small, engineering-focused schools may provide better learning environments for female

engineering students than large public universities, such as increased opportunities for cooperative learning,[7] as well as a greater visibility of role models.

Future Studies

Several additional areas could be explored to enhance our understanding of the role of gender differences in self-evaluation, peer evaluation, and performance in an engineering design course:

1. We have compared our results to comparable measures for students at relatively large, multi-purpose universities, where male or female students with low self-efficacy in engineering could easily transfer to another major without transferring to another institution. Since this is not the case at the institution where this study was conducted, a logical next step would be to compare our results to peer institutions that are similar in both size and in having a nearly exclusive emphasis on engineering specialties.
2. Students provided written comments as well as numerical ratings for the peer evaluations. These comments might offer insight as to why the students who self-promote received lower grades in the course.
3. Peer evaluation exercises and ratings are continued in a second-year engineering design course at the same institution. Similar data could be collected and analyzed from the students in this second-year course to identify whether increased experience in college has any effect on team member self-efficacy and effectiveness.

Summary

1. The difference between self-evaluation and average peer evaluation scores did not differ as a function of gender. In other words, male students did not systematically show larger or smaller differences in these scores than female students.
2. Gender differences in final course grades were marginally significant; this could be a result of a low variability in final grades in this sample.
3. Students who self-promoted (i.e., had higher self-evaluation scores than the average peer evaluation score from their team) actually received lower course grades than those who under-valued their contribution to their team.

Bibliography

1. J. Bento, J. P. Duarte, M. V. Heitor and W.J. Mitchell, 2004. *Collaborative Design and Learning: Competence Building and Innovation*, International Series on Technology Policy and Innovation, Westport, CT, USA, pp. 1-5.
2. C. Tenopir and D.W. King, 2004. *Communication Patterns of Engineers*, IEEE Press and Wiley Interscience, New York, NY, USA, pp. 24, 50-51, 85-86.
3. B. Oakley, R.M. Felder, R. Brent, and I. Elhajj, 2004. Turning Student Groups into Effective Teams. Journal of Student Centered Learning, 2: 9-34.
4. R. Marra, K. Rodgers, D. Shen, and B. Bogue, 2009. Women Engineering Students and Self-Efficacy: A Multi-Year, Multi-Institution Study of Women Engineering Student Self-Efficacy. Journal of Engineering Education, 98: 27-38.
5. A. Bandura, 1977. Self- Efficacy: Toward a Unifying Theory of Behavioral Change. Psychological Review, 84: 191-215.
6. A. Carberry, H-S Lee, and M. Ohland. 2010. Measuring Engineering Design Self-Efficacy. Journal of Engineering Education, 99: 71-79.

7. R. Felder, G. Felder, M. Mauney, C. Hamrin, and E. Dietz, 1995. A Longitudinal Study of Engineering Student Performance and Retention. III. Gender Differences in Student Performance and Attitudes. Journal of Engineering Education, 84: 151-163.
8. M. Chemers, L.Hu, and B.F. Garcia, 2001. Academic Self-Efficacy and First-Year College Student Performance and Adjustment. Journal of Educational Psychology, 93: 55-64.
9. J.S. Hyde, 1996. Meta-Analysis and the Psychology of Gender Differences, in B. Laslett, S.G. Kohlstedt, H. Longino and E. Hammonds, eds., *Gender and Scientific Authority*, University of Chicago Press, Chicago, IL, USA: 302-320.
10. G. Sonnert, 1995. *Who Succeeds in Science?*, Rutgers University Press, New Brunswick, NJ, USA: 1-15, 138-157.
11. L. Schiebinger, 1997. Creating Sustainable Science, in S.G. Kohlstedt and H. Longino, eds., *Women, Gender and Science: New Directions.* Chicago, IL, USA: University of Chicago Press, pp. 201-216.
12. L. Rudman, 1998. Self-Promotion as a Risk Factor for Women: The Costs and Benefits of Counterstereotypical Impression Management. Journal of Personality and Social Psychology, 74: 29-645.
13. Babcock, L and S. Laschever, 2003. *Women Don't Ask.* Princeton, NJ: Princeton University Press, pp. 85-89.
14. Nye, 2000. Strong Objectivity and the Language of Science, in J. Bart, ed., *Women Succeeding in the Sciences – Theories and Practices Across Disciplines*, Purdue University Press, West Lafayette, IN, USA: 195-203.
15. J. Wolfe and E. Powell, 2009. Biases in Interpersonal Communication: How Engineering Students Perceive Gender Typical Speech Acts in Teamwork. Journal of Engineering Education, 98: 5-16.
16. S.V. Rosser, 1995. Introduction: Reaching the Majority, Retaining Women in the Pipeline, in S.V. Rosser, ed., *Teaching the Majority: Breaking the Gender Barrier in Science, Mathematics and Engineering*. Teachers College Press, New York, NY, USA: 1-21.
17. S.V. Rosser, 1995. Conclusion: Changing Curriculum and Pedagogy to Reach the Majority Results in a Positive Upward Spiral, in S.V. Rosser, ed., *Teaching the Majority: Breaking the Gender Barrier in Science, Mathematics and Engineering*, Teachers College Press, New York, NY, USA: 220-229.
18. T. Schultz, 1998. Practical Problems in Organizing Students into Groups, in *Proceedings of Frontiers in Education Conference*. Tempe, AZ, USA, 1: 242-245.
19. J. Jaccard and M.A. Becker, 2003. *Statistics for the Behavioral Sciences*, Belmont, CA: Wadsworth/Thomas Learning.