SELF-PERCEPTION DIFFERENCES BASED ON GENDER AND PERSONALITY TYPE IN TEAM PROJECTS

Peter M. Ostafichuk, Stephen Mattucci, Agnes d’Entremont, Navid Shirzad, and Carol Naylor
University of British Columbia
ostafichuk@mech.ubc.ca

Abstract Factors affecting student self-perception in the context of engineering design team work are examined in this paper. Specifically, how students of different gender and personality type rank their own contributions to their team relative to how their teammates rank their contributions is considered. Gender- and personality-based differences in self-serving bias – an individual’s tendency to attribute positive outcomes to their own actions and negative outcomes to external factors – are known to exist.

This two-part study examines these factors in the context of self-evaluation and peer evaluation scores received in a second year mechanical engineering design project course. Four evaluation events were conducted in January, 2015 (Part 1), and followed by an in-class intervention (presentation) and three more evaluation events in April, 2015 (Part 2). In Part 1, self-serving biases were measured by examining the difference between self-evaluation scores and average peer evaluation scores received from teammates. Separate t-tests and mixed-linear model statistical analysis were used to compare the average self-peer bias in evaluation scores versus gender and each of the four MBTI domain scales.

Data showed a statistically significant increase in self-serving bias as the course progressed. Differences were also noted for gender (males initially had a higher self-serving bias than females, but this difference disappeared in time), and MBTI domains of Introversion/Extraversion and Thinking/Feeling (students with a preference for Extraversion and Thinking had higher self-serving bias). The differences for gender and personality type were statistically significant with t-tests but not with mixed-linear models, suggesting the observed effects were driven by a small number of individuals with large self-serving bias. Following the intervention – consisting of a short in-class presentation describing the observed effects from Part 1 – reduced self-serving bias was observed in Part 2, but it is unclear if the change was due to the intervention or due to other factors.

Keywords: teamwork, peer evaluation, self-evaluation, self-serving bias, self-efficacy, personality type, gender, Myers-Briggs

1. INTRODUCTION

One of the most important aspects of an undergraduate engineering curriculum are the culminating projects, which present unique opportunities for students to learn to problem solve as a group and develop teamwork skills. This study centres around differences between individuals’ perceptions of their own contributions to their team and the perceptions of the rest of the team towards those same contributions.

Self-efficacy refers to an individual’s confidence in successfully completing a given set of tasks [1]. Self-efficacy can have a positive effect amongst individuals, as it correlates to work-related performance [2]. Self-serving attribution bias (or simply self-serving bias) occurs when an individual attributes success to one’s own actions, and attributes failures to the actions of others [3]. This effect can be especially pronounced in groups with established leader-member dynamics, as there is an obvious target for blame [4]. These effects can also be reversed depending on one’s attribution style; most commonly either optimistic or pessimistic. Optimistic refers to attributing success to internal and stable factors (i.e. self), and failures to external and unstable factors (i.e. others), where pessimistic refers to attributing success to external factors, and failures to internal factors [5]. Self-efficacy, performance, and self-serving attribution bias are all interrelated, as self-serving attributions increase self-efficacy which further enhances performance [6].

There is little in the literature that links these cognitive effects to other factors such as personality type or gender; however, they can be investigated using self- and peer evaluation in a university-level project group setting.

Peer evaluation is a commonly used tool to foster better team experiences and to equitably recognize individual student’s contributions to their team’s success [7]. Peer evaluation can be used in a formative way (not for marks), a summative way (at the end of an experience for marks), or some blend of the two. Formative peer evaluation is generally used to encourage positive teaming behaviours and reform or decrease poor behaviors. Summative peer evaluation helps to temper group grades and to ensure that students get what they deserve – highly performing students that contribute well to their team are rewarded and students that do not
The behaviorally anchored rating scales (BARS) approach is a commonly used peer evaluation method. This method uses an evaluation instrument with predefined criteria and descriptors (anchors) for different levels of performance in those criteria [9],[10]. This focuses the evaluation on specific factors important for team success, it helps communicate constructive team behaviors to students, and aims to reduce subjectivity and improve inter-rater reliability [10],[11]. The scores are usually accompanied by written comments. Evaluatees typically receive their aggregate score and/or comments in an anonymous, randomly-ordered fashion at the end of the evaluation period, end of the course, or not at all.

In evaluating others, biases depending on evaluator and evaluatee gender, personality type, and culture have been shown to exist. In a meta-analysis of leaders and managers, Eagly et al. showed men’s leadership and agentic behavior were evaluated more favorably than those of women [12]; effects of role congruity were suggested as the source of prejudice towards female leaders [13]. Harsh reported a gender bias amongst 290 senior undergraduates in business in which evaluators tended to give higher evaluations, both in terms of performance and leadership, to managers of the same gender [14]; females received higher overall evaluations than males, in contrast to the findings of Eagly. In peer evaluations, Ruble showed the presence of positive same-gender bias [15]. In terms of personality influence, in a study of 196 sophomore and junior business communications students, May and Gueldenzoph [16] demonstrated a dependence between peer evaluation and personality, in this case described according to social style theory. In a separate study with 144 managerial communications students, May demonstrated a bias where evaluators tended to give more favourable peer evaluation scores to teammates with the same personality type, again measured by social style [17]. Additionally, in an eight year study, the authors have shown statistically significant differences in peer evaluation scores received based on certain gender-Meyers Briggs Type Indicator (MBTI) combinations [18]. Finally, individuals of a common culture tend to make self-serving attributions regarding the performance of each other, whereas they make unfavourable attributions about members from other cultures [19],[20].

This paper will explore the presence of self-serving attribution bias, as measured by the difference in self-evaluation score given to average peer-evaluation score received in an intensive second year mechanical engineering design course. In addition, the influence of gender and personality type, as measured by the MBTI, on self-serving bias will be examined. This study extends the existing body of knowledge by addressing four main research questions:

- Is self-serving bias present in an engineering design course, and, if so, how does it change with time,
- Does the evaluatee’s gender influence self-serving bias,
- Does the evaluatee’s personality type influence self-serving bias, and
- Is it possible to mitigate effects of self-serving bias by delivering a short intervention to raise student awareness of the potential for self-serving bias

The methodology for this study is outlined in the following section, including the course context, the project teams, peer evaluation instrument used, the subjects, and the experimental methods. The results of the statistical analyses are then provided in four parts: for all students together, for comparing effects of gender, for comparing effects of personality type, and for assessing the impact of the intervention. The paper finishes with discussion of the results and conclusions.

2. METHODOLOGY

2.1. Course Context

This study was conducted 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 [21]. The 2015 cohort, consisting of 121 students, was considered in this study. The course is delivered using the Team-Based Learning (TBL) approach [22],[23], with course-specific details regarding this TBL implementation extensively documented [24],[25],[26]. All students attend a common lecture section (i.e. there are 121 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, the course is seven weeks in duration and, other than an integrated course in technical communication, 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). Each design project culminates in a class-wide competition, putting extra stressors on teams to not only complete the project...
on time, but perform competition objectives to a high degree.

2.2. Project Teams

Following recommended practice, teams of six to seven students (20 teams each year) were instructor-formed [27] in order to maximize diversity [28],[29] and to minimize previously established subgroups [22],[29]. 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. With the exception of the Feeling preference, the remaining seven MBTI preferences were represented by at least two members on each team; there were not enough students reporting a preference for Feeling to ensure they were represented on 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.

2.3. Peer Evaluation Instrument

As part of course requirements, students completed six mandatory peer evaluations using the iPeer online software tool (http://ipeer.cltl.ubc.ca/). Teams used a BARS peer evaluation approach (i.e. rubric-based evaluation) in which they evaluated themselves and their teammates on four criteria:
- Communication: the individual communicates effectively with other team members
- Initiative: the individual displays initiative and contributes to team management and goal setting
- Responsibility: the individual assumes responsibility for own work, and participates equitably
- Professional behavior: the individual arrives to team meetings and class on time; completes work to a professional standard

Evaluations were completed using four levels of mastery: below expectations, marginal, meets expectations, and exceeds expectations. Before the first evaluation, teams worked through a scheduled activity in which they defined for their team the specific behaviors and evidence they would look for to assign ratings in each of the categories to clarify team expectations for each member. The raw numerical scores for the completed BARS evaluation were normalized on an evaluator-by-evaluator basis to an average of 100 (i.e. each evaluator’s average score given was normalized to 100). The evaluations required the students to provide comments to justify the scores they assigned. After each evaluation, students received aggregate scores and anonymous, randomly-ordered comments from their teammates. For each student, the average peer evaluation score 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. The self-evaluations completed on iPeer did not influence students’ grades.

2.4. Subjects

There were 121 students in the course. The gender distribution and distribution of reported MBTI preferences is shown in Table 1. Not all students completed the team formation survey so summing the numbers for MBTI preference pairs in the table yields a slightly lower value than 121.

<table>
<thead>
<tr>
<th>MBTI</th>
<th>Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
</tr>
<tr>
<td></td>
<td>I</td>
</tr>
<tr>
<td></td>
<td>S</td>
</tr>
<tr>
<td></td>
<td>T</td>
</tr>
<tr>
<td></td>
<td>J</td>
</tr>
<tr>
<td>Introversion</td>
<td>65</td>
</tr>
</tbody>
</table>

2.5. Experimental Method and Intervention

Peer evaluations were conducted on a weekly basis in the course. Evaluations were completed online, as described in Section 2.3, and were available for a total of two weeks. The first evaluation was released on the first Tuesday in the course (Week 1) and due the second Tuesday (Week 2) at 8:00 am. Evaluatees were penalized for late evaluations in order to motivate timeliness. All evaluations followed the same pattern of one week deadline plus one week extended deadline with grade penalty. All analysis in this study is based on raw evaluation scores (i.e. without late penalties applied). Students did not receive feedback from their peers on an evaluation until they had submitted their own evaluations and the deadline had passed.

As described in Section 2.1, the course is split into a four-week portion in January and a three-week portion in April. At the start of the April portion of the course, a short in-class presentation (the “intervention”) was given to the students. It consisted of data on self-serving bias drawn from the January 2015 portion of the course and summarized in Sections 3.1 to 3.3 below. The presentation was delivered during a workshop on team
dynamics and it was emphasized that the findings were trends across large groups and were not predictive of any one person’s behaviour. It was also emphasized that there was nothing inherently positive or negative about any of the results. For example, it was described that a positive self-serving bias could originate from a high degree of self-confidence or from someone who put in a lot of effort to the projects during time spent away from the team; similarly, a negative self-serving bias could originate from a high degree of humility or from someone who emphasized the team over the individual.

Peer evaluation data at the conclusion of each period was processed using the statistical methods described below.

2.6. Statistical Methods

Statistical analyses were performed two-tailed t-tests (in Excel) and mixed-linear models (in STATA, StataCorp, College Station, TX). The t-tests were done individually by time point but did not consider that the data was based on repeated measurements with the same people; the mixed-linear models are more robust as they account for the fact that repeated measures over time with the same person are related to each other. Gender and MBTI domain were included as separate random effects in the analyses. In STATA, observations were nested within individual students. The responses from each student were individually modelled as a linear function, allowing a repeated-measures type analysis, and combined into a mean model, thus providing an intercept and slope. The statistical modelling in STATA allowed for missing data; therefore, students without TypeFocus results were included in the analysis – they contributed to the mean results where specific observations were missing. In addition, STATA included weighting of data based on proximity to the mean model values such that the influence of outliers was diminished. A Bayesian information criterion (BIC) was used in STATA to determine optimal models (where lower BIC indicates a better model). The optimal models by BIC allowed individual model results to vary in both slope and intercept within analyzed groups. A 5% significance level was used for all analyses.

3. RESULTS AND DISCUSSION

The results of the statistical analyses are provided in four parts: for all students together, for identifying effects of gender, for identifying effects of personality type, and for assessing the impact of the intervention. The first three parts consider the January portion of the course only, while the intervention considers January and April.

3.1. Presence of Self-Serving Bias

Self-serving bias was observed in evaluations. Over the first four evaluations, the bias was initially small but grew at a rate of several percent per evaluation (i.e. per week) on average, as shown in Fig. 1. It remained positive at all times, indicating individuals rated their personal contributions more favourably compared to how their teammates rated those contributions, on average. Put another way, for all results taken on average, individuals rated their own contributions increasingly more favourably over time than they rated the contributions of their teammates. This finding was statistically significant (mixed-linear model slope, $p = 0.003$).

![Fig. 1. Presence of self-serving bias](image)

The initially small self-serving bias is consistent with expectations from Tuckman’s forming-storming-norming-performing model of team development [31] in that the initial team interactions tend to focus on politeness, promoting acceptance, and not upsetting team harmony. The presence of self-serving bias in rating individual contributions to team function was also observed previously by the lead author through a different self- and team-evaluation instrument [30].

3.2. Self-Serving Bias and Gender

The differences in self-serving bias based on gender are shown in Fig. 2. There are three observations of note in this data. First, the same upward trend in self-serving bias with time is observed for both male and female students. Next, the male students initially start with a small positive self-serving bias while the female students start with a modest negative self-serving bias. In other words, male students initially give themselves higher evaluations than those they receive from their peers, while female students give themselves lower evaluations than those they receive from their peers. Finally, the change in self-serving bias is more pronounced for the female students and, by the end of the first project, they have a
higher self-serving bias than the male students. The observed difference in evaluation 1 was statistically significant in the t-test (p = 0.003) but not with the mixed-linear model (difference in intercept, p = 0.11). Other differences were not statistically significant.

These results are of particular interest given the effect of time on self-serving bias. There have been numerous studies that show males tend to have a higher self-serving bias [6],[19],[20] (which agrees with the first and second evaluations) none of these studies have investigated the effect of time. The lack of statistical significance indicates additional data is required to verify the trends in Fig. 2 suggesting a time-dependence on the gender differences in self-serving bias.

3.3. Self-Serving Bias and Personality Type

In terms of personality type, measured by the MBTI, statistically significant differences in self-serving bias were observed between students with a preference for introversion and those with a preference for extraversion, as shown in Fig. 3. In particular, the students with a preference for extraversion (characterized by an outward focus) tended to have a higher degree of self-serving bias on average compared to those with a preference for introversion (characterized by an inward focus). The extraversion preference is associated with a comfort in working with others and an interest in active involvement while the introversion preference is associated with an interest in ideas and concepts and comfort in working alone or in small groups. Differences in evaluations 1, 2, and 4 were statistically significant with the t-test (p = 0.02, 0.02, and 0.05, respectively) but differences in slopes (p = 0.68) and intercepts (p = 0.18) were not significant with the mixed-linear models.

Fig. 2. Influence of gender on self-serving bias

Differences were also noted between the MBTI thinking and feeling domains, as shown in Fig. 4. The thinking domain is characterized by decision-making based on logic and objectivity while the feeling domain is characterized by decision-making based on values and subjectivity. The differences for evaluations 1, 2, and 3 were statistically significant (p = 0.008, 0.01, and 0.02, respectively) for the t-test but differences in slopes (p = 0.93) and intercepts (p = 0.12) were not significant for the mixed-linear models.

Fig. 3. Influence of MBTI I-E on self-serving bias

Fig. 4. Influence of MBTI T-F on self-serving bias

Statistically significant differences were not observed for either the sensing-intuition or the judging-perceiving domains of the MBTI under either model.

3.4. The Influence of an Intervention on Self-Serving Bias

Following the intervention given in the team dynamics workshop to start Project 2 in April, changes in self-serving bias were noted. As shown in Fig. 5, after the intervention, self-serving bias reduced to the same levels observed at the start of the course as the teams were just formed. One week into the second project, bias then returned to a similar level as the conclusion of Project 1. Interestingly, for the final evaluation, self-serving bias again returned to the near-zero levels at the start of each
project. The presence of statistically significant self-serving bias (i.e. self-peer bias different from zero) was observed for evaluations 3 ($p = 0.004$), 4 ($p \leq 0.001$), and 6 ($p = 0.003$) only, based on the mixed-linear models. The same results were also found using t-test analysis, but with different p-values.

The intervention is one possible cause for the change in self-serving bias. Another possible explanation is the long break (9 weeks) between evaluations 4 and 5, since evaluations 1-4 were for Project 1 and evaluations 5-7 were for Project 2. Given the marked decrease in self-serving bias in evaluation 7, it is likely that the project stage plays a significant role in the trends shown. Early in each project, time pressures and tensions between members tend to be much lower. The drop at the end of Project 2 not seen in Project 1 may be attributed to the fact that evaluation 7 was completed at the end of the course and the end of the term, and for most students, after final exams had been written.

4. CONCLUSIONS

A comparison of self-evaluation and peer evaluation scores was used to measure self-serving bias in a team-based second year mechanical engineering design course. Considering the research questions of this study

- Self-serving bias was found to be present and measurable, and was found to increase with time during a four-week team project.
- Statistically significant differences in self-serving bias between male and female students were observed in the first peer evaluation in the course, but subsequent differences were smaller and not statistically significant.
- Statistically significant differences in self-serving bias between students with preferences for Introversion and Extraversion, as well as those with preferences for Thinking and Feeling, were found to exist. Extraversion and Thinking had higher self-serving bias.
- The short in-class intervention was followed by a measured reduction in self-serving bias. It is unclear if the observed effect was due to the intervention or simply due to the normal changes in stresses and tensions that would be expected with a demanding team project.

The findings for differences based on gender and personality type were significant for t-tests but not mixed-linear models, suggesting effects were dominated by a small number of individuals. Further analysis on additional cohorts will be required to verify the significance of the findings.

Future steps to undertake include comparing self-serving bias to performance (expected grades), as there may be a relationship over time with self-efficacy and performance, resulting in a positive feedback loop, increasing self-serving bias. More precise tracking of the timing of evaluations relative to major course milestones (i.e. end of project, completion of exams, etc.) will help to correlate observed bias and student stress.

References


