EXAMINING GENDER AND NATIONALITY BIAS IN DECISION-MAKING BY ENGINEERING STUDENT TEAMS

Peter M. Ostafichuk, Masoud Malakoutian, and Mahsa Khalili
Department of Mechanical Engineering, University of British Columbia
ostafichuk@mech.ubc.ca

Abstract – This study uses two-stage team quizzes to assess differences in team decision-making based on the factors gender and nationality. Over 200 teams in two different engineering design courses delivered using Team-Based Learning across five years were considered. In the two-stage quizzes, individuals first committed to their own answers, and then the team discussed the same questions and answered as a group. Cases where an individual was incorrect and the team adopted that same incorrect answer were used as a measure of influence of that individual on team decision-making (i.e., “pushing” behaviour by the individual). Similarly, cases where an individual was correct but the team adopted a different (incorrect) answer were used as a measure of lack of influence (i.e., “switching” behaviour by the individual). Overall, no significant gender or nationality differences were found in pushing behaviours. Male students and international students were found to be more likely to engage in switching behaviours. The overall differences in switching were modest (0.3-0.4% difference per question), but this translates to between 5 and 15 more male/international students engaging in switching behaviours in a typical 75- to 150-student course.

Keywords: bias, gender, nationality, engineering teams, decision-making, two-stage quizzes, Team-Based Learning

1. INTRODUCTION

Diversity refers to the collection of all ways people are different and unique, and includes factors of gender identity, race, ethnicity, beliefs, age, personality type, and more. Teams benefit from diversity, as more diverse teams produce better work and make better decisions than less diverse teams [1]-[3]. A diverse team has a broader range of perspectives and experiences to draw from. However, diversity can also lead to bias; this is the unequal treatment of some individuals compared to others based on the same factors that lead to diversity. Through bias—whether it is explicit or implicit—some team members can be made to feel marginalized or less included. This not only impacts those individuals, but it also diminishes team cohesiveness, and, in turn, team effectiveness [4]. The presence of bias in post-secondary course teams is well documented. Macke et al. [5] showed that race played a role in peer evaluations in a Team-Based Learning (TBL) course in a manner inconsistent with academic performance. This impacted students’ feelings of isolation from the team [6]. Others have documented similar results for bias due to gender, socioeconomic status, and other factors [7],[8].

Related to team decision-making in TBL courses, in a study of the development of team trust in a TBL course, Preast [9] noted that absolute performance was less important than a team member’s “ability to help the team choose the correct answer and understand the material.” She used the example of an academically low performing student who knows when to speak up with useful information (and when to remain silent) as being more trustworthy in decision-making than an academically high performing student who interrupts teammates or fails to indicate when they are unsure about an answer. Additionally, Preast found on TBL teams that the number of members with the same answer is one of the most important factors in determining which answer a team ultimately selects in decision-making. One way a correct individual could sway the team appeared related to that person’s ability to clearly communicate a strong rationale for an answer. Another way an individual could sway the team appeared related to that individual’s ability to question or challenge other’s ideas. Considering the likelihood of individuals switching answers on TBL quizzes, Reinig [10] found students were more likely to do so after the first quiz in a course, but found no significant effects related to gender or English fluency.

The current study seeks to confirm the absence of bias in team decision-making due to factors of gender and nationality. Nine cohorts in two different TBL courses (comprising over 1100 students on over 200 teams, and responding to over 800 quiz questions) are considered using quantitative methods. Team decision-making is evaluated through the responses to team quizzes. In these quizzes, team members were required to reach consensus on answers to graded multiple-choice questions.
Differences between what each individual student felt the correct answer was and which answer the team ultimately selected was used to assess bias in the team’s decision-making.

2. RESEARCH QUESTIONS

There are four primary research questions considered in this study related to factors of gender and nationality in team decision-making:

- RQ1: Does an individual’s gender influence the likelihood that a team will follow that individual in making an incorrect decision?

- RQ2: Does an individual’s nationality influence the likelihood that a team will follow that individual in making an incorrect decision?

- RQ3: Does an individual’s gender influence the likelihood that their team will make an incorrect choice in decisions where the individual knows the correct choice (i.e., not follow the individual).

- RQ4: Does an individual’s nationality influence the likelihood that their team will make an incorrect choice in decisions where the individual knows the correct choice.

To support the four research questions above, this study also explored general trends based on timing of quizzes in the term and course.

3. CONTEXT

This study was conducted in two different TBL courses at the University of British Columbia. Decision-making behaviours were determined through individual and team responses on two-stage quizzes in both courses. The mechanics of the two-stage quizzes, the details of the two courses, and a summary of the demographics of the students in the courses are summarized below.

3.1. Two-stage quizzes

The courses considered in this study are taught using Team-Based Learning [11],[12]. This is a flipped classroom approach [13] in which students participate in a two-stage [14] multiple-choice readiness assurance test (RAT) at the start of each module. The first stage of the quiz is done individually and is commonly referred to as the iRAT (individual readiness assessment test); the second stage repeats the same questions as the first, but is done in teams and is referred to as the tRAT (team readiness assessment test). Both RATs are closed book, and are based on content from readings assigned in advance. RAT questions most commonly draw from lower levels of Bloom’s Taxonomy [15], including remembering, comprehending, and applying. For the two courses considered in this study, five item questions (a single correct answer and four distractors) were used. In the iRAT, students were provided a paper copy of the questions and then indicated their answers on a computer scan card (Apperson Datalink 1200 compatible). Students were allowed to add annotations to their test paper (such as to circle the answers they selected or to add questions or notes). At the end of the iRAT stage, students submitted their computer scan cards but held onto their test paper. Teams were then given an Immediate Feedback Assessment Technique (IF-AT) [16] “scratch card,” similar to that shown in Fig. 1, to use for the tRAT. Following the tRAT, teams handed in their IF-AT and all test papers.

![Fig. 1. IF-AT "scratch" card](image)

The use of the IF-AT cards was as follows:

1. The team discusses a tRAT question and reaches consensus on what they believe to be the correct answer. They scratch the corresponding box on the card.

2. If the team is correct (such as with choice C for question 1 in the figure), they will see a star, and they then move onto the next question.

3. If the team is incorrect (such as with choice B for question 2), they will see a blank. At this point, they discuss the question and answer choices again, and then they scratch a second time (such as with question 2 choice E).

4. The score the team receives is dependent on how many scratches it takes them to identify the correct answer. In these courses, scoring was 4 points for 1 scratch, 2 points for 2 scratches, 1 point for 3 scratches, and 0 points for 4 or 5 scratches.

The key pedagogical benefit of this technique is that students immediately receive feedback on whether or not they correctly answered a question, and if not, they immediately return to discussing this topic until they do get the correct answer. In addition, this technique provides an
opportunity for collaborative learning, where students teach each other based on their own understanding and emerging competency; it provides feedback to the instructor on student understanding; it provides a grade and a social incentive to come to class having reviewed the preparatory material; and most students find it fun and engaging.

In this study, the IF-AT cards provided a record of team decision-making. For example, referring to Fig. 1, the IF-AT card shows the team made a correct decision on Question 1, but an incorrect decision on Question 2, which was later corrected with feedback. The specific interpretation and use of the IF-AT data in this study is examined further in Section 4, after a review of the two courses in which this study was based.

3.2. Course A: MECH 223

MECH 223 (hereafter referred to as Course A) is a second-year mechanical engineering design course with two large team design projects [17]. It focuses on the topics of design process and design tools, but also covers project management, team functioning, effective communication, and so on. The course is part of the Mech 2 Program [18], and it operates full-time for four weeks in January and again for three weeks in April.

There are 20 teams in the course, with a nominal team size of six students per team (occasionally there are teams of five or seven). Following best practices, teams are formed by instructors [17],[20] with the goal of maximizing heterogeneity [21] across dimensions such as academic ability, personality type, practical skills, and so on [22]. Also following recommended practice in engineering, female students are not placed alone on teams [20]. Nationality data was neither available nor used in team formation. The course demographics for Course A are shown by year in Table 1. In the table and throughout the paper, F and M are used to refer to female and male, respectively, and C and I are used to refer to Canadian nationality and international (non-Canadian) nationality. Totals do not always add to 100% as demographic data were not available in all cases for all students. (Note: data files from the individual quizzes in 2017 were corrupted, and that year is not considered in this study.)

| Year | N   | F   | M   | C   | I   |
|------|-----|-----|-----|-----|-----|
| 2015 | 121 | 28  | 93  | 90  | 31  |
|      |     | (23%)| (77%)| (74%)| (26%)| |
| 2016 | 126 | 33  | 92  | 83  | 42  |
|      |     | (26%)| (73%)| (66%)| (33%)| |
| 2018 | 129 | 35  | 94  | 87  | 42  |
|      |     | (27%)| (73%)| (67%)| (33%)| |
| 2019 | 125 | 18  | 107 | 87  | 38  |
|      |     | (14%)| (86%)| (70%)| (30%)| |

3.3. Course B: MECH 326

The second course considered in this study, MECH 326 (hereafter referred to as Course B), is a third-year machine design course. It focuses on design for fracture and fatigue prevention, shaft design, weld design, and an introduction to finite element analysis in design. It is scheduled in a traditional manner, with two 50-minute classes per week, and two 80-minute tutorials per week. The course begins with several weeks of review content, and then is divided into five modules corresponding to the topics above. Each module has assigned readings from the course textbook (Shigley’s Mechanical Engineering Design, [24]), a series of active learning classes, and an associated team mini-project.

The nominal team size in Course B is 4 or 5 students. As with Course A, teams are instructor-formed and efforts are made to ensure there are not lone female students, but other steps from Course A to maximize heterogeneity are not employed. In order to make it more convenient for students to meet outside of scheduled class time, students are assigned to teams with others with similar schedules, which means placing students with others from the same program or option. The course demographics for Course B are shown by year in Table 2.

| Year | N   | F   | M   | C   | I   |
|------|-----|-----|-----|-----|-----|
| 2015 | 202 | 41  | 160 | 150 | 51  |
|      |     | (20%)| (79%)| (74%)| (25%)| |
| 2016 | 98  | 22  | 76  | 76  | 22  |
|      |     | (22%)| (78%)| (78%)| (22%)| |
| 2017 | 98  | 22  | 76  | 68  | 30  |
|      |     | (22%)| (78%)| (70%)| (31%)| |
| 2018 | 155 | 34  | 110 | 108 | 36  |
|      |     | (24%)| (76%)| (75%)| (25%)| |
| 2019 | 98  | 25  | 73  | 58  | 40  |
|      |     | (26%)| (74%)| (59%)| (41%)| |

3.4. Nationality data

Using citizenship data, student nationality was categorized by amalgamating regions from the United Nations geoscheme and M49 coding classification. Table 3 lists the regions used and the number of students from each region, by course. The largest contributing country is listed for each region. It is important to recognize that nationality is not the same as ethnicity or race. As mentioned above, nationality was not considered in team formation and can be considered random across teams.
Table 3: Nationality data by course

| Region                      | Largest      | [A] | [B] |
|-----------------------------|--------------|-----|-----|
| N America                   | Canada       | 353 | 466 |
| E Asia                      | China        | 75  | 79  |
| S Asia                      | India        | 20  | 26  |
| SE Asia                     | Malaysia     | 22  | 20  |
| W Asia                      | Iran         | 8   | 12  |
| W Europe                    | UK           | 2   | 15  |
| S & Cen. America            | Mexico       | 8   | 8   |
| Africa                      | Egypt        | 7   | 9   |
| E Europe & Cen. Asia        | Kazakhstan   | 5   | 4   |

4. METHODS

Data in this study were analyzed using IBM SPSS 25 and 26. Differences between groups were determined using an independent samples t-test for comparison of means. Levene’s test for equality of variances was applied in these cases, and p-values reported are based on the outcome of that test (unequal variances between groups was used when Levene’s test was significant, and equal variances were used otherwise).

There were two main undesirable team decision-making behaviours considered in this study: when an individual convinced the team to adopt an incorrect answer, and when an individual knew the correct answer yet allowed the team to select an incorrect answer. To describe these cases, the terms “pusher” and “switcher,” borrowed from switching behaviour theory, were adopted.

- **Pusher** was the term used to describe team members who had the same incorrect answer on an iRAT that the team had on the tRAT. The interpretation was that that individual had influence on the team’s final decision on the tRAT, and may have swayed the team towards the wrong answer. The pusher score used throughout this paper represents the percent chance on a single question that an individual would push their team to an incorrect decision. The pusher score was reduced proportional to the number of team members who had the same incorrect answer on a particular iRAT and tRAT question pair. (In other words, if only one student on a team had the same incorrect iRAT and tRAT pair, they would have a pusher score of 1; however, if four students had that same incorrect pair, they would each have a pusher score of 0.25.)

- **Switcher** referred to when an individual had the correct answer on the iRAT, but their team had an incorrect answer on the tRAT. The interpretation in this case was that the individual knew the correct answer, but they switched their response and allowed the team to submit an incorrect answer, presumably based on the influence of pushers. The switcher score used represents the percent chance on a single question that an individual would switch from a correct iRAT answer to an incorrect tRAT answer. It was not adjusted based on the number of switchers for a given question.

Pushing behaviour was related to research questions RQ1 and RQ2, and switching to RQ3 and RQ4. Analysis was conducted to determine the following:
- The relationship between pushing and switching behaviours and iRAT grades
- The trend in pusher score and switching score by quiz number, separately for Courses A and B
- Differences in pusher score and switcher score for female and male students, overall and by quiz
- Differences in pusher score and in switcher score for Canadian and international students, overall and by quiz number
- Differences in net score (pusher – switcher) based on gender, course, and nationality.

5. RESULTS AND DISCUSSION

Results and discussion relating to the analyses listed above are presented below in Sections 5.1 to 5.6.

5.1. Influence of Academic Performance

The results of a multiple linear regression of pusher and switcher scores versus individual average iRAT grades and team average iRAT grades was conducted. The results are shown in Table 4 and reveal that an individual’s pushing and switching behaviours are weakly related to that person’s iRAT grades, and strongly related to average iRAT grades of all students on that person’s team.

Table 4: Regression analysis with performance

| Coeff. | Pusher Score | Switcher Score |
|--------|--------------|----------------|
| $B_0$  | 9.305 (p < 0.001) | 8.059 (p < 0.001) |
| $B_{indiv}$ | 0.010 (p = 0.002) | 0.025 (p < 0.001) |
| $B_{team}$ | -0.117 (p < 0.001) | -0.101 (p < 0.001) |
| $r$    | 0.348        | 0.166          |

The regression from above is of the form

$$Score = B_0 + B_{indiv}(iRAT) + B_{team}(iRAT \text{ for team})$$

The better an individual does on the iRATs, the more likely they are to be a pusher on the tRATs ($B_{indiv}$ is positive with pusher score). This could be due to them being more confident in their answers or the team trusting them more. However, increased iRAT grades are also positively correlated with more switching behaviour (again, $B_{indiv}$ is positive for switching), which is inconsistent with a higher
degree of self-confidence, and suggests it is the latter explanation (the team trusts high performers more) that is at play. On the surface, this seems to be in contrast to the conclusions of Preast [9]. Furthermore, the better a person’s team does collectively on the iRATs, the less likely individuals on that team are to engage in pushing and switching. This makes sense in that the more questions a team gets correct, the fewer questions there are where pushing or switching can take place.

The above regressions were generally similar when considering males and females separately. Of note, pushing behaviour for women was more strongly dependent on their own iRAT performance (women: B_{indiv} = 0.017, p = 0.010 versus men: B_{indiv}=0.007, p = 0.042), and switching behaviour for men was more strongly dependent on the average team iRAT performance (B_{team} = -0.104, p < 0.001 versus B_{team} =-0.087, p < 0.001 for women). Similarly, considering Canadian and international students, Canadian students’ pushing behaviours were less dependent on team performance (B_{team} = -0.110, p < 0.001, versus -0.132, p < 0.001 for international), and Canadian students’ switching behaviour was less dependent on their own performance (B_{indiv} = -0.021, p < 0.001 versus 0.038, p < 0.001).

5.2. Pushing and Switching Behaviour by Quiz Number

The trends by quiz for Course A and Course B are shown for pusher behaviour (Fig. 2) and switcher behaviour (Fig. 3). Both courses show a similar pattern with a peak in pusher and switcher behaviour in Quiz 2. The fact that trends in the two figures appear to be correlated makes sense in that more pushing behaviour should tend to lead to more switching.

In terms of the peak observed for Quiz 2, there are at least three possible explanations. It is possible that Quiz 2 in each course is more difficult than Quizzes 1 and 3, and the observed peak is related more to quiz difficulty than to group behaviour. Alternatively, students may be better prepared on Quiz 1, due to the additional time to review material in advance, and then they may still be adjusting to how to best study for the quizzes at the time of Quiz 2. These two explanations would suggest quiz grades should be lower on Quiz 2 than on Quizzes 1 and 3. This is indeed consistent with average quiz grades in Course B, but not Course A. A third possibility is based on changing team dynamics as team members learn to work together. Tuckman’s model of team development [25] would suggest that during Quiz 1 teams would be in the Forming stage and team members would be polite and reserved, tending not to challenge one another. The Storming stage would be expected to roughly align with Quiz 2, and would involve more conflict between teammates and more willingness to challenge each other. As teams move to the Norming and Performing phases, team norms would develop, and mechanisms for resolving differences would emerge, and mutual trust would increase.

The other observation is that there is marked increase in pusher behaviour in Course B compared to Course A. The average unscaled iRAT grades are lower in Course B than Course A (approximately 62% and 74%, respectively), corresponding to a 46% increase in the number of iRAT errors in Course B. However, the increase in the pusher values from Course A to Course B is roughly three times this amount, suggesting there is also a change in student behaviours and/or attitudes that occurs between the second- and third-year courses.

5.3. Gender differences by quiz

The gender differences in pusher and switcher scores by quiz number (for both courses combined) are shown in Table 5. There is a statistically significant gender difference on Quiz 1 in terms of pusher behaviour (males are more likely to be pushers than females). This effect was observed in Course B (average F – M difference on Quiz 1 was -0.814, p = 0.006), but not Course A (average F – M difference was -0.006, p = 0.947). Overall, as shown in the last row, no significant gender difference in pushing behaviour was observed.

In terms of switching behaviour, an overall difference was observed, with males more likely to switch than females. A significant difference was noted on some quizzes (2 and 5, both with males more likely to switch)
but not others. Considering the courses separately, in Course A, a significant overall difference was observed (F - M difference was -0.397, p = 0.014), while for Course B, a similar magnitude difference was not significant (-0.372, p = 0.123). Overall, the data suggests there is a slight gender difference in switching behaviour.

Table 5: Gender dependence of pusher and switcher behaviours for both courses (Quiz 6 for Course A only)

| Quiz | Pusher | Switcher |
|------|--------|----------|
|      | F - M  | p-value  | F - M  | p-value  |
| 1    | -0.588 | 0.003    | -0.303 | 0.322    |
| 2    | +0.317 | 0.126    | -0.859 | 0.020    |
| 3    | -0.000 | 0.999    | +0.107 | 0.717    |
| 4    | +0.254 | 0.218    | -0.100 | 0.748    |
| 5    | -0.339 | 0.119    | -0.649 | 0.027    |
| 6    | -0.018 | 0.794    | -0.127 | 0.472    |
| All  | -0.067 | 0.585    | -0.386 | 0.013    |

5.4. Nationality differences by quiz

A similar analysis was completed for nationality at the level of Canadian/international and is shown in Table 6.

Table 6: Nationality dependence of pusher and switcher behaviours for both courses (Quiz 6 for Course A only)

| Quiz | Pusher | Switcher |
|------|--------|----------|
|      | C - I  | p-value  | C - I  | p-value  |
| 1    | +0.189 | 0.345    | -0.243 | 0.427    |
| 2    | -0.143 | 0.451    | +0.053 | 0.877    |
| 3    | -0.104 | 0.521    | -0.305 | 0.264    |
| 4    | +0.318 | 0.094    | -0.321 | 0.264    |
| 5    | +0.189 | 0.345    | -0.819 | 0.007    |
| 6    | -0.004 | 0.952    | -0.125 | 0.437    |
| All  | +0.089 | 0.427    | -0.308 | 0.032    |

In this case, no statistically significant differences in pusher behaviour based on nationality were observed. A difference in pushing behaviour was observed on Quiz 5 in Course A, but this does not appear to be part of a trend as the overall average C – I differences by course were +0.082 (p = 0.233) for Course A and -0.019 (p = 0.900) for Course B. A statistically significant difference in switching behaviour was observed overall, with international students more likely to switch than Canadians. The differences in Course A and Course B were -0.348 (p = 0.018) and -0.326 (p = 0.146) respectively (i.e., similar differences, but only Course A was statistically significant).

5.5. Net differences by group

The net scores (push-switch) were compared for different factors, and are shown in Table 7. A positive net value means an individual engages in more pushing than switching behaviour, while a negative indicates more switching than pushing.

The first cluster of columns shows mean net scores for females, males, the difference between females and males, and the significance. No statistically significant gender differences were observed in either course, or when considering only Canadian or international students separately, or overall.

The middle cluster of columns shows net scores for Course A and Course B. This reveals a statistically significant increase in net push-switch behaviour from Course A to Course B for both males and females, for both Canadian and international students, and overall.

Finally, considering nationality, an overall statistically significant difference was observed with Canadians having a higher net score than international students. The difference was similar in magnitude in all comparisons, but only statistically significant in Course A (not Course B), and for male students (not for female students).

5.6. Nationality differences by region

Nationality was also considered using the amalgamated United Nations regions listed in Table 3. In this case, some significant differences between regions were noted. Consistent with above, North Americans (98% Canadian and 2% Americans) had the lowest average switching scores, and were among the highest pushing and net (pusher-switcher) scores. Compared to North Americans,
only two of the eight other regions had statistically lower net scores (-1.6, p = 0.005, and -2.0, p = 0.021) accompanied by statistically higher switcher scores (+1.3, p = 0.001, and +1.4, p = 0.049). This suggests that regional differences in pushing and switching behaviour may exist; however, due to the small sample sizes for some regions, it would be inappropriate to conclude anything further at this point. (For this same reason, the names of the two regions described above are withheld.)

6. CONCLUSIONS AND IMPLICATIONS

Overall, some modest effects were noted in this study. A strong difference in behaviours was noted between the second-year Course A (MECH 223) and third-year Course B (MECH 326), with students more likely to engage in pushing and switching behaviours in Course B. The increased difficulty of the content and quizzes in Course B appears to account for some, but not all, of this difference. It is possible that by third year students are more comfortable trying to persuade their peers, even when they are not fully confident in the answer themselves.

Turning to the four research questions in this study:

- **RQ1**: as measured by the pusher behaviours, no significant gender differences were noted in the likelihood that a team would follow an individual in making a poor decision. A statistically significant difference was noted only on the first quiz of a second TBL course, which may suggest if bias is present on a new team, it promptly diminishes.

- **RQ2**: as with gender, no significant differences were noted in the likelihood a team would follow an individual in an incorrect decision based on that individual’s nationality.

- **RQ3**: in determining if an individual’s gender would influence the likelihood that they would allow their team to make an incorrect decision knowing the correct answer themselves, a gender difference was observed. Males were more likely than females to exhibit this switching behaviour. This could be interpreted that the team was less willing to listen to the male students in these cases, or that the male students were less likely to speak up and advocate for what they believed to be the correct answer, or some combination of both.

- **RQ4**: similar to RQ3, a statistically significant difference was observed when considering if an individual’s nationality was related to whether or not a team would make an incorrect decision when the individual knew the correct decision. International students were more likely to exhibit switching than Canadian students.

The differences observed in RQ3 and RQ4 were approximately 0.3-0.4%. This represents the difference in probability of switching behaviour on a given question. With five to six quizzes per course, 15-20 questions per quiz, and 20-30 teams per course, this equates to roughly 5 to 15 students per course experiencing these types of bias events. It is difficult to quantify what the impact of this bias is on an individual, or the impact on the health and cohesiveness of the team. To some individuals and on some teams, these incidents may go unnoticed; in other cases, a single such event could be very damaging. The other important consideration is that many of the decisions in this study (i.e., many of the quiz questions) are straightforward and easy, but in more realistic and practical engineering scenarios, decisions can be more challenging and carry higher stakes. Put another way, although the number of bias incidents observed seems to be small, this should not be interpreted to mean that the impacts from any event are also small.

The pushing and switching behaviours were found to depend strongly on team quiz performance; the higher the team average, the less pushing and switching that occurred. Curiously, the better an individual did on the quizzes, the more they engaged in both pushing and switching behaviours. An increase in an individual’s confidence due to better quiz performance would explain an increase in pushing behaviour, but switching behaviour would be expected to drop, not increase. These findings suggest that it is the team’s confidence in the individual—not the individual’s confidence in themselves—that increases as that person does better on the quizzes. Significant gender differences were not observed here, although a small effect was observed with nationality (international students’ behaviour was more strongly dependent on their own performance). This suggests future work should consider the reasons behind pushing and switching behaviour, whether that includes personality traits, individual confidence, team confidence, team dynamics, and so on.

In terms of practical implications, educators should be aware of the potential for bias in team decision-making. Even though the effects quantified in this study seem to be small, the impacts on individuals and on team cohesiveness could be significant. Making students aware of healthy decision-making behaviours [9], and raising student awareness for the potential of bias [26] appear to be appropriate interventions as part of a course with a large team decision-making element.

References

[1] Galinsky, A. et al., “Maximizing the Gains and Minimizing the Pains of Diversity: A Policy Perspective,” Perspectives on Psychological Science. 10, pp. 742-748, 2015.

[2] Kearney, E., Gebert, D., and Voelpel, S. C., “When and how diversity benefits teams: The importance of team
members' need for cognition," Academy of Management Journal, 52(3), pp. 581–598, 2009.

[3] Stevens, F., Plaut, V., and Sanchez-Burks, J. “Unlocking the Benefits of Diversity: All-Inclusive Multiculturalism and Positive Organizational Change,” The Journal of Applied Behavioral Science. 44, pp. 116-133, 2008

[4] Pramlal, Ashika, The Influence of Team Cohesiveness on Team Effectiveness. U Kwa-Zulu Natal, Masters Dissertation. https://doi.org/10.13140/2.1.2350.0160

[5] Macke, Caroline, James Canfield, Karen Tapp, and Vanessa Hunn. “Outcomes for Black Students in Team-Based Learning Courses.” Journal of Black Studies 50, no. 1, pp. 66-86, January, 2019.

[6] C. Macke, J. Averitt Taylor, J.E. Taylor, K. Tapp and J. Canfield, “Social Work Students’ Perceptions of Team-Based Learning,” Journal of Teaching in Social Work, 35:5, pp. 454-470, 2015.

[7] Stonewall, J., Dorneich, M., Dorius, C., and Rongerude, J. “A Review of Bias in Peer Assessment,” 2018 CoNECD, Crystal City, VA, April 2018.

[8] Ostafichuk, P.M. and Sibley, J., “Self-bias and Gender-Bias in Student Peer Evaluation: An Expanded Study,” Proc. Of CEEA-ACEG 2019, Ottawa, ON, June, 2019.

[9] Preast, Vanessa, “The Development of Team Trust Over Time and its Effect on Performance When Using Michaelsen's Team-Based Learning," Dissertation, 2012. DOI: doio/10.31274/edt-180810-1692.

[10] B.A. Reinig, G. E. Whittenburg and I. Horowitz, “Modelling Performance Improvement and Switching Behaviour in Team Learning,” Accounting Education, 18:4-5, pp. 487-504, 2009.

[11] Sibley, J.E., and Ostafichuk, P. Getting Started with Team-Based Learning. Sterling, Virginia: Stylus, 2014. [ISBN: 9781620361962]

[12] Micheelsen, L.K., Knight, A.B., and Fink, L.D. (Eds.). Team-Based Learning: A transformative use of small groups in college teaching. Sterling, VA: Stylus. 2004. [ISBN: 9781579220860]

[13] Nouri, J., “The flipped classroom: for active, effective and increased learning – especially for low achievers,” Int J Educ Technol High Educ 13, 33, 2016.

[14] Gilley, B.H., and B. Clarkston, “Collaborative testing: Evidence of learning in a controlled in-class study of undergraduate students,” Journal of College Science Teaching 43.3, pp. 83-91, 2014.

[15] Anderson, L.W., Krathwohl, D.R., and Airasian, P.W., “A Taxonomy for Learning, Teaching, and Assessing: A

Revision of Bloom's Taxonomy of Educational Objectives,” New York: Longman, 2000. [ISBN: 080131903X}

[16] Epstein, M.L., Lazarus, A.D., Calvano, T.B., Matthews, K.A., Hendel R.A., Epstein, B.B., and Brosvic, G.M., “Immediate feedback assessment technique promotes learning and corrects inaccurate first responses,” Psychol Rec, 52: pp. 187–201, 2002.

[17] Hodgson, A.J., and P.M. Ostafichuk, “Team-Based Learning in the Design Modules of a New, Integrated, 2nd Year Curriculum at UBC,” in Proc. CDEN 2005, Kananaskis, AB, July, 2005.

[18] Ostafichuk, P.M., E.A. Croft, S.I. Green, G.S. Schajer and S.N. Rogak, “Analysis of Mech 2: An Award-Winning Second Year Mechanical Engineering Curriculum,” in Proc. of EE2008, Loughborough, UK, July 2008.

[19] Feichtner, S., and E. Davis, “Why Some Groups Fail: A Survey of Students’ Experiences with Learning Groups,” The Organizational Behavior Teaching Review, Vol. 9 No. 4, 58-73, (1984).

[20] Oakley, B., Rebecca, B., Felder, R., and Imad, E., “Turning student groups into effective teams,” Journal of Student Centered Learning, 2004.

[21] Brickell, J., Porter, A., Reynolds, M., and R. Cosgrove, “Assigning Students to Groups for Engineering Design Projects: A Comparison of Five Methods,” Journal of Engineering Education, Vol. 83, Issue 3, 259-62, July 1994.

[22] Ostafichuk, P.M., Fengler, M., and Naylor, C., “Measuring the Influence of Team Functioning on Design Project Outcomes,” Proc. CEEA14, Canmore, AB, June 2014.

[23] Ostafichuk, P.M., Fengler, M., and Hodgson, T., The Engineering Design Process: An Introduction for Mechanical Engineers, 3rd ed. Lulu, 2019. [ISBN: 9780992058739]

[24] Budynas, R., Nisbett, J. and Shigley, J., “Shigley's mechanical engineering design,” 10th ed. New York: McGraw-Hill, 2018. [ISBN: 9780073398211]

[25] Tuckman, B.W., and Jensen, M.A.C., “Stages of Small-Group Development Revisited,” Group & Organization Studies, 2(4), pp. 419–427, 1977.

[26] Sukhera, J. and Chris W., “A Framework for Integrating Implicit Bias Recognition into Health Professions Education.” Academic Medicine 93 pp. 35–40, 2018.