UNDERSTANDING THE RELATIONSHIP BETWEEN TEAM DYNAMICS ON PEER EVALUATIONS AND TEAM EFFECTIVENESS

by

Behzad Beigpourian

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

School of Engineering Education

West Lafayette, Indiana

August 2020

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Matthew W. Ohland, Chair

School of Engineering Education

Dr. Allison F. Godwin

School of Engineering Education

Dr. Brent K. Jesiek

School of Engineering Education

School of Electrical and Computer Engineering

Dr. Anne P. Gardner

School of Professional Practice and Leadership, University of Technology Sydney

Approved by:

Dr. Donna M. Riley

To Sahar

My loving wife and my best friend

ACKNOWLEDGMENTS

I would not be able to finish my study without my advisor, committee members, and the CATME team.

There are no words that can express my deepest gratitude to my advisor, Professor Matthew Ohland, for his kindness, caring, guidance, knowledge, and encouragement. He has been more than an advisor and helped me to overcome all challenges throughout my study. He provided me excellent opportunities and always organized my thoughts and research with his thoughtful feedback and guidance. He is an excellent source of wisdom that remarkably scaffolded my skills in doing research.

I would also like to thank my committee members:

Dr. Allison Godwin, who inspired me by her excellent research papers, provided thoughtful feedback and helped me to organize my dissertation. She also was my instructor and taught me how to conduct significant research.

Dr. Anne Gardner, an Australian engineering education leader, gave me insight into my research. I am glad she joined my committee member and assisted me with her feedback and guidance.

Dr. Brent Jesiek assisted me with his feedback and guidance. He is a very knowledgeable and kind person, and I am grateful for having him as my committee member.

I also want to thank my cohort and ENE community that always were kindly present when I needed help.

Finally, I want to thank the CATME team, especially Dr. Daniel Ferguson, who always encouraged me throughout my study and improved my leadership and management skills.

This material is based upon work supported by the National Science Foundation under Grant No. 1431694. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Engineering students are expected to develop professional skills in addition to their technical knowledge as outcomes of accredited engineering programs. Among the most critical professional skills is the ability to work effectively in a team. Working effectively in teams has learning benefits and also provides an environment for developing other professional skills such as communication, leadership skills, and time management. However, students will develop those skills only if their teams function effectively.

This dissertation includes three studies that together inform team formation and management practices to improve team dynamics. The first study investigates mixed-gender team dynamics to determine whether those teams are realizing their potential. The second study explores the relationship of individual psychological safety and students' team member effectiveness and the moderating effects of team-level psychological safety. The third study explores self-rating bias among first-year engineering students and its relationship to student characteristics and dimensions of team-member effectiveness.

Although mixed-gender teams had equal team dynamics with all-male teams, more team facilitation and training are needed to improve the experience of mixed-gender teams. Asian, Black, and Hispanic/Latino students, as well as students with lower GPA, report lower psychological safety, which is associated with lower team-member effectiveness. Team-level psychological safety moderated this effect for Asian and Hispanic/Latino students. Students' effort in teams was associated with lower self-rating bias, likely an indication of greater self-awareness. Together, these studies and their findings contribute to a broader understanding that there are interrelationships among team composition, team dynamics, and team-member effectiveness, and that these relationships differ based on student characteristics such as race/ethnicity, gender, and prior knowledge. This work adds to the body of research demonstrating the importance of teaching students about effective teamwork, conducting regular peer evaluations of team functioning, and interpreting those peer evaluations carefully to avoid perpetuating any biases. This work also demonstrates the usefulness of psychological safety as an important indicator of marginalization.

1. INTRODUCTION

Engineering programs aim to graduate students who not only have developed individual competence in mathematics and design but have also developed collaboration skills in teaming contexts. ABET's accreditation criteria require that engineering programs support the development of these multidisciplinary teamwork skills (ABET, 2019). Teaming skills are essential for engineers because they have to spend much time in collaboration and working with diverse individuals (Passow, 2012; Trevelyan, 2014). Teamwork also improves other professional skills such as communication, time management, and problem-solving. To improve these professional skills, instructors should use purposeful team formation approaches and team facilitation in a way to make sure students have equal opportunities for learning (Abrami et al., 1994; Felder & Brent, 2001; Johnson, Johnson, & Smith, 2007; Oakley, Felder, Brent, & Elhajj, 2004; Shimazoe & Aldrich, 2010; Windschitl, 1999).

1.1 Problem Statement

Students should have effective teaming experiences and use professional working practices for success in developing this important engineering skill. If instructors form teams without any facilitation, students might develop counterproductive work practices. For example, engineering students might exclude from team activities team members who have lower skills and knowledge. Students can also develop and strengthen preferences to work individually instead of collaborating with team members because they think a competent engineer is a person who finishes the task alone (Leonardi, Jackson, & Diwan, 2009). To avoid these counterproductive working practices, it is important for instructors to evaluate students' teaming skills and monitor team dynamics, and students should actively participate in teams.

Teamwork is difficult to evaluate because it is challenging for instructors to observe all activities and interactions in teams, especially in large classrooms, and especially when significant team interaction occurs outside the classroom. Peer evaluation and evaluation of team dynamics are useful ways to find out what is happening in teams. Through peer evaluation, instructors can measure individual team members' effectiveness, and in team dynamics, instructors can measure the teaming process (Borrego, Karlin, McNair, & Beddoes, 2013; Ohland et al., 2012). To have a valid assessment, student involvement in the assessment process requires utilizing a practical and scientific approach and facilitation.

Mixed-gender teams have high potential to perform better than same-gender teams (Lee, Choi, & Kim, 2018), but research indicates that more effort is needed to utilize the broad skillset of mixed-gender teams (Fila & Purzer, 2014; Hirshfield & Fowler, 2020; Laeser, Moskal, Knecht, & Lasich, 2003). Poor team dynamics in mixed-gender teams might hinder teams from benefiting from mixed-gender composition. Conversely, positive measures of team dynamics are an indication that students are generally benefiting from collaboration and learning within the teams (Asio, Cross, & Ekwaro-Osire, 2018; Beigpourian, Ferguson, Berry, Ohland, & Wei, 2019; Borrego et al., 2013). Self-evaluation is a common approach to assist students in learning team skills through feedback and reflection, but several drawbacks can reduce the effectiveness of self-evaluations (Alba-Flores & Rios, 2019; Cinar & Biglin, 2011). For example, some students have biases in their self-evaluation and are not competent in assessing their skills (Ostafichuk, Mattucci, D'Entremont, Shirzad, & Naylor, 2015).

Diversity in engineering classrooms is an important issue because diversity results in broader skillset in engineering and make the engineering classrooms more inclusive (Ohland et al., 2011). So, engineering education researchers are looking for ways to increase this diversity (Pawley, Schimpf, & Nelson, 2016). Teams in engineering classrooms include students who are minoritized—marginalized by underrepresentation and/or social power—on the basis of race/ethnicity, gender, and citizenship status. To leverage diversity in the engineering classroom to make engineering more inclusive, it would be useful to study the experience of minoritized students in teams through self and peer evaluation and team dynamics.

1.2 Existing Research and Gap

Several engineering education researchers have studied gender dynamics in engineering teams, including teaming interaction, peer evaluation, mode of collaboration, communication, team formation, and leadership self-efficacy (Beigpourian & Ohland, 2019; Flynn, Savage, Penti, Brown, & Watke, 1991; Laeser et al., 2003; Rosch & Collier, 2014; Wolfe & Powell, 2009). I discuss these studies in more detail as appropriate through the dissertation. There is a gap in the literature in comparing the dynamics of all-male teams to those of mixed-gender teams and in determining which measures of team dynamics can best identify gender issues in teams.

Although team dynamics have been of interest to some engineering education researchers for many years (Lewis, Aldridge, & Swamidass, 1998; Natishan, Schmidt, & Mead, 2000; Whitman et al., 2005), recently, the subject has attracted more interest. The last decade has seen multiple research teams exploring such diverse team dynamics as trust, interdependence, conflict, cohesion, and satisfaction in engineering teams (Asio et al., 2018; Baughman, Hassall, & Xu, 2019; Beigpourian, Luchini, Ohland, & Ferguson, 2019; Borrego, Foster, & Froyd, 2014; Neumeyer & Mckenna, 2011). Despite increasing interest in team dynamics, there has been little use of psychological safety, a well-known team dynamic that measures the extent to which team members feel safe to take interpersonal risks and speak up in teams (Edmondson, 1999).

Similarly, various researchers in the past decade have studied the advantages and disadvantages of self and peer evaluation. Self and peer evaluation improves critical thinking, team performance, and problem-solving skills, encourages students to communicate more in teams, and provides an opportunity to learn through reflection and feedback (Alba-Flores & Rios, 2019; Cinar & Biglin, 2011; Hansen, 2006; Wang, Hsieh, & Chen, 2016; Willey & Gardner, 2010). However, some studies raised some concerns about fairness and accuracy in self and peer evaluation (Cinar & Biglin, 2011; Falchikov, 2004; Marcangelo, Cartney, & Barnes, 2010). In spite of these concerns, there has been little focused study of self-evaluation bias and its possible causes and effects.

1.3 Studies Overview

This dissertation includes three studies that together inform team formation and management practices to improve team dynamics. These studies add to teaming knowledge in this area and fill existing gaps (mentioned in the previous section) in engineering education research about the first-year engineering students' experience by focusing on peer evaluation and team dynamics. I decided to study first-year engineering students because first-year and capstone are the most common points in the curriculum when team-based learning occurs. First-year students are novices, and they need to practice teaming, and their team experiences lay a foundation for the rest of their education. If they learn the teaming practice and have a positive experience in the first-year, they will have more successful experiences in the later their teaming experiences in senior years or after graduation (Leonardi et al., 2009).

Here, I shortly described the objectives of three studies:

Study 1: This study addresses the gap in mixed-gender team dynamics by investigating whether mixed-gender teams have different team dynamics than all-male teams. Also, I explore how team composition affects the satisfaction of women in teams.

Study 2: This study fills the gap about psychological safety of minoritized students in team contexts. First, I look at differences in students' psychological safety based on race/ethnicity, gender, citizenship status, and grade-point average (GPA). Next, I investigate the relationship between psychological safety and team-member effectiveness controlling for race/ethnicity, gender, citizenship status, and GPA. Lastly, I study the relationship between team average psychological safety and team-member effectiveness.

Study 3: This study examines students' self and peer evaluations to see whether students' gender, race, citizenship status, GPA, and their teammates' perceptions of their contributions to the team's work and interactions with teammates influence their self-rating of their knowledge, skills, and abilities (KSAs).

1.4 Significance

These studies described here are significant for several reasons. First, it appears from the literature that the benefits of mixed-gender teams are not being realized in student learning and, more generally, engineering is not benefiting from the gender diversity that it has. Improving team dynamics in mixed-gender teams would lead to better performance of teams and team members (Lee et al., 2018). This study leads to suggestions to improve the facilitation or training approaches in mixed-gender teams.

To support the generation of diverse ideas in teams, all team members should participate in idea generation. Based on Edmondson's (1999) definition of psychological safety, students will not bring their ideas into teams if they have low psychological safety. Describing group differences and influences on psychological safety identifies areas for improvement and provides tools for measuring the impact of course-level innovations designed to address any systematic biases.

Although industry will always need to invest in on-boarding and on-the-job training more generally, improving teamwork and self-evaluation skills would be a significant contribution that academia could make before students graduate rather than allowing misconceptions to become more strongly rooted (Leonardi et al., 2009). This study helps identify factors leading to bias in self-evaluation and lack of self-awareness, and consequently increasing the students teaming skill knowledge.

Improving team experiences could encourage students, particularly minoritized students, to persist in engineering. Retention among minoritized students is low in some universities, and many engineering students leave engineering programs in the first and second year of study (Ohland et al., 2011). Since early engineering courses commonly include a focus on teamwork, even at institutions that do not have a formal first-year-engineering program (Reid, Reeping, & Spingola, 2018), improving team experiences is a high-impact place to improve retention.

1.5 Common Assumptions

Several assumptions guide my research are explicitly described below to foreground their role in the work presented in this dissertation. These assumptions make the results valid and reliable.

1. Reliable and valid instruments have been used to collect the data.
2. Participants complete surveys independently and without getting any help from other students.
3. The learning environment is equivalent for all students so observed differences are associated with teaming experiences.
4. The amount and quality of previous teaming experiences are distributed randomly among teams.

1.6 Limitations

The data used for this research have some limitations, which I provide below. The same data are used for all three studies in this dissertation. These limitations influence the result interpretation, which has been considered through the discussion. Suggestions for future research have also been proposed that can address some of these limitations.

1. The data is collected from a single university. This university's students might not be representative of students at other universities.
2. This study relies on student perceptions of team member effectiveness and team dynamics rather than expert observation.

3. In the first study, the intersection of race/ethnicity and gender is not considered, so the result most likely represents the experience of White women.
4. Gender is treated as binary in the first study because the number of "Other" responses is prohibitively small.
5. Studies 2 and 3 ignore important variables such as language, personality, and socioeconomic status because data were unavailable.
6. International students are from different countries with different backgrounds and cultures, but they are aggregated in these studies.
7. Instructor and Teaching Assistant effects are not considered.
8. Some students have not answered the survey questions.

1.7 Data collection

Data were collected using the web-based interface of CATME, the Comprehensive Assessment of Team-Member Effectiveness). CATME is an integrated system that collects data team formation, peer evaluation, and team dynamics (Ohland et al., 2012). CATME has been available for use in higher education since 2005, and has been used by over 1.5 million students of over 20,000 instructors across 2400 institutions in 87 countries.

CATME's Team-Maker tool forms teams based on criteria assigned by instructors. When instructors want to form teams, they can ask questions such as race/ethnicity, gender, citizenship status, GPA, schedule, and software skills. After students completed a Team-Maker survey, instructors assign a weight to each question according to the criteria's importance. They also can decide whether they want to group students based on similarities or differences in each criterion (Layton, Loughry, Ohland, & Ricco, 2010). All participants in this research completed a Team-Maker survey and answered questions about gender, race/ethnicity, citizenship status (domestic or international), and GPA. In this study, I use those data as participant demographics.

Once teams are formed, instructors can ask students to evaluate their team members at different stages. CATME's peer evaluation tool includes team-member effectiveness questions and follow-up questions, including team dynamic questions. Team-member effectiveness is measured using five behaviorally anchored rating scales (Ohland et al., 2012): contributing to team's work, interacting with teammates, keeping the team on track, expecting quality, and having relevant knowledge, skills, and abilities. Students decide which category of behaviors is the closest match to the behavior of their team members (including themselves). The CATME peer evaluation instrument is provided in Appendix A. Team process measures can be collected along with the peer evaluation data. More details about participants and data collection processes are available in each study.

1.8 Course and Team Project Context

Participants were enrolled in a required first-year engineering course. The high-level learning objectives of the course are to apply basic programming skills to solve engineering problems, to develop and interpret mathematical models of engineering problems, to function as an effective team member, and act professionally. Team-based activities, including work completed as a team as well as measures of team-member effectiveness comprise approximately 50 percent of the grade in this class.

All students participated in an 8-week project with multiple milestones. Before starting the project, team members engaged in multiple team-based activities, including developing a document describing their team processes (commonly called a team charter or a team contract). Each team member also engaged in pair programming exercises with each of the other members of their team, so each dyad within a team had independent work experience before starting the project as a whole team. In the larger project, students engaged in a modeling activity using MATLAB. The project was sufficiently open-ended that teams were asked to select and justify their solution strategy. As the project progressed, the dataset modeled became increasingly complex, requiring students to improve their algorithm.

Students evaluated their team member effectiveness and their teams' dynamics in four stages over the course of the class. The first self/peer evaluation and interdependence were collected in the fifth week of semester, the second self/peer evaluation and conflict were collected in twelfth week, the third self/peer evaluation and psychological safety in the fourteenth, and the final self/peer evaluation and satisfaction in the last week of semester, when all team assignments had been turned in.

1.9 Interrater Agreement

In Study 1 and Study 2, I measure some team constructs by averaging students' answers in their perception about their teams. Aggregating individual data to the team level assumes that there is sufficient agreement among different raters. There are two compositions models that describe these team-level constructs. In the direct consensus model, individual-level measurements are aggregated to create a group level measurement. For example, the satisfaction measures an individual student's satisfaction. In the referent shift-consensus model, students evaluate a team-level construct, such as conflict, which measure students' perceptions of whether there is conflict in the team. In both models, the aggregation must present the shared perceptions among team members, and there is an agreement about team dynamics (Chan, 1998; Klein, Conn, Smith, & Sorra, 2001; Woehr, Loignon, Schmidt, Loughry, & Ohland, 2015).

The most common index for measuring within-unit agreement is r_{wg} (Bliese, 2000; James, 1982), which is suitable for multiple-item surveys comparing rating variance with the variance expected if there were no agreement. A higher r_{wg} indicates more agreements among units, and researchers rely on cut-off values to interpret the level of agreement. It is generally accepted that an r_{wg} above 0.7 represents an acceptable level of agreement. Researchers have also categorized the strength of agreement: lower than 0.3 as lack of agreement, 0.31-0.50 as weak agreement, 0.51-0.70 as moderate agreement, 0.71-0.90 as strong agreement, and higher than 0.90 as very strong agreement (James, Demaree, & Wolf, 1984; Lance, Butts, & Michels, 2006; LeBreton & Senter, 2008).

For this dissertation, since I am using team-level dynamics, r_{wg} was calculated for all team-level constructs. The r_{wg} values were 0.85 for interdependence, 0.90 for task conflict, 0.95 for relationship conflict, 0.93 for process conflict, 0.88 for psychological safety, and 0.84 for satisfaction. Based on the recommended cut-off values, all within-unit agreements are strong or very strong, so the data can be aggregated at the team level. These teaming constructs are described in Study 1 in detail.

2. STUDY 1

2.1 Abstract

Women are minoritized in U.S engineering programs and most engineering classrooms, and a variety of evidence indicates that they face inequities in team interactions. For example, male team members tend to assign female teammates to non-technical tasks or misinterpret their communication style as a sign of weakness (Beigpourian & Ohland, 2019; Wolfe & Powell, 2009). To reduce these possible biases, some researchers recommend that instructors avoid isolating women in engineering teams (Rosser, 1998). While there are studies of mixed-gender teams in engineering education, they have focused on team functions and team's final report (Laeser et al., 2003), peer evaluation ratings (Kaufman, Felder, & Fuller, 2000), leadership self-efficacy (Rosch & Collier, 2014), and mode of collaboration (Flynn et al., 1991). No comprehensive study of the dynamics of mixed-gender teams could be identified in the context of engineering education. This study investigates multiple measures of team dynamics in mixed-gender teams: task interdependence, conflict, psychological safety, and satisfaction. Further, this research explores the extent to which women's satisfaction improves if their male teammates have similar characteristics (e.g., race/ethnicity and citizenship status) and/or if their male teammates have higher GPAs. The participants of this study were enrolled in a first-year engineering course, who, four times during their teaming process, provided information about various team dynamics. Mann-Whitney U tests were used to explore differences in team dynamics between mixed-gender teams and all-male teams. A multiple regression model was used to predict the women's team satisfaction based on their male team members' characteristics. Mixed-gender teams reported higher levels of task interdependence, but reported similar levels of conflict, psychological safety, and satisfaction. Women tended to be more satisfied when they worked with men who have similar citizenship status (based on class demographics, domestic women are more satisfied if their male teammates are also domestic). The results of this study include recommendations for instructors to improve team formation and facilitation.

2.2 Introduction

Students in effective teams can create a shared identity, increase their positive attitude, develop better social skills, and increase their critical thinking skills. To achieve these benefits, instructors should use intentional approaches for team formation and team facilitation (Abrami et al., 1994; Felder & Brent, 2001; Johnson, Johnson, & Smith, 2007; Oakley, Felder, Brent, & Elhajj, 2004; Shimazoe & Aldrich, 2010; Windschitl, 1999). Additionally, interpersonal relationships among team members can influence team members' experiences and students' learning (Humphrey, Aime, Cushenbery, Hill, & Fairchild, 2017; Joshi, 2014; Krasikova & LeBreton, 2012), and those relationships are influenced by their perceptions. Members of student engineering teams tend to marginalize students perceived to have lower knowledge and skills, excluding them from team activities (Leonardi et al., 2009) and limit their opportunities within the team (Joshi, 2014). This problem is of greater concern if those perceptions are influenced by gender bias rather than the student's real expertise (Meadows & Sekaquaptewa, 2013). Considering the male-dominated setting of engineering classrooms and the importance of interpersonal interactions, it is important to study the tole of gender in team dynamics.

Research suggests that one way to possibly decrease bias against female students is to form teams with equal number of female and male students (Rosser, 1998). It is valuable to study the dynamics of these gender-balanced teams in greater detail, especially in light of a recent qualitative study that found that women are less satisfied in teams with two women than teams with one woman (Hirshfield & Fowler, 2020). This study aims to investigate dynamics in these gender-balanced mixed-gender teams to inform team formation approaches that could improve women's satisfaction in engineering teams.

2.3 Literature Review

The ideal literature would be drawn from engineering education and focus on gender dynamics and their relation to team composition. Since there are few papers that focus on that narrow scope, the literature review is divided into three sections; gender in engineering teams, team formation, and team dynamics.

2.3.1 Gender in engineering teams

Although women have been described as "good team players" (Bannerot, 2003) and tend to respond well to teamwork as a pedagogy, it can also have negative impacts on their experiences (Felder, Felder, Mauney, Hamrin, & Dietz, 1995; Hartman & Hartman, 2006; Hartman & Hartman, 2009; Rosser, 1990; Tonso, 2006). Among the negative impacts women have experienced are being ignored, undervalued, and assigned to non-technical tasks (Wolfe, Powell, Schlisserman, & Kirshon, 2016), having their speech patterns interpreted as a sign of weakness and lack of ability (Wolfe & Powell, 2009), higher conflict due to poor communication (Deep, Othman, & Salleh, 2016), and being assumed to be less interested in leadership (Leaper & Ayres, 2007; Meadows & Sekaquaptewa, 2013) or less fit for it (Rosch, Collier, & Zehr, 2018). Other research provides some possible explanations for these gender differences, such as gendered differences in the definition of leadership (Bailey, Swan, Coso, Creager, & Rowan-Kenyon, 2012) and preferred mode of collaboration (Flynn, Savage, Penti, Brown, & Watke, 1991; Stump, Hilpert, Husman, Chung, & Kim, 2011), and degree of career interest (Berenson, Slaten, Williams, & Ho, 2005).

The more concerning explanation for negative experiences is outright sexism and stereotyping. Some students and faculty believe that female students cannot be proficient in some areas simply because they are female (Mead et al., 1999). Although we might like to think that such ideas are outdated, these stereotypes persist even among some leaders in higher education (Jaschik, 2005). Natishan, Schmidt, and Mead (2000) found that women may need to prove themselves in teams to be seen as equals. Tonso (1996) argued that we should change the culture of engineering education before thinking about including more women because engineering culture increases women's difficulties in teams and classrooms. She suggested that additional studies be conducted on how engineering culture influences women, provide more opportunities for women to express their opinions without fear, and modify engineering practices and policies to expand the definition of engineering and the identity of an engineer. This research seeks to improve our understanding of women's experience in teams.

2.3.2 Team formation

Various team formation strategies have been used in engineering education, ranging from self-selection to instructor-assigned teams, including combinations of two or more approaches (Layton et al., 2010; Parker, Sangelkar, Swenson, & Ford, 2019). Some scholars noted that asking students to form their own teams gives them a feeling that they have ownership of the project, improving their chances for success (Kowalski & Smyser, 2017; Parker et al., 2019), whereas other researchers noted that self-selection can result in teams that lack diversity and teams with social cohesion rather than task cohesion (Bacon, Stewart, & Silver, 1999; Steiner & Kanai, 2016).

Just as there are various approaches to team formation generally, there is also disagreement in practice about how to consider gender when forming teams. Some professors believe that there is no need to consider gender in team formation because gender bias in teamwork is not an important issue, and women would encounter the same problem at the workplace (Beddoes & Panther, 2018). Mead et al. (1999) found that half of faculty believe that gender does not influence team activities. Beddoes and Panther (2018) found that professors rarely consider gender during team formation, but when they do, practice ranges from forming teams to avoid isolating women to forming same-gender teams.

Research does show that the gender composition of teams does affect students' team experiences. Cinar and Bilgin (2009) found that having more women in teams decreased peer evaluation bias within teams, and more students gave a full rating to their teammates. Lloyd and Szymakowski (2017) discovered that teams with a higher percentage of women had more conversations about tasks, more collaboration for solving a problem, and a higher level of verbal discussion in the class. Researchers at the Colorado School of Mines found that teams with equal numbers of men and women had lower performance in all team functions, raising questions about mixed-gender team formation (Laeser et al., 2003).

2.3.3 Team dynamics

Students' personality, previous teaming experience, and students' academic performance influence the team dynamic (Monaghan et al., 2015; Steiner & Kanai, 2016). More collaboration in teams generally results in better team performance (Menekse, Higashi, Schunn, & Baehr, 2017). The importance of team dynamics is well proven in organizational behavior and psychology. Interdependence, conflict, psychological safety, satisfaction, trust, and cohesion are the most commonly used measures of team dynamics in organizational behavior and psychology, and there are benefits in measuring them in the engineering education context (Beigpourian, Luchini, et al., 2019; Borrego et al., 2013). Measures of team dynamics included in this study are described in more detail below. Satisfaction, cohesion, conflict, and psychological safety have been studied in the context of engineering education (Asio et al., 2018; Baughman et al., 2019; Whitman et al., 2005) to study teams in flipped classrooms or virtual teams, or how these team dynamics influence the team innovations, but there is less research on team dynamics focused on the effects of gender composition. It is important to study the relationship of gender and team dynamics given the possibility that women report less satisfaction in mixed-gender teams (Hirshfield & Fowler, 2020) and particularly because team dynamics can provide an indicator of whether the teams are being managed well (Lee et al., 2018), making it possible for mixed-gender teams to achieve the benefit of diverse perspectives (DiStefano & Maznevski, 2000).

2.3.3.1 Team interdependence

Effective teams include members who depend on each other for completing the projects and achieving the team outcomes (Guzzo & Shea, 1992). Distribution of roles, skills, and resources, giving feedback, and having a reward system all influence team interdependence (Wageman, 1995). Team interdependence can be in the form of task interdependence or goal interdependence. Task interdependence is when team members need to use each other's knowledge and expertise to do their job, and goal interdependence is when the team has developed shared goals (Vegt, Emans, & Vliert, 2001).

2.3.3.2 Conflict

Three types of conflict are described in teams: relationship conflict, task conflict, and process conflict. Relationship conflict might happen because of different personalities in teams and create tension and friction in teams (Jehn, 1995). Task conflict starts when team members have different views about what the teamwork content is and when they have diverse ideas about how to accomplish the team's goal (Jehn, 1994). Process conflict occurs when there is disagreement among team members about how to use resources and assign responsibilities (Jehn & Bendersky, 2003). Conflict has a complex effect on team performance; relationship conflict usually affects team performance negatively, but task conflict might benefit a team because it can be a sign of an existing variety of ideas. Process conflict can have adverse results for teams, but can improve team performance if it occurs in the early stages of the team's work and is resolved (Amason, 1996; De

Dreu, 2006; Jehn, Greer, Levine, & Szulanski, 2008; Jehn & Mannix, 2001; Neumeyer & Santos, 2020; Pelled, Eisenhardt, & Xin, 1999).

2.3.3.3 Psychological safety

Psychological safety is about being unconcerned about the negative consequences of decision making or suggesting a new idea (Edmondson, 1999; Kahn, 1990). Improving psychological safety in teams increases team learning and results in better interaction, more knowledge exchange, and more significant contributions (Bienefeld & Grote, 2014; Leroy et al., 2012; Siemsen, Roth, Balasubramanian, & Anand, 2009; Zhang, Fang, Wei, & Chen, 2010). With high psychological safety, students show more commitment, better attitude, and better learning behaviors in teams, as well as less task conflict (Biswas & Bhatnagar, 2013; Kirk-Brown & Van Dijk, 2016; Liu, Hu, Li, Wang, & Lin, 2014; Rathert, Ishqaidef, & May, 2009; Ulusoy et al., 2016).

2.3.3.4 Satisfaction

Satisfaction in teams is an indication that team members' expectations are being met. If students' expectations are met about team members' ability and collaboration in teams, they will have higher satisfaction in teams and tend to want to work with that team again (Peeters, Rutte, Van Tuijl, & Reymen, 2006). Team characteristics, well-defined team objectives, assessment and rewards, effective leadership, and conflict management strategies all influence the team members' satisfaction (Barczak & Wilemon, 2001).

2.4 Theoretical Framework

Social role theory focuses on gender similarities and dissimilarities in social behaviors and the consequences of those similarities and differences. Based on social role theory, women might be generally better in some skills, and men might have higher abilities in other skills (Eagly, 1987). Having both men and women on a team therefore provides a mechanism for well-managed mixed gender teams to out-perform single-gender teams; women might attend more to feelings and interpersonal relationship in the team, while men tend to pay attention more to the problem itself (Apesteguia, Azmat, & Iriberri, 2012; Lee et al., 2018; Myaskovsky, Unikel, & Dew, 2005). This difference can also explain gendered differences in the definition of leadership (Bailey et al., 2012).

Mapping, Bridging, Integrating (MBI) theory posits that well-managed teams must understand and map these differences, bridge and integrate them (DiStefano & Maznevski, 2000).

Together, these two theories predict that mixed-gender teams should benefit from complementary skills, and their measures of team dynamics should be better than those of single-gender teams. Since mixed-gender teams need to be managed well (through mapping, bridging, and integrating gender differences) to achieve this outcome, if these benefits are not observed, the most likely explanation is that the mixed-gender teams are not managed well. Here, instructors have a role to facilitate teams so that students can understand and benefit from these differences. Figure 2.1 depicts the proposed theoretical framework for this study, which posits that mixed-gender teams must manage the gender differences that are present and that, if done well, they have the potential to outperform single-gender teams.

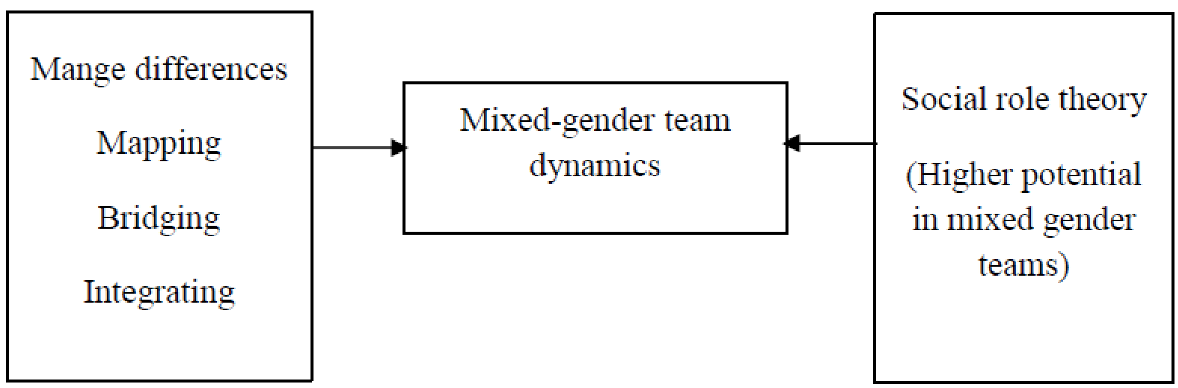

Figure 2.1. A theoretical framework describing dynamics in mixed-gender teams.

2.5 Scope and Research Questions

Reviewing the literature in engineering education revealed a gap in comparing the team dynamics of mixed-gender teams with those of all-male teams. While various studies suggest that female students should not be isolated in team formation, this study aims to describe a broader range of team dynamics and outcomes for mixed-gender teams with an equal number of female and male students. The result of this study can guide team formation strategies and improve instruction about effective teaming in engineering classrooms. The research questions (RQs) for this study are:

RQ1: Does task interdependence differ between mixed-gender teams and all-male teams?

RQ2: Does task conflict differ between mixed-gender teams and all-male teams?

RQ3: Does relationship conflict differ between mixed-gender teams and all-male teams?

RQ4: Does process conflict differ between mixed-gender teams and all-male teams?

RQ5: Does psychological safety differ between mixed-gender teams and all-male teams?

RQ6: Does satisfaction differ between mixed-gender teams and all-male teams?

RQ7: Does the male teammates' GPA and previous programming skills, and the degree of similarity of male race, and citizenship status to those of female students affect female student’ satisfaction in teams?

The level of diversity in teams is difficult to measure because of differences in seen and unseen attributes (e.g., race, gender, academic status, personality, culture, age, social-economic status, location, etc.) can create diversity in teams (Harrison & Klein, 2007). All these factors cannot be accounted for in the analyses, but with a sufficiently large sample, many of them will be randomly distributed, making it feasible to look at differences based on a student being on a mixed-gender or same-gender team. Also, for the last research question, I decided to measure satisfaction instead of performance because women can have high performance in teams, but still not to be satisfied about the teaming process.

2.6 Method

2.6.1 Participants of the study

Participants of this study were from two separate cohorts of a first-year engineering class, including 35 sections, over four semesters (Fall 2017, Spring 2018, Fall 2018, Spring 2019), taught by 14 instructors at a large Midwestern public university. These students completed various measures of team dynamics at four different points in their time together as a team. In total, 3651 students were in the cohorts that participated in this study. For the first to sixth research questions, the study includes only participants in four-member teams with either exclusively male team members (male teams) or exactly two female and two male team members (mixed-gender teams). These are the two most common team compositions in our sample size, accounting for the team experience of 2968 students. For the last research question, only female and male students in four-person mixed-gender teams are included in the study (940 students). After further removing record for students with incomplete data, the demographics of the remaining participants are presented in Table 2.1 for each research question. For the first to sixth research questions, there were 507 all-male teams and 235 mixed-gender teams. The students' self-reported racial and ethnic as citizenship demographics for RQ7 are reported in Table 2.2 and 2.3, respectively. The number of students included in each study varies, because data were collected at multiple time points with varying response rates. The sample size and response rates are high, but could not be assumed to be missing completely at random (MCAR), so I considered imputation, which replaces the missing data with multiple sets of possible values. Among several methods of imputation, multiple imputation was chosen because it has less bias and more statistical power by retaining all data (Jakobsen, Gluud, Wetterslev, & Winkel, 2017; McCleary, 2002).

Table 2.1. Gender reported by study participants

	Female		Male in all-male teams		Male in mixed-gender teams		Missing
Gender	Number	Percent (%)	Number	Percent (%)	Number	Percent (%)	Percent (%)
RQ1	436	14.7	1824	61.5	422	14.2	9.6
RQ2-4	423	14.3	1735	58.5	412	13.9	13.4
RQ5	434	14.6	1820	61.3	415	14	10
RQ6	441	14.8	1862	62.8	427	14.4	8
RQ7	441	46.9	-	-	427	45.4	7.7

Since I am considering race/ethnicity and citizenship status for the last research question, additional demographics are provided for this research question:

Table 2.2. Race/ethnicity reported by RQ7 study participants

	Both Female and male students		Female students	
Race	Number	Percent (%)	Number	Percent (%)
Asian	703	25.8	123	27.9
Black	54	2	9	2
Hispanic	290	10.6	48	10.9
Native	6	0.2	0	0
White	1504	55.1	231	52.4
Other	108	4	16	3.6
Prefer not to answer	65	2.4	14	3.2

Table 2.3. Citizenship reported by RQ7 study participants

	Both Female and male students		Female students	
Race	Number	Percent (%)	Number	Percent (%)
International	513	18.8	86	19.5
Domestic	2217	81.2	355	80.5

2.6.2 Data Sources

The data were collected using CATME, web-based software to form teams and make peer evaluation surveys (Ohland et al., 2012). The CATME system was also used to collect team dynamics measures after the peer evaluation. The data collection procedure was the same in all classes and both semesters; students completed four peer evaluations throughout their teaming process. Students rated their perceptions about team interdependence and conflict in the first and second peer evaluations, respectively. In the third peer evaluation, they reported their level of psychological safety in their teams. In the last peer evaluation, at the end of their team project (or close to their project submission), they rated their team satisfaction.

For the first to sixth research questions, Mann-Whitney U tests for two independent samples were used to compare the sample of all-male teams with the sample of mixed-gender teams, reporting effect sizes for all significant results (Cohen, 2013). Effect size was measured using Cohen's *d*, for which an effect size less than 0.2 is small, between 0.2 and 0.5 is medium, between 0.5 and 0.8 is large, and above 0.8 is very large. The dependent variables for these research questions are task interdependence, task conflict, relationship conflict, process conflict, psychological safety, and satisfaction. These dependent variables were collected using instruments with strong validity evidence and a history of use in engineering education. Cronbach's alpha for these instruments are provided in Table 2.4. Having Cronbach's higher than 0.8 is good, and higher than 0.7 is acceptable (Edmondson, 1999; Jehn & Mannix, 2001; Loughry & Tosi, 2008; Newman, Donohue, & Eva, 2017; Vegt et al., 2001). Vegt et al., (2001) designed the interdependence and satisfaction instruments in the context of teams in school and engineering firms. Jehn and Mannix (2001) developed the conflict instrument in the context of student teams, and psychological safety instrument is highly respected instrument which have been used by several researchers in a variety of disciplines (Edmondson, 1999). These instruments are included in Appendix A. Task interdependence is measured with five items, conflict with nine items (three for each type of conflict), psychological safety with seven items, and satisfaction with three items. CATME uses the original scale of these instruments. Interdependence, conflict, and satisfaction data were collected using a five-point Likert scale (1-5) and psychological safety data collected by a Likert seven-point scale (1-7).

For RQ7, female students' satisfaction ($FemaleS_i$) is the dependent variable. For independent variables, the number of male students in the team whose race matches a female student's is the first independent variable (Race $Similarity_i$), and a similar process was used for citizenship status to find the second independent variable (CS $Similarity_i$). The average GPA and average previous MATLAB skills of each woman's male teammates were the third and fourth independent variables. The model for RQ7 is:

$$FemaleS_i = \beta_0 + \beta_1(\text{Race Similarity}_i) + \beta_2(\text{CS Similarity}_i) + \beta_3(\text{MenGPA}_i) + \beta_4(\text{MenMATLAB}_i) + r_{ij} \quad \text{(Eqn 2.1)}$$

β_0, β_1, β_2, β_3, β_4, are respectively the intercept level of female' satisfaction and the main effects of race similarity, citizenship status similarity, men's average GPA, and men's previous MATLAB skill average. Lastly, r_{ij} is the random error.

Table 2.4 and Table 2.5 summarize the information about variables and their alignment with research questions. Task interdependence, task conflict, relationship conflict, process conflict, psychological safety, and satisfaction variables are collected at the individual level. These variables are collected through multi-item instruments. Each item is measured ordinally, and the ratings from those items are averaged, created a continuous variable for each student. Students provide their individual perspective on a team-level construct. These values are averaged for a team, which is a common and acceptable way among researchers when analyzing several ordinal items measuring a construct (Johnson & Creech, 1983; Norman, 2010; Sullivan & Artino, 2013; Zumbo & Zimmerman, 1993).

Table 2.4. levels of measurement for variables

Variable	Level of Measurement	Number of Items	Cronbach's Alpha
Task Interdependence	Scale average: Continuous Items: Ordinal	5	0.76
Task Conflict	Scale average: Continuous Items: Ordinal	3	0.94
Relationship Conflict	Scale average: Continuous Items: Ordinal	3	0.94
Process Conflict	Scale average: Continuous Items: Ordinal	3	0.93
Psychological Safety	Scale average: Continuous Items: Ordinal	7	0.82
Satisfaction	Scale average: Continuous Items: Ordinal	3	0.84
Females' Satisfaction	Scale average: Continuous Items: Ordinal	3	0.84
Race Similarity	Scale: Ratio		
Citizenship Status Similarity	Scale: Ratio		
Men Average GPA	Scale: Continuous		
Men previous MATLAB skill	Scale average: Continuous Items: Ordinal*		

*The original MATLAB questions was categorical from the lowest skill to the highest skill. In this study, I treated them as ordinal and averaged them in team to make them continuous.

Table 2.5. Alignment of research questions, dependent variable, independent variables, and data analysis

Research Question	**Dependent variable**	**Independent variables**	**Data Analyses**
RQ1: How does task interdependence differ between mixed-gender teams and all-male teams?	Task interdependence measured individually (Vegt et al., 2001) and averaged for the team.	• Mixed Gender Teams • All Male Teams	Mann-Whitney U test
RQ2: How does task conflict differ between mixed-gender teams and all-male teams?	Team task conflict measured individually (Jehn & Mannix, 2001) and averaged for the team.	• Mixed Gender Teams • All Male Teams	Mann-Whitney U test
RQ3: How does relationship conflict differ between mixed-gender teams and all-male teams?	Team relationship conflict measured individually (Jehn & Mannix, 2001) and averaged for the team.	• Mixed Gender Teams • All Male Teams	Mann-Whitney U test
RQ4: How does process conflict differ between mixed-gender teams and all-male teams?	Team process conflict measured individually (Jehn & Mannix, 2001) and averaged for the team.	• Mixed Gender Teams • All Male Teams	Mann-Whitney U test
RQ5: How does psychological safety differ between mixed-gender teams and all-male teams?	Team psychological safety measured individually (Edmondson, 1999) and averaged for the team.	• Mixed Gender Teams • All Male Teams	Mann-Whitney U test
RQ6: How does satisfaction differ between mixed-gender teams and all-male teams?	Team satisfaction measured individually (Vegt et al., 2001) and averaged for the team.	• Mixed Gender Teams • All Male Teams	Mann-Whitney U test
RQ7: How are the GPAs of male teammates and female student' similarity to their male teammates' race and citizenship status related to female student' satisfaction in their teams?	Team satisfaction measured individually (Vegt et al., 2001) female students	• Race similarity • Citizenship Status Similarity • Men's Average GPA in the Team • Men's MATLAB skill	Multiple Linear Regression

2.7 Results

The data were checked for normality, using the Shapiro–Wilk test, and no dependent variable was normally distributed. All assumptions for Mann-Whitney U tests were met, so they were used for RQ1-6 to identify significant differences between all-male teams and mixed-gender teams in term of task interdependence (RQ1), task conflict (RQ2), relationship conflict (RQ3), process conflict (RQ4), psychological safety (RQ5), and satisfaction (RQ6). Since this constituted six analyses on highly related samples, the Bonferroni correction was used to address the issue of multiple comparisons.

Mann-Whitney U tests indicated that task interdependence was greater for mixed-gender teams (Mdn = 3.28) than all-male teams (Mdn = 3.2), U= 49647.5, $p < 0.0183$, with a small-to-medium effect size ($d = 0.27$). Task conflict was the same for mixed-gender teams (Mdn = 1.17) and all-male teams (Mdn = 1.17), U = 56488, $p = 0.48$. Relationship conflict was similar for mixed-gender teams (Mdn = 1.75) and all-male teams (Mdn = 1.75), U = 54871.5, $p = 0.08$. There was no significant difference in process conflict for mixed-gender teams (Mdn = 1.33) and all-male teams (Mdn = 1.33), U = 55694.5, $p = 0.32$. Psychological safety between mixed-gender teams (Mdn = 6) and all-male teams (Mdn = 6.1) was not significant, U = 56149.5, $p = 0.21$. Finally, satisfaction was equal for mixed-gender teams (Mdn = 4.33) and all-male teams (Mdn = 4.33), U= 55657.5, $p = 0.15$.

To predict female students' satisfaction in mixed-gender teams (RQ7), multiple linear regression was used. The data for this research question met the assumptions of multicollinearity, multivariate normality, and homoscedasticity, and the dependent and independent variables were continuous. The data were winsorized to omit extreme outliers. Female students' citizenship status similarity with their men team members had significant effect on female students' satisfaction in mixed-gender teams. However, men's average GPA, men's MATLAB skills, and female students' racial/ethnic similarity with their male team members did not show any significant result.

However, R square for this multiple regression analysis was very low, suggesting that there are other variables influencing the women's satisfaction in teams. Table 2.6 summarizes the result for this analysis.

Table 2.6. Predicting female students' satisfaction based on their male team members' characteristics

Variable	*B*	*SE B*
Intercept	3.29**	0.35
Men's average GPA	0.19	0.10
Race similarity	0.03	0.06
Citizenship status similarity	0.18**	0.07
Men's MATLAB skills	-0.01	0.05
R^2	0.02	
F for change in R^2	3.36	

Note: B is the unstandardized beta, and SE B is the standard error for the unstandardized beta.
*$p < .05$, **$p < .005$

2.8 Discussion

This study aimed to compare the dynamics of mixed-gender teams with all male teams. Based on proposed theoretical framework, mixed-gender teams have more potential for having better team dynamics than all-male teams. Based on the results, in the first stage of teaming, task interdependence was significantly higher for mixed-gender teams than all-male teams. It shows that in the initial stages of teaming, members of mixed-gender teams rely more on each other's unique expertise, which has been shown to result in better team dynamics (Lee et al., 2018). Nevertheless, the subsequent stages of teaming showed no differences between team dynamics of mixed-gender teams and all-male teams in terms of task conflict, relationship conflict, process conflict, psychological safety, or satisfaction. For conflict, it is still a somewhat positive result because the theoretical framework would predict that mixed-gender teams would have higher conflict due to their higher level of diversity. While no negative effects for mixed-gender teams relative to psychological safety and satisfaction were observed, mixed-gender teams did not demonstrate the improved team dynamics that some have observed. These results suggest that there is more that instructors can do so facilitate mixed-gender teams to reach their potential (Fila & Purzer, 2014; Laeser et al., 2003).

The last research question aimed to identify approaches in team formation related to the satisfaction of female students. The R-squared for this model was very low, so I discuss the results here with caution. The only notable result is that female students were more satisfied if their male teammates had the same citizenship status. Since few teams had more than one international student, this result is best interpreted as domestic female students having a preference for working with domestic male students over international male students. It may be that adding cultural (and possibly language) differences complicates the gender differences being managed, resulting in a less satisfying team experience. Racial/ethnic similarity, men's average GPA in teams, and men's MATLAB skills did not have any significant effect on female students' satisfaction, suggesting that there are other unmeasured factors influencing female students' satisfaction.

2.9 Conclusion

Engineering education scholars have conducted a variety of studies about gender in undergraduate engineering teams. However, there was a gap about mixed-gender team dynamics. Based on proposed theoretical framework in this study, mixed-gender teams have higher potential than same-gender teams to have better team dynamics, but for utilizing this potential, their team members should be able to understand the differences, communicate and integrate these differences. This study found that mixed gender teams had better team dynamics in the beginning of teaming, but their subsequent team dynamics were similar to those of all-male teams. Using Social Role Theory and MBI theory, it appears that improvements in facilitating mixed-gender teams are needed to realize the expected improvements in team dynamics. This might take the form of a priori instruction or mentoring throughout ongoing team interaction. Instructors can provide additional training for students, teach them the language of team dynamics, and encourage them to capitalize on the diversity in their teams.

Managing gender differences in mixed-gender teams appears to be complicated by the introduction of cultural (and possibly language) differences. This challenge could be addressed through encouraging students to communicate across both culture and gender differences to find a way to integrate these differences. Students might be more willing to understand these differences if they realize that this diversity is beneficial. As a result, it will be important to address cultural issues as well, either through instruction or team facilitation.

Notably, there are factors affecting female students' satisfaction that were not included in this study that could be addressed by additional studies. The results of this study are addressed further in the final chapter in light of the additional findings from Study 2 and Study 3.

3. STUDY 2

3.1 Abstract

Higher psychological safety (i.e., common feeling among team members to take interpersonal risks without any fear of negative consequences) increases learning in teams (Edmondson, 1999; Newman et al., 2017), yet there is no comprehensive study in engineering education to see how individual and team-average psychological safety are related to students' team experience, and particularly how that relationship differs based on gender, race/ethnicity, citizenship status, previous MATLAB skills, attended semester, and students' GPA. This study addresses these questions using multi-linear regression and multi-level modeling. Individual psychological safety was significantly related to racial identifications of Asian, Black, and Hispanic/Latino, being an international student, attended semester, and GPA. Team member effectiveness was significantly related to GPA, previous MATLAB skills, attended semester, a racial identification of Black, and team-average psychological safety. The results of this study help engineering instructors understand what is happening in teams and add to our knowledge about the team experiences of minoritized students.

3.2 Introduction

Engineering students should acquire multidisciplinary teaming skills before graduation (ABET, 2019), so it is essential to include teaming experiences in engineering curricula and courses. These teaming opportunities should start in the first-year because students can learn teaming skills much faster and easier in the first year and because teaching teamwork early might prevent the enculturation of poor practices (Leonardi et al., 2009). To foster effective teaming practices in engineering classrooms, it is important to understand team members' and teams' experiences. Based on research described in the Introduction, psychological safety increases team learning significantly. Psychological safety gives team members more confidence to take risks in teams, express their ideas, and participate actively. This study explores the extent to which first-year engineering students' characteristics or background might be related to their psychological safety, and whether feeling more psychologically safe and working in teams with high average psychological safety are related with students' team-member effectiveness.

3.3 Literature Review

Among teamwork studies, there are different team input-process-output effectiveness models (Hackman, 1987; McGrath, 1964; Takai & Esterman, 2019; Tannenbaum, Beard, & Salas, 1992; Tucker et al., 2014). Each part (i.e., input, process, or outcome) and their relationship includes different factors influencing team effectiveness. Different ways to measure the teamwork effectiveness have been investigated by several engineering education researchers, such as peer evaluation or team process and outcome measures. Peer evaluation can be used for both formative and summative assessments (Ohland et al., 2012), and can provide feedback to improve communication, leadership, teamwork skills, academic results, and learning (Alba-Flores & Rios, 2019; Wang et al., 2016; Willey & Gardner, 2010).

Team dynamics are strongly related to team effectiveness (Takai & Esterman, 2019; Tucker et al., 2014). Quality of collaboration, personality, experience, and academic performance are all related to team performance (Menekse et al., 2017; Monaghan et al., 2015; Steiner & Kanai, 2016). Trust, interdependence, and conflict are team process assessments that are widely used in studies of organizational behavior and industrial and organizational psychology, and are useful in engineering education as well (Borrego et al., 2013). Students perceive this importance as well, believing it would be beneficial to be educated about team processes (Natishan et al., 2000). Nevertheless, students and faculty have not focused as much on team dynamics (Lewis et al., 1998), and there is need for further study of team processes in engineering education.

Whitman et al. (2005) found that students in virtual teams have less satisfaction and perceived peer performance compared to students who had face-to-face communication in their teams. Asio, Cross, and Ekwaro-Osire (2018) found that cohesion, conflict, and psychological safety were related to perceptions of team innovation. Baughman, Hassall, and Xu (2019) reported that in flipped classrooms, students had improved satisfaction and perceived peer performance after teaming compared to non-flipped classroom teams, but did not notice any difference in terms of conflict. Neumeyer and Santos (2020) reported that conflict decreases the perception of team performances.

Psychological safety is recognized as a significant indicator of a team’s success (Duhigg, 2016), but there is a lack of comprehensive research in this area in engineering education. Psychological safety is defined as feeling safe to take interpersonal risk, express opinions, and act without being worried about the possible negative consequences (Edmondson, 1999; Kahn, 1990).

Psychological safety can increase the learning behaviors in teams, meaning that team members do more activities to improve their performance in teams, such as asking team members for help, seeking feedback, and sharing information with teammates (Edmondson, 2004b). Google identified psychological safety as the most critical factor for increasing team performance (Bergmann & Schaeppi, 2016). Although psychological safety is related to trust, they are conceptually different. Trust measures the extent to team members trust each other individually, while psychological safety goes further than trust and it is a belief about the whole team. There are numerous studies in several disciplines that investigated psychological safety quantitatively and qualitatively at the individual, the team level, and the organization level (Newman et al., 2017).

Higher psychological safety in teams is associated with more interaction and better knowledge sharing (Leroy et al., 2012; Siemsen et al., 2009; Zhang et al., 2010), helps team members to speak up and provide candid feedback (Bienefeld & Grote, 2014; Liang, Farh, & Farh, 2012; Tynan, 2005), increases team member's commitment and attitudes (Biswas & Bhatnagar, 2013; Kirk-Brown & Van Dijk, 2016; Rathert et al., 2009; Ulusoy et al., 2016), decreases relationship and task conflict (Wilkens & London, 2006), and mitigates problems typical of virtual teams (Gibson & Gibbs, 2006). High psychological safety in teams mediates the relationship between team learning orientation (common perception among team members to develop skills and competence) and team learning when the team's open-mindedness (the extent to which a team is curious about new ideas) is low (Harvey, Johnson, Roloff, & Edmondson, 2019). Similarly, psychological safety mediates the relationship between high performance and creativity implementation (Agarwal & Farndale, 2017).

Individuals who work in a psychologically safe environment are more willing to report their errors (Derickson, Fishman, Osatuke, Teclaw, & Ramsel, 2015; Leroy et al., 2012), and this benefits the organizations because when errors are reported, organizations become aware of the errors, the consequence of errors, the circumstances that caused the errors (Reason, 2000; Tax & Brown, 1998; Zhao & Olivera, 2006). There are many studies in engineering education about the importance of learning in teams (Adams, 2001; Grulke, Beert, & Lane, 2001; Hirsch & Mckenna, 2008; Natishan et al., 2000; Nepal, 2016), and engineering students can benefit similarly by learning from their errors in their teams.

The way teams are managed affects psychological safety at the individual and team level. Leadership behaviors such as providing openness in teams, building trust in teams, and having behavioral integrity (consistency of leader's words and actions) has been found to increase psychological safety in teams and team members (Detert & Burris, 2007; Madjar & Ortiz-Walters, 2009; Palanski & Vogelgesang, 2011).

Psychological safety might be particularly useful indicator of the experience of minoritized students. Engineering teams includes students who are minoritized in terms of gender, race/ethnicity, citizenship status, and culture, and student should learn to work effectively in those teams to respect differences without marginalizing minoritized students (Felder, Felder, & Dietz, 2002; Godwin et al., 2017; Natishan et al., 2000; Steiner & Kanai, 2016). Psychological safety encourages members to learn from mistakes and failures (Carmeli & Gittell, 2009; Edmondson, 2004a), is related to team performance (Schaubroeck, Lam, & Peng, 2011), and can mediate the relationship between diversity climate and individual performance, a relationship that may be more robust for minoritized students (Singh, Winkel, & Selvarajan, 2013). Singh and Winkel (2012) found that in organizations where the perceptions of psychological safety are high, minoritized employees exhibit more helping behaviors, meaning that they will help their coworkers more when their coworkers have difficulty in their assignments or falling behind their tasks. Psychological safety has also been found to help employees meet the expectations of their job, and this effect was more substantial for racially minoritized employees (Singh et al., 2013). No research could be identified that explored the psychological safety of minoritized students in engineering teams, so there is a gap in engineering education literature in measuring and studying psychological safety, particularly for studying the experiences of minoritized students. This study aims to fill that gap by describing the extent to which team members' effectiveness is related to psychological safety.

3.4 Theoretical Framework

Team members who experience high psychological safety reflect on the team's work, speak up about concerns, offer ideas without any fear of negative consequences, and are comfortable giving and receive feedback. In contrast, if they feel low psychological safety, they are less likely to share their opinions with others, prefer to be silent, do not give teammates feedback, and contribute less to the team generally (Edmondson, 1999; Roussin, Larraz, Jamieson, & Maestre, 2018). In the following sections, I explain some research and theories to combine them and develop a theoretical framework for guiding my research questions.

3.4.1 Conservation of resources theory (COR)

The Conservation of Resources (COR) theory depicts how depletion or investment of resources in a team or organization can influence the individuals' psychological safety and subsequently decrease or increase team performance (Newman et al., 2017). Based on the COR theory, individuals look for resources like social supports, autonomy, and rewards. Individuals who are deprived of a particular kind of resource, will compensate for their resource loss by seeking to gain other resources. One way individuals can acquire or lose resources is through their relationships in teams (Hobfoll, 1989). Newman et al. (2017) adapted the COR theory for studying psychological safety, arguing that individuals with access to the resources that lead to higher psychological safety (e.g., being able to voice ideas and share knowledge, receiving helpful feedback from teammates, feeling safe to provide feedback to others) have a more positive work outcome. On the contrary, depriving individuals of these resources results in a negative work outcome and lower team member effectiveness.

3.4.2 Diversity-based barriers and psychological safety

Although diversity can be beneficial for teams, it can also result in barriers to effective collaboration (Lau & Murnighan, 2005). Edmondson and Roloff (2009) proposed a conceptual framework to illustrate the relationship between diversity-based barriers and psychological safety. Three types of diversity have been described: separation diversity, variety diversity, and disparity diversity (Harrison & Klein, 2007). When a difference in an attribute in teams is related to the horizontal continuum, such as time zone or location, the diversity is of the separation type. When the difference in teams is categorical (such as race/ethnicity and gender), the diversity is characterized as the variety type. Differences in life experience belong to this category. Disparity diversity is related to a vertical continuum difference, and creates the most challenging barriers for collaboration. Differences in expertise can be considered as a disparity where opportunities to gain these experiences are systemically unequal. Team members with high levels of relevant expertise may want to lead their team, while individuals with lower levels of relevant expertise might feel lower psychological safety and be marginalized (Leonardi et al., 2009). Team members with lower social standing can have a hard time speaking up and sharing their ideas. To reduce this barrier to collaboration, nurturing psychological safety is important. Lack of psychological safety in disparity-diverse teams can lead to lower team learning. While demographic diversity is of the variety type based on social constructs, if different levels of social status are associated with particular demographic groups, disparity diversity results, which might result in more barriers to collaboration and lower psychological safety.

Based on this theoretical framework, individuals with minoritized demographic attributes might have an unequal chance of gaining psychological safety in teams if race/ethnicity, gender, and citizenship status are ascribed a lower social status. In engineering teams, it is likely that minoritized students feel lower psychological safety. Based on this theoretical framework, however, even if minoritized students lose individual psychological safety as a resource, higher psychological safety at the team level team might help compensate for the lack of individual psychological safety, resulting in increased team-member effectiveness.

3.5 Scope and Research Questions

Engineering schools seek ways to improve students' professional skills, in addition to their technical skills (ABET, 2019). Working in team is an efficient and effective way to develop both professional and technical skills, but low psychological safety can prevent realizing those benefits. Psychological safety and its relationship to team-member effectiveness has not been studied sufficiently in the context of engineering education, particularly to explore group differences in psychological safety. This study aims to answer the following research questions:

RQ1: To what extent engineering students' demographics (gender, race, GPA, citizenship status, attended semester) related to their perceptions about psychological safety in their teams?

RQ2: To what extent perceived team-member effectiveness related to engineering students' demographics (gender, race, GPA, citizenship status, attended semester) and their individual psychological safety?

RQ3: To what extent perceived team-member effectiveness related to team-average psychological safety?

This study is significant because it focuses on conditions that affect learning during the team interactions of engineering students. The theoretical framework of this study posits that psychological safety influences team member effectiveness, so if minoritized students are deprived of individual and team psychological safety, it will affect their opportunity to learn in teams.

Confirming the proposed theoretical framework is a warning for engineering education community to pay attention to students' psychological safety to make engineering programs more inclusive.

3.5.1 Proposed theoretical framework

Integrating the previous sections, I depict a theoretical framework in Figure 3.1 linked to the research questions posed in this study. In this theoretical framework, demographic diversity influences individual psychological safety, which is investigated with the first research question. Both individual characteristics and individual psychological safety affect team member effectiveness, a relationship explored in the second research question. Finally, team psychological safety can act as a positive resource and improve the students team member effectiveness, the focus of the third research question.

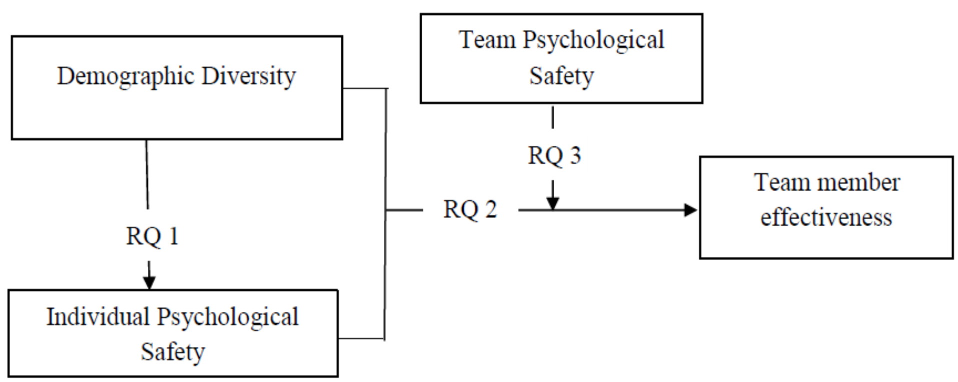

Figure 3.1. Modeled relationship of individual psychological safety, team psychological safety, and team-member effectiveness.

3.6 Method

3.6.1 Participants of the study

The dataset used in this study comprises four cohorts of students enrolled in a first-year engineering during the Fall 2017, Spring 2018, Fall 2018, and Spring 2019 semesters. Nine students retook the class during that time, so there is little overlap among the cohorts. Each class offering followed the same process in assigning four peer evaluations and follow-up team dynamic questions. In all classes, participants completed the psychological safety instrument in the third peer evaluation, which occurred approximately three weeks before the completion of the project and course.

There were 3651 potential participants in this study, but only 3271 students answered all questions related to demographic information, psychological safety, and team member effectiveness. The data violated the assumptions of missing completely at random (MCAR), so multiple imputation was used to substitute for the missing data. The frequency of gender, race/ethnicity, citizenship status, and semester for the remaining participants are summarized in Table 3.1, Table 3.2, Table 3.3, and Table 3.4, respectively.

Table 3.1. Demographics of study participants based on gender

Gender	Number	Percent (%)
Female	758	23.2
Male	2492	76.2
Other	21	0.6

Table 3.2. Demographics of study participants based on race

Race/Ethnicity	Number	Percent (%)
Asian	784	24.0
Black	71	2.2
Prefer not to answer	82	2.5
Hispanic	320	9.8
Native	7	0.2
Other	114	3.5
White	1893	57.9

Table 3.3. Demographics of study participants based on citizenship status

Citizenship Status	Number	Percent (%)
International	2700	82.5
Domestic	571	17.5

Table 3.4. Demographics of study participants based on semester frequency

Attended Semester	Number	Percent (%)
Fall 17	187	5.7
Fall 18	153	4.7
Spring 18	1426	43.6
Spring 19	1505	46.0

3.6.2 Data Sources

Data for this study were collected using CATME, a web-based tool for team formation and peer evaluation (Ohland et al., 2012). Psychological safety is the dependent variable in the first research question (individual's psychological safety), and an independent variable in the second (individual's psychological safety) and third research question (team average psychological safety) and was collected using Edmondson's (1999) instrument. Students responded to seven statements using a seven-point Likert scale, and then the average of these statements (after reversing the score for negatively worded statements) was used as the psychological safety for an individual student, ranging from 1 to 7, with 7 being the highest. This instrument was used because it has strong validity evidence and has been used to measure psychological safety by a large number of scholars in different disciplines (Newman et al., 2017). In the second research question, the dependent variable is a student's perceived team-member effectiveness, which was collected using CATME's five peer evaluation dimensions in "*Contributing to the team's work*", "*Interacting with Teammates*", "*Keeping the team on track*", "*Expecting quality*", and "*Having relevant Knowledges, Skills, and Abilities*". I excluded the self-ratings and averaged these dimensions to use as a student's team-member effectiveness as perceived by their teammates. Self-ratings are excluded to focus on the perceived team member effectiveness by others and due to the possible self-rating inflation to reduce the bias in the results. These dimensions were measured using a behaviorally anchored rating scale (BARS). Following prior research, the CATME peer ratings were converted to a quantitative measure from 1 (the lowest team-member effectiveness) to 5 (the highest team-member effectiveness). Both the psychological safety measure and the BARS scale are available in Appendix A. Table 3.5 describes the levels of measurement for variables, and Table 3.6 presents alignment among our research questions, the measures for answering each question, variables, and how the data have been analyzed.

Table 3.5. levels of measurement for variables

Variable	Level of Measurement
Individual psychological safety	Scale average: Continuous Individual items: Ordinal
Gender	Scale: Nominal
Race	Scale: Nominal
Citizenship status	Scale: Dichotomous
GPA	Scale: Continuous
Previous MATLAB skill	Scale: Nominal
Attended semester	Scale: Nominal
Team psychological safety	Scale average: Continuous Individual items: Ordinal

MATLAB data collected in four categorical variables ranging from the lowest skill to the highest previous skills; no previous experience in MATLAB (None), basic calculation skills (Basic Calc), advanced calculation skills (Adv Calc), and coding experience (Coding).

Table 3.6. Alignment of research questions, dependent variable, independent variables, and data analysis

Research Question	Dependent variable	Independent variables	Data Analyses
RQ 1: How are engineering students' demographics (gender, race, GPA, citizenship status, attended semester) related to their perceptions about psychological safety in their teams?	Students' perceptions about the team's psychological safety (Edmondson, 1999)	• Gender • Race/ethnicity • Citizenship status • GPA • Attended semester • Previous MATLAB skill	Multiple linear regression model
RQ 2: How is perceived team-member effectiveness related to engineering students' demographics (gender, race, GPA, citizenship status, attended semester) and their individual psychological safety?	Peer evaluation ratings of team-member effectiveness from CATME, excluding self-ratings (Ohland et al., 2012)	• Gender • Race/ethnicity • Citizenship status • GPA • Attended semester • Individual psychological safety • Previous MATLAB skill	Hierarchical multilevel modeling – level 1 (Individual level)
RQ 3: How is perceived team-member effectiveness related to team-average psychological safety?	Peer evaluation ratings of team-member effectiveness from CATME, excluding self-ratings (Ohland et al., 2012)	• Team psychological safety (Average psychological safety of all team members)	Hierarchical multilevel modeling – level 2 (Team level)

3.6.3 Analytic Strategy

For the first research question, to predict a student's psychological safety based on a student's characteristics (Race, gender, citizenship status, GPA, attended semester, and previous MATLAB skills), I estimated a regression model using SPSS:

$$StudPS_i = \beta_0 + \beta_1(Gender_i) + \beta_2(Race_i) + \beta_3(CS_i) + \beta_4(GPA_i) + \beta_5(semester_i) + \beta_6(MATLAB_i) + r_{ij} \quad \text{(Eqn 3.1)}$$

In this model, $StudPS_i$ is a student's psychological safety, which is predicted by the sum of the intercept (β_0), the main effect of gender (β_1), race/ethnicity (β_2), CS which is citizenship status (β_3), GPA (β_4), attended semester (β_5), and previous MATLAB skills (β_6). r_{ij} accounts for random error.

For the second and third research questions, I used multilevel modeling with two levels to predict the effect of a student's characteristics and their team's average psychological safety (TeamPS) on a student's perceived team-member effectiveness (StudTE). I considered a hierarchical two-level model of change using the Mixed Procedure (PROC Mixed) in SAS. However, in level 2, I excluded the variables that were non-significant in the first model. This first level (individual level) model is specified as:

$$StudTE_{ij} = \beta_{0j} + \beta_{1j}(Gender_j) + \beta_{2j}(Race_j) + \beta_{3j}(CS_j) + \beta_{4j}(GPA_j - MeanGPA) + \beta_{5j}(semester_j) + \beta_{6j}(StudPS_j - MeanStudPS) + \beta_{7j}(MATLAB_j) + r_{ij} \quad \text{(Eqn 3.2)}$$

Where $StudTE_{ij}$, the perceived team-member effectiveness of the i^{th} student in the j^{th} team, is modeled as a function of β_{0j}, β_{1j}, β_{2j}, β_{3j}, β_{4j}, β_{5j}, β_{6j}, and β_{7j}, team- level coefficients indicating the mean of student's "perceived team-member effectiveness" score of team j, the effect of gender on the student's perceived team-member effectiveness in team j, the effect of race on the student's perceived team-member effectiveness in team j, the effect of citizenship status on the student's perceived team-member effectiveness in team j, the effect of a student's GPA relative to average GPA in team j, the effect of attended semester on the student's perceived team-member effectiveness in team j, the effect of a student's psychological safety relative to average of students psychological safety, and the effect of previous MATLAB skills on the student's perceived team-member effectiveness in team j, respectively, and r_{ij} is random error. Team- level coefficients were in turn modeled as Eqn 3.3 to 3.10:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(TeamPS_j - MeanTeamPS) + \mu_{0j} \quad \text{(Eqn 3.3)}$$

$$\beta_{1j} = \gamma_{10} + \mu_{1j} \quad \text{(Eqn 3.4)}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}(TeamPS_j - MeanTeamPS) + \mu_{2j} \quad \text{(Eqn 3.5)}$$

$$\beta_{3j} = \gamma_{30} + \gamma_{31} (TeamPS_j - MeanTeamPS) + \mu_{3j} \quad \text{(Eqn 3.6)}$$

$$\beta_{4j} = \gamma_{40} + \gamma_{41} (TeamPS_j - MeanTeamPS) + \mu_{4j} \quad \text{(Eqn 3.7)}$$

$$\beta_{5j} = \gamma_{50} + \gamma_{51} (TeamPS_j - MeanTeamPS) + \mu_{5j} \quad \text{(Eqn 3.8)}$$

$$\beta_{6j} = \gamma_{60} + \mu_{6j} \quad \text{(Eqn 3.9)}$$

$$\beta_{7j} = \gamma_{70} + \mu_{7j} \quad \text{(Eqn 3.10)}$$

where γ_{00}, γ_{10}, γ_{20}, γ_{30}, γ_{40}, γ_{50}, γ_{60}, γ_{70}, and γ_{01} are the sample average level of student's perceived team-member effectiveness, the main effect of gender on student's perceived team-member effectiveness, the main effect of race on student's perceived team-member effectiveness, the main effect of citizenship status on student's perceived team-member effectiveness, the main effect of student's GPA relative to average GPA on the student's perceived team-member effectiveness, the main effect of attended semester on student's perceived team-member effectiveness, the main effect of student's psychological safety relative to average of students psychological safety on the student's perceived team-member effectiveness, the main effect of average psychological safety in the student's team relative to the average team psychological safety in the entire sample, and the main effect of previous MATLAB skills on student's perceived team-member effectiveness, respectively. The values γ_{21}, γ_{31}, γ_{41}, and γ_{51} are the interactions of average psychological safety in the student's team relative to the average psychological safety among all teams with race, citizenship status, students' GPA relative to average GPA in their team, and attended semester, respectively. The values μ_{0j}, μ_{1j}, μ_{2j}, μ_{3j}, μ_{4j}, μ_{5j}, μ_{6j} and μ_{7j} are team-level errors in the estimated intercept and slope for gender, race, citizenship status, GPA, attended semester, individual psychological safety, and previous MATLAB skills, respectively.

3.7 Results

Data from first-year engineering students was modeled using multiple regression to characterize individual psychological safety (RQ1), so the assumptions for multiple regression were checked. The dependent variable is measured on a continuous scale, and there are several independent variables that are either continuous or categorical. The data were winsorized to remove extreme outliers, and the conditions of multivariate normality, multicollinearity, and homoscedasticity were all met, by checking Shapiro-Wilk tests, variance inflation factors, and residuals, respectively. R square in this model was low, suggesting that still there are many other variables predicting the individual psychological safety. Yet, the purpose was not to define a model to predict psychological safety, but rather to estimate the extent to which students' demographics influence their individual psychological safety. Higher student GPA was associated with a significantly higher psychological safety. Black, Asian, and Hispanic/Latino students had significantly lower psychological safety. Gender did not have any relationship to psychological safety. Being an international student had the most significant negative relationship to student psychological safety. Students enrolled Spring 2019 had the highest psychological safety. Students' previous experience in MATLAB had no relationship to psychological safety. Table 3.7 shows the result for this multiple regression.

Table 3.7. Predicting psychological safety of students based on their demographic's information

Variable	*B*	*SE B*	
Intercept	5.933**	0.085	
GPA	0.083**	0.025	
Female	-0.030	0.027	
Other Gender	-0.102	0.187	
Asian	-0.166**	0.032	
Black	-0.241**	0.076	
Hispanic	-0.086*	0.039	
Native	0.043	0.205	
Other Race	-0.112	0.061	
International	-0.300**	0.037	
Fall17	-0.301**	0.054	
Fall18	-0.253**	0.055	
Spring18	-0.173**	0.024	
MATLAB (Adv Calc)	0.038	0.043	
MATLAB (Coding)	0.063	0.044	
MATLAB (None)	-0.015	0.031	
R^2			0.10
F for change in R^2			23.882

Note: B is unstandardized beta, and SE B is the standard error for the unstandardized beta.
*$p < .05$, **$p < .005$

Next, the effect of students' demographics, psychological safety, and team psychological safety on student's team member effectiveness were investigated. First, an unconditional model was run without any level 1 or level 2 predictors. Intracluster correlation coefficient (ICC) in this model shows that 24% of the variance in student's team member effectiveness was related to between-team differences, and 76% of the variance was attributable to within-student differences. The second step comprised a multilevel model of change by adding level 1 (individual) predictors (gender, race/ethnicity, citizenship status, GPA, previous MATLAB experience, attended semester, and individual psychological safety). This model accounted for 16% of within-student variation of student team-member effectiveness. Finally, level 2 (team level) added average psychological safety in teams as a predictor variable, and this model accounted for 8.7% of the within-team variation in student team-member effectiveness.

The fixed effect in the unconditional model was significant meaning that a student's team-member effectiveness is significantly greater than zero ($p < .001$). The random effects for the team-level were also significant indicating that student team-member effectiveness varies significantly across teams ($p < .001$). The residual was also significant, which means that there is significant within-student variation in student team-member effectiveness ($p <. 001$).

For the second model (adding individual-level predictors), the fixed effect for intercept was significant indicating that student team-member effectiveness is significantly greater than zero. For individual-level predictors, student GPA and individual psychological safety are significantly related to their team-member effectiveness ($p <. 001$). Students who specified their race as Black or Other had significantly lower team-member effectiveness than students of other races. Students enrolled Spring 2019 had significantly higher team-member effectiveness compared to previous semesters. Also, students with "Basic Calculations" or "Coding" experience with MATLAB programming environments had significantly higher team member effectiveness than students with other previous programming skill levels. Similar to the first model, the residual and the random effects for the team level were significant.

For the final model (adding team-level predictor), the fixed effects for intercept, GPA, Black students, students who specified their race as Other, attended semester, previous MATLAB skills, and the effect of average psychological safety in teams on student team-member effectiveness were significant meaning that student team-member effectiveness is significantly greater than zero, students' GPA significantly affect their team member effectiveness, Black students and students who identified their race as Other had significantly lower team member effectiveness than students with other races, students enrolled Spring 2019 had significantly higher team member effectiveness than students in previous semesters, students with previous MATLAB coding experience significantly had higher team member effectiveness than students with lower previous MATLAB skills, and average psychological safety in teams significantly affect the students team member effectiveness ($p <. 001$). The random effects for the team level were significant, indicating that student team-member effectiveness varies significantly across teams (p<.001). Residual also was significant meaning the student's team member effectiveness varies significantly across students (p <. 001). For the interaction main effects, only the interaction between GPA and team psychological safety was significant. See Table 3.8 for the full results of final model.

Table 3.8. Estimates two level model for the effect of students' demographics, individual psychological safety, and team psychological safety on student's team member effectiveness

	Team member effectiveness	
Fixed effects	Estimate	SE
Intercept	3.98***	0.02
GPA	0.32***	0.02
Individual Psychological Safety	0.03	0.01
Gender (Female)	0.02	0.02
Gender (other)	-0.20	0.11
Race (Asian)	-0.03	0.02
Race (Black)	-0.13*	0.06
Race (Hispanic)	-0.03	0.03
Race (Native)	-0.24	0.19
Race (other)	-0.12**	0.05
CS (International)	-0.02	0.03
Attended Semester (Fall 17)	0.43***	0.05
Attended Semester (Fall 18)	-0.21***	0.05
Attended Semester (Spring 19)	-0.05*	0.02
MATLAB (Basic Calculation)	0.04	0.02
MATLAB (Advanced calculation)	0.05	0.03
MATLAB (Coding)	0.13***	0.02
Team Psychological Safety	0.31***	0.03
Interaction of GPA and Team Psychological safety	-0.12***	0.04
Interaction of Asian and Team Psychological safety	0.00	0.06
Interaction of Black and Team Psychological safety	-0.07	0.15

Table 3.8 continued

Interaction of Hispanic and Team Psychological safety	-0.01	0.07
Interaction of International and Team Psychological safety	0.04	0.06
Random effects:		
Variance intercept team level	0.04***	0.01
Variance linear slope (team psychological safety)	0.02	0.02
Variance linear slope (Interaction of GPA and Team Psychological safety)	0.07*	0.04
Variance linear slope (Interaction of Race (Asian) and Team Psychological safety)	0.15**	0.05
Variance linear slope (Interaction of Race (Black) and Team Psychological safety)	0.11	0.13
Variance linear slope (Interaction of Race (Hispanic) and Team Psychological safety)	0.06	0.05
Variance linear slope (Interaction of CS (International) and Team Psychological)	0.06	0.06
Residual variance	0.18***	0.01
–2LL	5010.8	
AIC	5072.8	

Note. *AIC* = Akaike Information Criterion; *–2LL* = –2 Log Likelihood, relative model fit statistics.
* $p < .05$, ** $p < .01$, *** $p < .001$.

Finally, since I was interested to find out more about the interaction of team psychological safety on students' race. I repeated the final model removing the team psychological safety fixed effect. Removing this fixed effect identified additional results related to race/ethnicity; the interaction of team psychological safety with race/ethnicity was significant for Asian and Hispanic/Latino students but non-significant for Black students meaning that team psychological safety is positively related to the team-member effectiveness of Asian and Hispanic/Latino students, but this effect is not observed for Black students.

3.8 Discussion

The purpose of this study was to frame and investigate research questions examining the relationship of student demographics to their psychological safety and team-member effectiveness, followed by investigating the effect of individual and team psychological safety on student team-member effectiveness. Black students experienced the strongest negative effect in psychological safety and it was related to a lower perception of team-member effectiveness. This result is in consistent with the theoretical framework and the theory suggested by Newman et al. (2017). These results indicate that Black students have lower psychological safety in teams, so they may be less comfortable to speak up, provide feedback, or share their idea in the teams. This lower psychological safety is related to lower team-member effectiveness and team learning. Asian and Hispanic/Latino students also had lower psychological safety, but their lower responses did not result in significantly lower team-member effectiveness. Based on the proposed theoretical framework, those two racial/ethnic groups are gaining access to other resources may compensate for their lower psychological safety; specifically, they benefited from a higher team average psychological safety. Black students were not able to conserve resources because they did not benefit from working in teams with high team average psychological safety. For Black students in these engineering teams, diversity is functioning as disparity diversity rather than variety diversity, and social power differences prevent Black students from gaining the required resources to feel psychologically safe and have higher levels of team-member effectiveness (Edmondson & Roloff, 2009). This result might indicate that Black students need more resources in engineering teams to improve their individual psychological safety and team member effectiveness. Team composition strategies that avoid putting Black students in teams in which all other teammates are White might help—variability in team racial composition was not explored in this work. Instructors might also suggest informal activities that would improve social cohesion of teams and make Black students feel more included within the teams.

Gender was not related to student psychological safety and team-member effectiveness. It might be because most teams formed in a way to avoid isolating female students in teams as suggest by Rosser (1998). Individual psychological safety was another factor related to higher team member effectiveness, yet this effect was non-significant when considering the team psychological safety, suggesting that improvements to individual psychological safety are best achieved through interventions that target team psychological safety. This finding is consistent with studies about the influence of psychological safety on team learning (Leroy et al., 2012; Siemsen et al., 2009; Zhang et al., 2010).

International students had the lowest estimate in predicting psychological safety. The possible explanation for that is international students for whom English is not their first language have lower English-language skills in the first-year of study. As a result, they might have real and perceived difficulties in participating in team activities, which reduces their psychological safety. Also, due to cultural and language differences, they might answer the psychological safety questions differently. Cultural differences might also account for their lower psychological safety. International students may still be adjusting to U.S. culture, which makes them feel more uncomfortable in their teams. Nevertheless, the large reduction in the psychological safety of international students did not influence their team member effectiveness. According to our theoretical framework and the theory suggested by Newman et al. (2017), a possible reason for that could be that certain technical skills are serving as a resource gain, which helps them to function as and be identified as an effective team member.

GPA was another factor influencing the students' psychological safety and team member effectiveness. Students with low GPA had lower psychological safety and team-member effectiveness. Interestingly, when both GPA and team psychological safety are high, student team-member effectiveness decreases. Two possible explanations are offered; first, when students have a higher GPA and are more comfortable in their teams, they may be more likely to be critical of each other, resulting in lower ratings of team-member effectiveness relative to their higher expectations. A second explanation might be that students with lower GPA suffered a smaller reduction to their psychological safety in teams with high psychological safety.

Finally, MATLAB previous skills and attended semester had some significant effects. MATLAB previous experience was not related to psychological safety, but students with the highest level of previous experience in MATLAB had higher team member effectiveness. The result for team-member effectiveness was expected because the project required programming using MATLAB, so students with prior coding experience had advantage. The lack of relationship to psychological safety might indicate that teams were generally inclusive for students with lower coding skills. The higher observed psychological safety for students enrolled in the Spring 2019 semester may be an indication of improvement in team management practices over time.

3.9 Conclusion

In engineering programs, certain demographic groups are underrepresented and can be subject to minoritization. This study intended to explore the dynamics of psychological safety and its relationship to team-member effectiveness, because these are an indication of whether students have equal opportunities to learn in a team environment. Psychological safety is a team climate measure positively related to team learning and to students' ability to participate actively in teams. This study addressed a gap in the engineering education literature to describe the relationship of individual- and team-level psychological safety and team-member effectiveness for students with different demographics. Multilevel modeling revealed that some minoritized students did not benefit from individual psychological safety as much as other students. Asian, Black, Hispanic/Latino, international students, and students with lower GPA felt lower psychological safety comparing to their peers. These effects are additive, so the lowest psychological safety would be experienced by students who are Black, international, and have low GPA. Among these students, students who are Black or have lower GPA also had lower team-member effectiveness.

Team average psychological safety was positively related to team-member effectiveness. Based on these overall results, both individual psychological safety and team psychological safety are resources that increase the engineering students' team-member effectiveness, but team psychological safety was a stronger factor than individual psychological safety. Additional work is needed to provide Black students with a psychological safe team environment. Until that is achieved, ratings of their team-member effectiveness should be interpreted carefully to avoid acting on stereotypes and biases that may exist at the team level. Because team-average psychological safety has a stronger effect than individual psychological safety, instructional approaches are best focused on making the team environment psychologically safe. The lack of team psychological safety can also be addressed through facilitation of teamwork to make sure all team members can speak up, providing feedback without any fear, and respect new ideas. For instance, instructors can explain about the concept of psychological safety and its importance in team functioning. Then, students can practice idea generation in teams, and exchange feedback about whether they felt psychologically safe to speak up. Female students had high psychological safety and team-member effectiveness, possibly because the course that was the focus of this study forms teams to avoid isolating female students. These findings have some implication for the engineering education community, which are described in the last chapter.

4. STUDY 3

4.1 Abstract

Research shows that if engineering students do not learn how to properly and effectively work in teams, they will adopt counterproductive practices based on their misconceptions about teamwork. These misconceptions become more robust in the final years of undergraduate study if they are not addressed. Self-rating inflation is one of these problems, and I hypothesize that overestimation of self-ratings is related to lower teamwork skills and knowledge. This study investigates this premise using a multilinear regression model. This study further explores the relationship of self-rating inflation to student characteristics such as gender, race/ethnicity, citizenship status, previous MATLAB skills, and GPA. Female students and students who contributed and interacted more in their teams were observed to have less self-rating inflation. Students with higher GPAs and a higher level of prior programming skill had higher self-rating inflation. The results of this study provided useful information about how to interpret student self-evaluations. The results inform the extent to which students in this particular context are laying a foundation for the counterproductive practices seen in other research.

4.2 Introduction

Professional skills are widely studied and acknowledged to be a critical companion to technical skills (Downing, 2001; Florman, 1997; Lang, Cruse, McVey, & McMasters, 1999; Meier, Williams, & Humphreys, 2000; Passow & Passow, 2017; Scott & Yates, 2002). In recent decades, engineering curricula have introduced more professional skills than before and adopted more student-centered pedagogies. Teamwork is one of the professional skills specifically included in accreditation criteria. Engineering graduates should have the ability to function effectively on a team (ABET, 2019). Teamwork also helps students to acquire other competencies as well, such as communication, problem-solving skills, and time management.

To ensure that team experiences fulfill their potential as learning experiences, student teams must function effectively. Team learning is the process of acquiring teamwork knowledge and skills through collaboration and working cohesively on a specific task to reach the team's goal (Nepal, 2016). Team learning does not occur by just simply putting students in a team and asking them to finish a task. Instructors should inform students about the importance of teamwork, teach students teaming skills, improve cohesion in teams using teaming exercises, use an appropriate method for team formation, set a clear goal for team projects, provide feedback, and consider peer evaluations in grading (Hansen, 2006). Students also can improve their team learning by participating more in team activities and increasing their self-awareness about their teaming skills.

Peer evaluation, as a tool that contributes to formative and summative assessment, has a significant role in team learning, so it is important to explore issues that arise in peer evaluation. This study includes greater discussion of the peer evaluation process and related research findings followed by an investigation of what factors are related to self-rating inflation by engineering students.

4.3 Literature Review

4.3.1 Teamwork in engineering education

Due to the importance of teamwork skills, there are many studies of teamwork in the engineering context. Some of that literature focuses on strategies for forming teams to be successful (Kowalski & Smyser, 2017; Parker et al., 2019; Layton et al., 2010; Parker et al., 2019; Steiner & Kanai, 2016). Other research has focused on strategies for teaching students to work in teams (Cinar & Biglin, 2011; Downing, 2001; Fiegel & Denatale, 2011; Godwin et al., 2017; Luechtefeld et al., 2008; Natishan et al., 2000; Nepal, 2016; Novoselich & Knight, 2019). Still other researchers have studied various team dynamics (Beigpourian, Ferguson, et al., 2019; Lewis et al., 1998; Natishan et al., 2000; Asio et al., 2018; Baughman et al., 2019; Beigpourian, Luchini, et al., 2019; Borrego et al., 2013; Neumeyer & Santos, 2020; Whitman et al., 2005).

Yet even if student teams are well-formed and well-taught, regular monitoring of team functioning is important to ongoing success. Free-riding is a common problem in teams and occurs for a variety of reasons. Free-riding occurs when some team members benefit from the team's work without contributing their fair share (Tenenberg, 2019). Regardless of the cause, a common approach to observing and remediating free-riding is the use of peer evaluation.

4.3.2 Self and peer evaluation

Self and peer evaluation not only alert instructors to free riding, but also help instructors identify conflict, adjust grades on team assignments, notice low-performing teams, encourage students to communicate more in teams and learn about teaming skills, develop students' problem-solving and self-regulation skills, and improve students' academic results (Alba-Flores & Rios, 2019; Cinar & Biglin, 2011; Felder & Brent, 2007; Hansen, 2006; Jassawalla, Sashittal, & Sashittal, 2009; Kilic, 2016; Wang et al., 2016; Willey & Gardner, 2010). Alba-Flores and Rios (2019) used peer evaluation in a capstone course to assess students' writing assignments, presentations, and team performances. They found that peer evaluation was useful in increasing students' presentation skills. Peer evaluation is a valid tool to measure individual performance and can improve the learning process in the classroom and students' critical thinking, motivation, and monitoring (Cinar & Biglin, 2011; Liu, Lin, Chiu, & Yuan, 2001; Wang et al., 2016). Using peer evaluation, students can help their peers to identify and capitalize on opportunities for personal development (Topping, 1998). It gives students more opportunities to learn and a different perspective from an instructor's evaluation (Liu et al., 2001) and can be even more effective if used frequently (Willey & Gardner, 2010).

Despite many similarities, some researchers distinguished self and peer evaluation effectiveness (Nepal, 2018; Sande & Godino-Llorente, 2014). Both self and peer evaluation offer students unique opportunities. Peer evaluation is a useful tool for reciprocal learning, and self-evaluation is good for insightful learning, (Marcangelo et al., 2010). Sande and Godino-Llorente (2014) compared the effectiveness of self and peer evaluation for students' problem-solving skills. They found that peer evaluation is more similar to an evaluation conducted by instructors and is more likely to improve students' performance than self-evaluation.

However, there are some challenges in self and peer evaluation. According to some studies, there might be bias in self and peer evaluation (Luthar, 1996; Ostafichuk & Sibley, 2019; Walker, 2001). Cinar and Bilgin (2011) found that international (non-Australian) students tend to be more lenient compared to Australian students in rating everyone in teams (including themselves). Students might be worried about the fairness of assessment or be unconfident in giving honest feedback; they might think that assessment is an instructor's job; a personal relationship might influence the ratings; instructors might be skeptical about the validity of peer evaluation; and instructors might be uncomfortable to give students the power to evaluate their peers (Falchikov, 2004; Marcangelo et al., 2010). Some students disrupt the peer evaluation process by giving the maximum rating to all team members or giving themselves disproportionately high scores (Cinar & Biglin, 2011).

Research in a mechanical engineering class found that students inflated their self-rating, and this self-rating inflation increased in subsequent peer evaluations. They also found that gender and personality influenced the self-rating bias. Men had a greater self-bias in first peer evaluation, but there was no gender difference in the later peer evaluations. Helping to explain the concern that some students have for evaluating peers, they also found that introverted students and students with a "Feeling" personality type (base on Myers-Briggs Type Indicator scale) had a lower self-rating bias (Ostafichuk et al., 2015).

Although some adjustment factors, mathematical models, normalization, and team formation strategies can reduce or adjust for biases (Cinar & Biglin, 2011; Nepal, 2018; Ohland et al., 2012), the better approach is to teach students about self and peer evaluation a priori (Loignon et al., 2017). Peer evaluation is most effective if instructors give students clear criteria, help them interpret those criteria, and monitor the peer evaluation process. Self-evaluation improves if instructors scaffold students' reflection and self-awareness skills, reward students for identifying and acting on opportunities for growth, and urge students to have regular debriefings (Beaman, 1998; Marcangelo et al., 2010; Rust, Price, & O'Donovan, 2003).

Some studies have proven that students who participate more in team activities evaluate themselves more accurately (or even underestimate themselves), but students with less involved in their teams tend to overestimate their skills (Atwater & Yammrino, 1992; Lejk & Wyvill, 2001), which is consistent with a general tendency to reduce self-rating bias as expertise increases, known as the Dunning-Kruger effect (Kruger & Dunning, 1999). This study adds to the engineering education literature to study the extent of self-rating bias in that context and its relationship to student characteristics.

4.3.3 Scope and research questions

The review of the literature on peer evaluation indicates the extent to which self and peer evaluation are necessary for team management and the challenges inherent in conducting them. Knowing more about peer evaluation helps researchers and instructors develop better practices for forming teams and teaching teamwork. Based on previous literature, this study aims to investigate the factors that are related to self-rating inflation in a first-year engineering class.

Students begin engineering study with varied teaming skills and conceptions about proper teaming practice. This study explores how students' knowledge and teaming skills are related to their self-evaluations, as well as the relationship of gender, race/ethnicity, citizenship status, and GPA to self-rating inflation. I investigate these two research questions:

> RQ1: To what extent are a student's interactions with teammates (as assed by teammates), gender, race, GPA, and citizenship status related to self-rating bias (the ratio of self to peer ratings) of a student's knowledge, skills, and abilities?
>
> RQ2: To what extent are a student's contribution to their team (as assed by teammates), gender, race, GPA, and citizenship status related to self-rating bias (the ratio of self to peer ratings) of a student's knowledge, skills and abilities?

This study adds value to engineering education research and practice in multiple ways. First, by finding factors that are related to self-rating inflation, engineering instructors can design approaches to reduce this self-rating inflation to encourage more effective team experiences, improving their employability and making them more productive employees. Second, this study targets first-year engineering students because, according to research, it is vital to address counterproductive practices early (Leonardi et al., 2009). Third, the relationship of team-member effectiveness and self-rating inflation has not been studied previously in engineering students. Finally, by studying the relationship of self-rating inflation to individual characteristics, this research will inform the interpretation of self and peer evaluation ratings.

4.4 Theoretical Framework

Kruger and Dunning (1999) show in multiple contexts that individuals who lack a particular skill are particularly unable to assess their own skill level. Individuals with low skills overestimate their skill level significantly, and this self-assessment bias decreases with increased skill until it becomes negative; individuals with high skills typically underestimate their skill level. If this finding holds in the domain of teamwork, students with higher teamwork-related skills would be less likely to overestimate their skills. Leonardi et al. (2009) found that students start undergraduate engineering with misconceptions about proper teamwork practice. For example, students viewed individual accomplishment as the measure of success for an engineer. They found that due to the lack of adequate teaming knowledge among engineering students, students deepen their commitment to practices counterproductive to teamwork in the later years of their education. Integrating the findings of Kruger and Dunning into the framework by Leonardi et al., I expect students with low knowledge about teaming skills might have more bias in self-evaluation, and that bias would worsen if those students do not receive adequate instruction in teamwork.

Measuring bias in peer evaluation is not an easy task because there is no way to measure students' true scores unless an expert observes all of a team's behaviors. Self-other agreement (SOA) theory proposes that the extent to which a person's self-rating is in agreement with ratings by others is good indicator of self-awareness (Atwater & Yammarino, 1997). Similarity of self-evaluations and ratings by others are related to a better perception of skills (Fleenor, Smither, Atwater, Braddy, & Sturm, 2010). The relation of self-ratings to ratings by others is therefore a useful indicator of bias in peer evaluation. Research using SOA also showed that individual characteristics such as gender, race/ethnicity, position, and culture can influence self-other rating agreement (Fleenor et al., 2010). For example, men tend to overestimate their rating more than women (Brutus, Fleenor, & McCauley, 1999; Moshavi, Brown, & Dodd, 2003; Vecchio & Anderson, 2009; Visser, Ashton, & Vernon, 2008). Figure 4.1 depicts the proposed theoretical framework for this study.

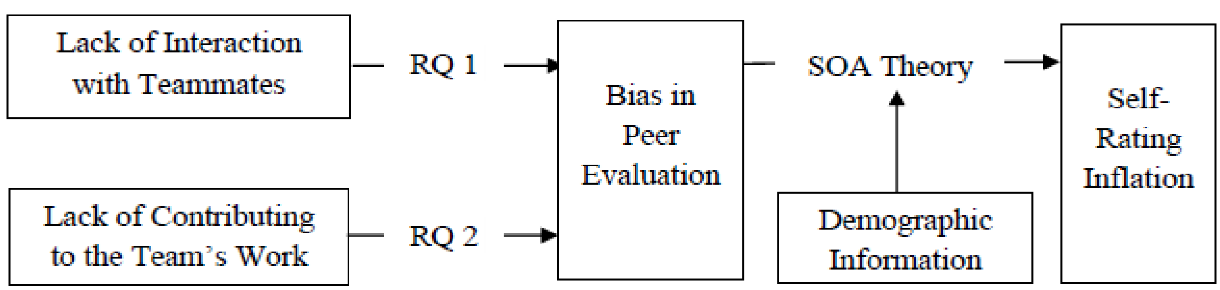

Figure 4.1. Theoretical framework for self-rating inflation

Based on this theoretical framework, lower levels of interaction with teammates (RQ1) and lower of contribution to the team's work (RQ2) are associated with a larger bias in peer evaluation. These biases are affected by students' demographics as well. The bias is also measured by self-rating inflation based on SOA theory.

4.5 Method

4.5.1 Participants of the Study

I conducted this study on students who attended a first-year engineering class. This data includes 3651 engineering students, of which 3271 students and 940 teams completed the peer evaluation instrument. Among the participants, 187 students were enrolled in Fall 2017, 1426 in Spring 2018, 153 in Fall 2018, and 1505 in Spring 2019. The sample included 758 students who identified as female, 2492 as male, and 21 selected "Other/Prefer not to answer". For students' race/ethnicity, 784 reported Asian, 71 Black, 320 Hispanic/Latino, 7 Native American, 1893 White, 114 Other/Mixed-heritage, and 82 students selected Other/Prefer not to answer. Finally, 571 students were international and 2700 were domestic students.

4.5.2 Data Sources

Data were collected at the end of the semester, after all teaming activities were completed, using CATME (the Comprehensive Assessment of Team Member Effectiveness). CATME is a web-based system that includes tools for team formation and peer evaluation (Layton et al., 2010; Ohland et al., 2012). CATME's peer evaluation tool measures team-member effectiveness in five dimensions; "*Contributing to the team's work*," "*Interacting with Teammates*," "*Keeping the team on track*", "*Expecting quality*," and "*Having Relevant Knowledge, Skills, and Abilities*." CATME uses a behaviorally anchored rating scale (BARS) in which after reading behavioral descriptions in each dimension, students identify which set of descriptions best describes the behavior of each team member. These ratings are converted to the ordinal numbers ranging from 1 (equal to the lowest team member effectiveness skill) to 5 (equal to the highest team member effectiveness skill). These dimensions scales are available in Appendix A.

For the dependent variable in both research questions, the ratio of students' self-ratings to ratings by peers on "*Having Relevant Knowledge, Skills, and Abilities*" was calculated. This measures self-rating inflation because based on SOA theory, students with a better knowledge of their skills will have ratings similar to their ratings by others. Gender, race, GPA, previous programming skills, and citizenship status as our independent variables. Students specified their previous programming skill from no previous experience in MATLAB (None) to coding experience (Coding). Basic calculation skills (Basic Calc) and advanced calculation skills (Adv

Calc) were the two middle options. For the first research question, I added the rating that students received by their team members on the "*Interacting with Teammates*" dimension as additional independent variable. In the second research question, I added the rating that students received by their team members on the "*Contributing to the team's work*" dimension as an additional independent variable. Tables 4.1 and 4.2 describe the variables in this study and their alignments with the research questions.

Table 4.1. Levels of measurement for variables

Variable	**Level of Measurement**
Student's interaction with teammates	Scale average: Continuous Individual items: Ordinal
Student's contributing to teams	Scale average: Continuous Individual items: Ordinal
Gender	Scale: Nominal
Race/Ethnicity	Scale: Nominal
Citizenship status	Scale: Nominal
GPA	Scale: Continuous
Previous MATLAB skills	Scale: Nominal
Self to peer rating on KSAs	Scale average: Continuous Individual items: Ordinal

The first research question was modeled as:

$$SO_i = \beta_0 + \beta_1(\text{Gender}_i) + \beta_2(\text{Race}_i) + \beta_3(\text{CS}_i) + \beta_4(\text{GPA}_i) + \beta_5(\text{Stud Interaction}_i) + \beta_6(\text{MATLAB}_i) + r_{ij} \quad \text{(Eqn 4.1)}$$

In this model, SO_i is a student's self-rating of teamwork divided by their rating by peers on the student's related knowledge, skills, and abilities, which is predicted by the model intercept (β_0), the main effects of gender (β_1), race/ethnicity (β_2), CS which is citizenship status (β_3), GPA (β_4), students' interaction with teammates (β_5), and previous programming skills (β_6). The value r_{ij} represents random error.

The second research question was modeled as:

$$SO_i = \beta_0 + \beta_1(\text{Gender}_i) + \beta_2(\text{Race}_i) + \beta_3(\text{CS}_i) + \beta_4(\text{GPA}_i) + \beta_5(\text{Stud Contribution}_i) + \beta_6(\text{MATLAB}_i) + r_{ij} \quad \text{(Eqn 4.2)}$$

This model is similar to the first model except that interaction with teammates has been replaced by students' contributions to the team's work (β_5).

Table 4.2. Alignment of research questions, dependent variables, independent variables, and data analysis

Research Question	Dependent variable	Independent variables	Data Analyses
RQ 1: To what extent are a student's interactions with teammates (as assed by teammates), gender, race, GPA, and citizenship status related to self-rating bias (the ratio of self to peer ratings) of a student's knowledge, skills, and abilities?	Ratio of self-ratings of teamwork to ratings by peers on "*Having Relevant Knowledge, Skills, and Abilities*" dimension of CATME.	• Gender • Race • Citizenship status • GPA • Previous MATLAB skill • Other students' ratings of "*Interacting with Teammates*" dimension of CATME	Multiple linear regression model
RQ 2: To what extent are a student's contributing to teams (as assed by teammates), gender, race, GPA, and citizenship status related to self-rating bias (the ratio of self to peer ratings) of a student's knowledge, skills and abilities?	Ratio of self-ratings of teamwork to ratings by peers on "*Having Relevant Knowledge, Skills, and Abilities*" dimension of CATME.	• Gender • Race • Citizenship status • GPA • Previous MATLAB skill • Other students' ratings of "*Contributing to the team's work*" dimension of CATME	Multiple linear regression model

4.6 Results

The data were checked for multiple regression assumptions. The dependent variable was in continuous scale and the independent variables were either continuous or categorical. The data were winsorized to address extreme outliers, and the data met the requirements related to multivariate normality, multicollinearity, and homoscedasticity. Then, both multiple regression models were conducted to predict the first-year engineering students' self-rating bias based on their demographics' information, their scores in the interaction with teammates (RQ1), and contributing to the team's work (RQ2). In the first regression model, the interaction with teammates score and previous programming experience using advanced calculations was associated with a significantly reduced self-rating bias on students KSAs. No significant relationship with race/ethnicity, gender, GPA, or citizenship status was observed in this model. Table 4.3 summarizes the result for the first multiple regression model.

Table 4.3. Modeling students' self-rating bias in terms of their interaction with teammates score and demographic information

Variable	*B*	*SE B*
Intercept	1.54**	0.03
Interaction with teammates	-0.12**	0.01
GPA	-0.01	0.01
Female	-0.01	0.01
Other/Prefer not to answer	-0.02	0.04
Asian	0.01	0.01
Black	0.00	0.02
Hispanic/Latino	0.014	0.01
Native American	0.03	0.07
Other/Prefer not to answer	0.05	0.02
International	0.01	0.01
MATLAB (Adv Calc)	-0.03*	0.01
MATLAB (Coding)	0.00	0.01
MATLAB (None)	-0.01	0.01
R^2	0.14	
F for change in R^2	40.64	

Note: B is unstandardized beta, and SE B is the standard error for the unstandardized beta. *$p < .05$, **$p < .005$

For the second regression model, students' GPA was significantly related to self-rating bias on student's KSAs. Contributing to the team's work score, identifying as female, and having no previous programming experience were associated with a significantly reduced self-rating bias on students KSAs. The effect of race, and citizenship status was not significant in this model. Table 4.4 presents the result for the second multiple regression model.

Table 4.4. Modeling students' self-rating bias in terms of their contributing the team's work score and demographic information

Variable	*B*	*SE B*
Intercept	1.49**	0.03
Contributing to the team's work	-0.13**	0.01
GPA	0.02*	0.01
Female	-0.03**	0.01
Other/Prefer not to answer	-0.04	0.04
Asian	0.02	0.01
Black	-0.01	0.02
Hispanic	0.01	0.01
Native	0.02	0.07
Other/Prefer not to answer	0.03	0.02
International	0.01	0.01
MATLAB (Adv Calc)	-0.02	0.01
MATLAB (Coding)	0.01	0.01
MATLAB (None)	-0.02*	0.01
R^2	0.18	
F for change in R^2	59.28	

Note: B is unstandardized beta, and SE B is the standard error for the unstandardized beta. *$p < .05$, **$p < .005$

4.7 Discussion

Results supported the proposed theoretical framework. Low ratings of interaction with teammates and contributing to the team's work were associated with increased bias in self-rating on student's KSAs. A general bias in self-evaluations was observed (the intercepts were significant), but as ratings of interaction with teammates and contributing to the team's work increased, this bias decreased. When students are more effective at interacting with teammates, they may receive more feedback from their team members and are more aware about their weaknesses. On contrary, students who do not interact well with their teammates may be likely to receive and active on advice or help from teammates, and are more likely to overestimate their skills. This finding is consistent with Leonardi et al.'s (2009) findings that engineering students who prefer to accomplish work alone have lower knowledge about their skills. Similarly, when students make more and better contributions to the team's work, the quality of feedback receive from their teammates will be better, leading to a better self-awareness about their skills. This result is similar to another study's result about role of self-awareness in decreasing bias in self-evaluation (Fleenor et al., 2010).

In the second model, when I considered the effect of contributing to the team's work, female students had significantly less self-evaluation bias than male students. This finding was consistent with other research (Brutus et al., 1999; Moshavi et al., 2003; Vecchio & Anderson, 2009; Visser et al., 2008), but this effect was not observed in the first model. It may be that students identifying as female students generally have higher ratings of their interaction with teammates, an effect that cannot be observed when the interaction score is a dependent variable. The statistical interaction of gender and the interaction with teammates score was not included in this study because it violated the regression assumptions.

Considering the effect of contributing to the team's work (second model), GPA had a significant relationship to students' self-rating bias. Students with higher GPA had more bias in their self-rating. GPA is a measure of academic achievement in other courses, whereas Leonardi et al. (2009) and others would suggest that teamwork is a separate skill that requires instruction and practice. Similar to GPA, in the second model, students with lower previous programming skills had lower bias than students with highest previous programming skills. In the first model, however, considering the effect of interaction with teammates, students with previous programming experience with advanced calculation had lower self-rating bias than students with previous coding experience. It is possible that students that had experience with advanced calculations are confident enough to interact more effectively, provide more feedback to their team members, and learn more about their teaming skills, reducing self-rating bias. At higher levels of previous programming skills, this reduction in self-bias may be lost because students become dissatisfied with working with students of lower skill levels as predicted by Leonardi, et al. (2009).

4.8 Conclusion

This study integrated the findings of multiple prior research studies to examine the predictions that students who contribute less to the team's work and/or are not as good at interacting with teammates will be less self-aware about their skills and have a higher self-rating bias. Based on other research, it was anticipated that this bias could also be associated with students' demographic identification and academic background. The findings show some support for the theoretical framework in data collected on the self-rating behaviors of first-year engineering students. These results highlight the importance of training students about the proper teaming skills. If instructors just form teams without any training or facilitation and ask students to evaluate themselves, their self-ratings are not likely to be valid because their lack of knowledge about proper teaming practice will make it impossible for them to evaluate teamwork skills in themselves and others. This self-rating bias would generally be worse among students with higher academic standing. Without the ability to accurately evaluate their teaming skills, they would inflate their self-ratings by assuming that they would be on par with their technical skills, academic GPA, or prior programming skill. Students learn teaming skills and have lower self-evaluation bias when they provide and seek feedback from team members and have higher levels of commitment to the team's work. When self-ratings are considered in adjusting student grades, it would be appropriate to consider the gender bias in self-evaluation; women should not be penalized for having a lower self-evaluation bias. Further implications of this study in light of the complete set of studies are provided in the last chapter.

5. CONCLUSION

5.1 General Discussion and Conclusion

Engineers spend most of their time using their collaboration skills rather than their technical skills (Trevelyan, 2014), and engineers identified teamwork as the most essential professional skill in their workplace (Passow, 2012). Due to the importance of teamwork, this dissertation serves to deepen understanding of engineering student team processes through research on team formation, team dynamics, and peer evaluation. In each of these studies, students' experiences have been investigated based on their gender, race/ethnicity, citizenship status, GPA, and programming skills. By including student demographics and knowledge in this study, the results inform approaches that can make teamwork in engineering education more inclusive.

Women's experiences in engineering classrooms have been of interest to many engineering education scholars. In this dissertation, I studied the relationship of mixed-gender teams on psychological safety and self-rating bias. Mixed-gender teams have a broader skillset that should make these teams more effective than single-gender teams if mixed-gender teams are managed well. The first study found that although mixed-gender teams have similar team dynamics to all-male teams, mixed-gender teams' potential to have better team dynamics has not been realized. This finding highlighted the need to teach students how to capitalize on gender differences in teams in way to understand each other skills and utilize them to improve team dynamics. Adopting the Mapping, Bridging, Integrating approach might help teams manage differences in terms of gender, race/ethnicity, and culture. This framing could lead to improved team dynamics and higher satisfaction. In this approach, students initially map differences that have the potential to improve innovation and performance. Then team members communicate across these differences to bridge and integrate them.

Women had a significantly lower self-rating bias than men, resulting from some combination of heightened self-awareness and (possibly culturally conditioned) modesty. A heightened self-awareness seems to be part of the explanation for women's lower self-rating bias, because women's higher levels of ability in interacting with teammates would tend to lead to that awareness. Instructors should avoid unintentionally penalizing women by considering the self-rating in grade adjustments without adjusting their scores based on self-rating bias. Whereas Study 1 did not find the higher levels of team dynamic expected for mixed-gender teams, Study 2 provides some positive indicators of the team climate for women in the class studied. Women reported similar levels of psychological safety as men as well as similar team member effectiveness. This result might support the practice of avoiding the formation of teams in which women are isolated, but this research did not study other team configurations. To the extent that this team formation practice leads to better outcomes for women, it is certainly the preferred approach in the early years of their education, when it is most important to ensure that women have the opportunity to learn and engage in positive team experiences.

Race/ethnicity did not moderate women's satisfaction in teams, and there was no self-rating bias based on race/ethnicity, but race/ethnicity was significantly and negatively related to psychological safety and team member effectiveness. Asian, Hispanic/Latino, and Black students reported significantly lower psychological safety. This is problematic because lower psychological safety was also associated with lower team-member effectiveness. Study 2 found that the average psychological safety of a team moderated that relationship for Asian and Hispanic/Latino students, but lower psychological safety was still related to lower team-member effectiveness for Black students. This result points to psychological safety as an important indicator of the climate for minoritized students in engineering teams. If minoritized students are not comfortable to speak up, share their ideas, and provide feedback, they will appear to be less effective in team ratings. These issues are particularly important in first-year engineering students, where students are developing habits and engaging interactions that affect student's future learning and their interest and willingness to persist.

International students had the lowest individual psychological safety, suggesting that they might have issues in teams regarding cultural differences and language skills as English is not their first language. This result was consistent with the finding in Study 1, where domestic women were less satisfied when they were placed in teams with male international students; teams need to map and bridge their gender and culture differences, and the wider the gender and culture gap, the more effort is needed. Despite the low individual psychological safety reported by international students, they did not have low reported team-member effectiveness as was the case for racial/ethnic diversity. International students had high team-member effectiveness, possibly because of their high coding skills. Whereas Leonardi, et al. (2009) found that certain counterproductive team practices worsen as students progress through the engineering curriculum, there are indications that team dynamics of multicultural teams improve even after a single multicultural team experience (Jimenez-Useche, Ohland, & Hoffmann, 2015).

These studies provide additional evidence for how previous knowledge and technical skills influence the engineering students' teaming experiences. While having a high GPA and previous programming skills were associated with higher psychological safety and team-member effectiveness, they were also related to a larger self-rating bias. This finding highlights the need to assess and give feedback on both team skills and technical skills. These high-performing students likely overestimate their team skills because their feedback from most academic assessments has been positive.

Instructor effects were not considered in this dissertation, but the higher psychological safety observed in Spring 2019 might suggest that instructional teams had made improvements over time either in teaching students to manage diversity or in facilitating teams. This finding might highlight the importance of engineering instructors' familiarity with effective team assessment and team dynamics.

In general, this dissertation shows that the team experiences of engineering students can be influenced by team formation, team dynamics, and peer evaluation. These influences might differ based on students' demographics and previous knowledge. It is crucial to consider both individual student experiences and team dynamics in forming teams, and instructors should incorporate regular peer evaluation and assessment of team dynamics. Peer evaluation is a useful tool for measuring individual team member effectiveness, whereas team dynamics are more related to the teaming process. With better team dynamics, students can work within their teams to increase their team-member effectiveness. In turn, team-member effectiveness enhances students' self-awareness of their teaming skills, which increases their ability to learn through teamwork. Feedback and facilitation is more important in teams with higher levels of diversity, such as mixed-gender teams and teams with international students. These teams need help to recognize, bridge, and integrate these differences.

5.2 Implication

Within the limitations reported, these studies addressed some existing gaps in the understanding of engineering student teams, resulting in several implications for engineering education:

1. Mixed-gender teams have a broader skillset and should have better team dynamics than same-gender teams. However, they were observed to have team dynamics similar to all-male teams. Engineering instructors should provide appropriate training for students to assist them in map, bridge, and integrate these gender differences in a way to improve the team dynamics and achieve that higher level of team-member effectiveness.
2. Women's satisfaction in mixed-gender teams is a complex issue, and mixing national/cultural diversity with gender diversity was associated with lower satisfaction. Students can mitigate this negative effect by communicating across both gender and culture differences, and understand that these differences result in broader skillsets in teams. Further research in this area, both qualitative and quantitative, is needed.
3. Psychological safety is a significant indicator of whether the team environment is inclusive of minoritized students. In addition to measuring psychological safety, instructors can observe teams for signs of low psychological safety and seek opportunities to facilitate.
4. Higher team psychological safety compensated for lower individual psychological safety, so instruction based on improving psychological safety generally should be productive. The team average psychological safety is also an indicator of teams requiring facilitation. If instructors address the low team psychological safety at the earlier stages of teaming, that can improve the team climate for individual students who feel less psychologically safe.
5. Black students were the most marginalized group based on individual psychological safety. Black students' psychological safety continued to be related to lower team-member effectiveness even when team psychological safety was included as a moderator. More systematic effort is needed to create a more inclusive climate for Black students in predominately White institutions. This might include purposeful team composition strategies that avoid isolating Black students in teams, which is not always possible due to low representation. Putting Black students in teams with

Hispanic/Latino and Native students might help address this, which should be the subject of further quantitative and qualitative study. Another strategy would be to suggest informal activities for engineering student teams to improve social cohesion. Monitoring Black students' perception of team dynamics could be useful to address issues earlier.

6. The free exchange of feedback improves team-member effectiveness, yet engineering students tend to exclude students they perceive as having inferior skills. This research provides evidence that this practice is hurting both low-GPA students, who have lower psychological safety and team-member effectiveness and high-GPA students, who have lower self-awareness resulting from a self-evaluation bias. Students need guidance on giving, receiving, and acting on feedback.
7. Both peer evaluation and team dynamic measurement are necessary for diagnosing and improving the students' experience in teams. Peer evaluation provides information about individual students and some information on the team as a whole, whereas team dynamics speak about the team as a whole. By using, instructors get a more complete picture of team functioning and its effect on individual students.
8. Students' misconceptions about teaming practice must be addressed early in the engineering curriculum. Education about proper team practice, period assessment, and feedback on teamwork can help students diagnose and repair their misconceptions.

5.3 Future Direction

Besides the implications for engineering education practice above, this research raises research questions for future research, which are summarized below:

1. In the first study, I investigated the experience of mixed-gender teams as a whole. Additional studies to address the experience of individual men and women in these teams would bring more value for engineering education, as would further research on different team compositions.
2. In the first study, all teams have worked on the same team task. Considering the effect of team task's complexity on mixed-gender team dynamic could be the subject of another study.

3. Whereas this dissertation addressed four significant team dynamics in organizational behavior, there are others that might be investigated in similar ways, such as cohesion and trust.
4. There are additional student characteristics that might be investigated as well. Students' personality, nationality/culture, language, and the effect of intersectionality are among those characteristics that would be useful to study.
5. The first study showed that there is more to learn about the factors that affect the satisfaction of women in engineering teams. Further study could identify appropriate frameworks and design a model to investigate this issue more.
6. This study investigated the psychological safety of the individual and team level. There is potential to explore this more broadly, such as at the section, instructor, or class level. Each would provide useful information, but there would certainly not be enough data to include all those levels at once.
7. Instructors and teaching assistants are likely to have a significant effect on students' psychological safety. This effect could be studied with a different level structure in the multilevel modeling, but might also be addressed through other research approaches such as participant diaries and observation. A research project underway in this setting is likely to provide additional insights on important subject (Pawley, Dickerson, Ohland, & Zywicki, 2019).
8. Students are not the only ones who need psychological safety to realize their potential. Psychological safety could be studied and improved among faculty and staff of academic institutions. In addition to the likelihood of race/ethnicity and gender affecting psychological safety in academic institutions, the power disparities associated with faculty rank almost certainly play a role, and psychological safety is likely to be a good measure of whether junior faculty and non-tenure-track faculty feel safe to share their ideas and concerns.
9. This dissertation addressed students' team experiences throughout one semester. A longitudinal study that tracks students' team skill development would provide useful insights to engineering education.
10. As with other skill development, students might improve their teaming misconceptions and develop more quickly as effective team members if they collaborate with students

who have more favorable views of teamwork and are more effective as a team member. The former might be measured in advance using a survey and considered in team formation. The survey would require careful development, however, to avoid having students choose what they believe to be the preferred answer rather than what they really believe. Framing the survey as a tool for gathering information about students' prior team experiences might avoid that concern.

11. One of these changes might be paying more attention to the social aspect of engineering. Where engineering identities and practices have been separated into "the social" or "the technical," most engineers might prefer the technical identity. Women may be more motivated in engineering teams if we address the social purposes of engineering projects because some women prefer design projects with an emphasis on social issues in addition to technical aspects. These changes show promise not only for women but also for men. Many men also indicate particular interest in engineering problems that address social problems (Faulkner, 2007; Ngambeki et al., 2012; Weber & Custer, 2005).

APPENDIX A. SURVEYS

A 5.1. Team dynamic survey questions

Question Type	Sub-Question
Psychological Safety* (Edmondson, 1999)	If you make a mistake on this team, it is often held against you. (reversed scale)
	Members of this team are able to bring up problems and tough issues.
	People on this team sometimes reject others for being different. (reversed)
	It is safe to take a risk on this team.
	It is difficult to ask other members of this team for help. (reversed scale)
	No one on this team would deliberately act in a way that undermines my efforts.
	Working with members of this team, my unique skills and talents are valued and utilized.
Task Conflict** (Jehn & Mannix, 2001)	How much conflict of ideas is there in your work group?
	How frequently do you have disagreements within your work group about the task of the project you are working on?
	How often do people in your work group have conflicting opinions about the project you are working on?
Relationship Conflict** (Jehn & Mannix, 2001)	How much relationship tension is there in your work group?
	How often do people get angry while working in your group?
	How much emotional conflict is there in your work group?
Process Conflict** (Jehn & Mannix, 2001)	How often are there disagreements about who should do what in your work group?
	How much conflict is there in your group about task responsibilities?
	How often do you disagree about resource allocation in your work group?
Interdependence** (Vegt et al., 2001) <u>(with minor modification)</u>	My teammates and I have to obtain information and advice from one another in order to complete our work
	I depend on my teammates for the completion of my work
	I have a one-person job; I rarely have to check or work with others (scale reversed)
	I have to work closely with my teammates to do my work properly
	In order to complete our work, my teammates and I have to collaborate extensively
Satisfaction** (Vegt et al., 2001) <u>(with minor modification)</u>	I am satisfied with my present teammates
	I am pleased with the way my teammates and I work together
	I am very satisfied with working in this team

*Using very inaccurate to very accurate scale (Seven-point Likert scale)
**Using strongly disagree to strongly agree scale (Five-point Likert scale)

A 5.2. Contributing to Team's Work dimension (BARS scale)

Score	Contributing to Team's Work
5	• Does more or higher-quality work than expected. • Makes important contributions that improve the team's work. • Helps teammates who are having difficulty completing their work.
4	Demonstrates behaviors described immediately above and below.
3	• Completes a fair share of the team's work with acceptable quality. • Keeps commitments and completes assignments on time. • Helps teammates who are having difficulty when it is easy or important.
2	Demonstrates behaviors described immediately above and below.
1	• Does not do a fair share of the team's work. Delivers sloppy or incomplete work. • Misses deadlines. Is late, unprepared, or absent for team meetings. • Does not assist teammates. Quits if the work becomes difficult.

A.3. Interacting with Teammates dimension (BARS scale)

Score	Interacting with Teammates
5	• Asks for and shows an interest in teammates' ideas and contributions. • Makes sure teammates stay informed and understand each other. • Provides encouragement or enthusiasm to the team. • Asks teammates for feedback and uses their suggestions to improve.
4	Demonstrates behaviors described immediately above and below.
3	• Listens to teammates and respects their contributions. • Communicates clearly. Shares information with teammates. • Participates fully in team activities. • Respects and responds to feedback from teammates.
2	Demonstrates behaviors described immediately above and below.
1	• Interrupts, ignores, bosses, or makes fun of teammates. • Takes actions that affect teammates without their input. Does not share information. • Complains, makes excuses, or does not interact with teammates. • Is defensive. Will not accept. • help or advice from teammates.

A.4. Keeping the Team on Track dimension (BARS scale)

Score	Keeping the Team on Track
5	• Watches conditions affecting the team and monitors the team's progress. • Makes sure that teammates are making appropriate progress. • Gives teammates specific, timely, and constructive feedback.
4	Demonstrates behaviors described immediately above and below.
3	• Notices changes that influence the team's success. • Knows what everyone on the team should be doing and notices problems.

	• Alerts teammates or suggests solutions when the team's success is threatened.
2	Demonstrates behaviors described immediately above and below.
1	• Is unaware of whether the team is meeting its goals. • Does not pay attention to teammates' progress. • Avoids discussing team problems, even when they are obvious.

A.5. Expecting Quality dimension (BARS scale)

Score	Expecting Quality
5	• Motivates the team to do excellent work. • Cares that the team does outstanding work, even if there is no additional reward. • Believes that the team can do excellent work.
4	Demonstrates behaviors described immediately above and below.
3	• Encourages the team to do good work that meets all requirements. • Wants the team to perform well enough to earn all available rewards. • Believes that the team can fully meet its responsibilities.
2	Demonstrates behaviors described immediately above and below.
1	• Satisfied even if the team does not meet assigned standards. • Wants the team to avoid work, even if it hurts the team. • Doubts that the team can meet its requirements.

A.6. Having Related Knowledge, Skills, and Abilities dimension (BARS scale)

Score	Having Related Knowledge, Skills, and Abilities
5	• Demonstrates the knowledge, skills, and abilities to do excellent work. • Acquires new knowledge or skills to improve the team's performance. • Able to perform the role of any team member if necessary.
4	Demonstrates behaviors described immediately above and below.
3	• Demonstrates sufficient knowledge, skills, and abilities to contribute to the team's work. • Acquires knowledge or skills as needed to meet requirements. • Able to perform some of the tasks normally done by other team members.
2	Demonstrates behaviors described immediately above and below.
1	• Missing basic qualifications needed to be a member of the team. • Unable or unwilling to develop knowledge or skills to contribute to the team. • Unable to perform any of the duties of other team members.

REFERENCES

ABET. (2019). *2018-19 Criteria for Accrediting Engineering Programs*. Baltimore, MD: ABET.

Abrami, P. C., Chambers, B., Poulsen, C., Kouros, C., Farrell, M., & D'Apollonia, S. (1994). Positive social interdependence and classroom climate. *Genetic, Social, and General Psychology Monographs*, *120*(3), 329—346. Retrieved from http://europepmc.org/abstract/MED/7926696

Adams, S. G. (2001). The effectiveness of the E-Team approach to invention and innovation. *Journal of Engineering Education*, *90*(4), 597–600. https://doi.org/10.1002/j.2168-9830.2001.tb00645.x

Agarwal, P., & Farndale, E. (2017). High-performance work systems and creativity implementation: The role of psychological capital and psychological safety. *Human Resource Management Journal*, *27*(3), 440–458. https://doi.org/10.1111/1748-8583.12148

Alba-Flores, R., & Rios, F. (2019). Incorporating peer review techniques to enhance students' communication skills and team performance in engineering capstone projects. *International Journal of Engineering Education*, *35*(6(B)), 1969–1982.

Amason, A. C. (1996). Distinguishing the effects of functional and dysfunctional conflict on strategic decision making: Resolving a paradox for top management teams. *Management*.

Apesteguia, J., Azmat, G., & Iriberri, N. (2012). The impact of gender composition on team performance and decision making: Evidence from the rield. *Management Science*, *58*(1), 78–93. https://doi.org/10.1287/mnsc.1110.1348

Asio, S. M., Cross, J. A., & Ekwaro-Osire, S. (2018). Factors affecting innovation in engineering design teams: An empirical investigation of student team perceptions. *International Journal of Engineering Education*, *34*(4), 1159–1173. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-85049807995&partnerID=40&md5=a0450ca1c7e3cd87bb9e0091671607d6

Atwater, L. E., & Yammarino, F. J. (1997). Self–other rating agreement. In G. R. Ferris (Ed.), *Research in personnel and human resources management* (p. 121−174). JAI Press.

Atwater, L. E., & Yammrino, F. J. (1992). Does self-other agreement on leadership perceptions moderate the validity of leadership and performance predictions? *Personnel Psychology*, *45*(1), 141–164. https://doi.org/10.1111/j.1744-6570.1992.tb00848.x

Bacon, D. R., Stewart, K. A., & Silver, W. S. (1999). Lessons from the best and worst student team experiences: How a teacher can make the difference. *Journal of Management Education*, *23*(5), 467–488. https://doi.org/10.1177/105256299902300503

Bailey, R. R., Swan, A. K., Coso, A., Creager, M. F., & Rowan-Kenyon, H. T. (2012). The role of gender in student perceptions of leadership on interdisciplinary engineering teams. *Journal of Women and Minorities in Science and Engineering*, *18*(2), 97–113. https://doi.org/10.1615/jwomenminorscieneng.2012003031

Bannerot, R. (2003). Who are the good team players? In *Proceedings of 2003 Annual Conference of ASEE Gulf Southwest Section*.

Barczak, G., & Wilemon, D. (2001). Factors influencing product development team satisfaction. *European Journal of Innovation Management*, *4*(1), 32–36. https://doi.org/10.1108/14601060110365556

Baughman, J., Hassall, L., & Xu, X. (2019). Comparison of student team dynamics between nonflipped and flipped versions of a large-enrollment sophomore design engineering course. *Journal of Engineering Education*, *108*(1), 103–118. https://doi.org/10.1002/jee.20251

Beaman, R. (1998). The unquiet... even loud, andragogy! alternative assessments for adult learners. *Innovative Higher Education*, *23*(1), 47–59. https://doi.org/10.1023/A:1022920513341

Beddoes, K., & Panther, G. (2018). Gender and teamwork: an analysis of professors' perspectives and practices. *European Journal of Engineering Education*, *43*(3), 330–343.

Beigpourian, B., Ferguson, D. M., Berry, F. C., Ohland, M. W., & Wei, S. (2019). Using CATME to document and improve the effectiveness of teamwork in capstone courses. 2019 ASEE Annual Conference & Exposition. Retrieved from https://www.asee.org/public/conferences/140/papers/25751/view

Beigpourian, B., Luchini, F., Ohland, M. W., & Ferguson, D. M. (2019). Psychological safety as an effective measurement in engineering classrooms. In *2019 ASEE Annual Conference & Exposition*. Retrieved from https://www.asee.org/public/conferences/140/papers/25418/view

Beigpourian, B., & Ohland, M. W. (2019). A systematized review: Gender and race in teamwork in undergraduate engineering classrooms. 2019 ASEE Annual Conference & Exposition. Retrieved from https://www.asee.org/public/conferences/140/papers/25494/view

Berenson, S. B., Slaten, K. M., Williams, L., & Ho, C.-W. (2005). Voices of women in a software engineering course: Reflections on collaboration. *Journal on Educational Resources in Computing*, *4*(1), 1–18. https://doi.org/10.1145/1060071.1060074

Bergmann, B., & Schaeppi, J. (2016). *A data-driven approach to group creativity*. Harvard Business Review, 12th July 2016.

Bienefeld, N., & Grote, G. (2014). Speaking up in ad hoc multiteam systems: Individual-level effects of psychological safety, status, and leadership within and across teams. *European Journal of Work and Organizational Psychology*, *23*(6), 930–945. https://doi.org/10.1080/1359432X.2013.808398

Biswas, S., & Bhatnagar, J. (2013). Mediator analysis of employee engagement: Role of perceived organizational support, P-O fit, organizational commitment and job satisfaction. *Vikalpa: The Journal for Decision Makers*, *38*(1), 27–40. https://doi.org/10.1177/0256090920130103

Bliese, P. D. (2000). Within group agreement, non-independence and reliability: Implications for data and analysis. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 349–381). San Francisco, CA: Jossey-Bass.

Borrego, M., Foster, M. J., & Froyd, J. E. (2014). Systematic literature reviews in engineering education and other developing interdisciplinary fields. *Journal of Engineering Education*, *103*(1), 45–76. https://doi.org/10.1002/jee.20038

Borrego, M., Karlin, J., McNair, L. D., & Beddoes, K. (2013). Team effectiveness theory from industrial and organizational psychology applied to engineering student project teams: A research review. *Journal of Engineering Education*, *102*(4), 472–512. https://doi.org/10.1002/jee.20023

Brutus, S., Fleenor, J. W., & McCauley, C. D. (1999). Demographic and personality predictors of congruence in multi-source ratings. *Journal of Management Development*, *18*(5), 417–435. https://doi.org/10.1108/02621719910273569

Carmeli, A., & Gittell, J. H. (2009). High-quality relationships, psychological safety, and learning from failures in work organizations. *Journal of Organizational Behavior*, *30*(6), 709–729. https://doi.org/10.1002/job.565

Chan, D. (1998). Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology*, *83*(2), 234–246. https://doi.org/10.1037/0021-9010.83.2.234

Cinar, Y., & Biglin, A. (2011). Peer assessment for undergraduate teamwork projects in petroleum engineering. *International Journal of Engineering Education*, *27*(2), 310–322.

Cinar, Y., & Bilgin, A. (2009). A study on peer assessment of teamwork projects in undergraduate petroleum engineering education. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers. https://doi.org/10.2118/123166-MS

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge. https://doi.org/10.4324/9780203771587

De Dreu, C. K. (2006). When too little or too much hurts: Evidence for a curvilinear relationship between task conflict and innovation in teams. *Journal of Management*, *32*(1), 83–107. https://doi.org/10.1177/0149206305277795

Deep, S., Othman, H., & Salleh, B. M. (2016). Potential causes and outcomes of communication conflicts at the workplace -A qualitative study in pakistan. *Journal of Management Info*, *3*(3), 1–5. https://doi.org/10.31580/jmi.v11i1.54

Derickson, R., Fishman, J., Osatuke, K., Teclaw, R., & Ramsel, D. (2015). Psychological safety and error reporting within veterans health administration hospitals. *Journal of Patient Safety*, *11*(1), 60–66. https://doi.org/10.1097/PTS.0000000000000082

Detert, J. R., & Burris, E. R. (2007). Leadership behavior and employee voice: Is the door really open? *Academy of Management Journal*, *50*(4), 869–884. https://doi.org/10.5465/amj.2007.26279183

DiStefano, J. J., & Maznevski, M. L. (2000). Creating value with diverse teams in global management. *Organizational Dynamics*, *29*(1), 45–63. https://doi.org/10.1016/S0090-2616(00)00012-7

Downing, C. G. (2001). Essential non-technical skills for teaming. *Journal of Engineering Education*, *90*(1), 113–117. https://doi.org/10.1002/j.2168-9830.2001.tb00577.x

Duhigg, C. (2016). What google learned from its quest to build the perfect team. *The New York Times Magazine*.

Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale , NJ: Erlbaum.

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, *44*(2), 350–383. https://doi.org/10.2307/2666999

Edmondson, A. C. (2004). Psychological safety, trust, and learning in organizations: A group-level lens. In R. M. Kramer & K. S. Cook (Eds.), *The Russell Sage Foundation series on trust. Trust and distrust in organizations: Dilemmas and approaches* (pp. 239–272). New York, NY: Russell Sage Foundation.

Edmondson, A C. (2004). Learning from failure in health care: Frequent opportunities, pervasive barriers. *Quality and Safety in Health Care*, *13*(suppl_2), ii3–ii9. https://doi.org/10.1136/qshc.2003.009597

Edmondson, Amy C, & Roloff, K. S. (2009). Overcoming barriers to collaboration: Psychological safety and learning in diverse teams. In *Team effectiveness in complex organizations: Cross-disciplinary perspectives and approaches.* (pp. 183–208). New York, NY, US: Routledge/Taylor & Francis Group.

Falchikov, N. (2004). Involving students in assessment. *Psychology Learning & Teaching*, *3*(2), 102–108. https://doi.org/10.2304/plat.2003.3.2.102

Faulkner, W. (2007). `Nuts and bolts and people': Gender-troubled engineering identities. *Social Studies of Science*, *37*(3), 331–356. https://doi.org/10.1177/0306312706072175

Felder, R.M., Felder, G. N., & Dietz, E. J. (2002). The effects of personality type on engineering student performance and attitudes. *Journal of Engineering Education*, *91*(1), 3–17. https://doi.org/10.1002/j.2168-9830.2002.tb00667.x

Felder, Richard M., Felder, G. N., Mauney, M., Hamrin, C. E., & Dietz, E. J. (1995). A longitudinal study of engineering student performance and retention. III. Gender differences in student performance and attitudes. *Journal of Engineering Education*, *84*(2), 151–163. https://doi.org/10.1002/j.2168-9830.1995.tb00162.x

Felder, Richard M, & Brent, R. (2001). Effective strategies for cooperative learning. *Journal of Cooperation & Collaboration in College Teaching*, *10*(2), 69–75.

Felder, Richard M, & Brent, R. (2007). Cooperative learning. In *Active learning* (Vol. 970, pp. 4–34). American Chemical Society. https://doi.org/doi:10.1021/bk-2007-0970.ch004

Fiegel, G. L., & Denatale, J. S. (2011). Civil engineering capstone design: Team formation, preparation, and performance. *International Journal of Engineering Education*, *27*(6), 1295–1307.

Fila, N. D., & Purzer, S. (2014). The relationship between team gender diversity, idea variety, and potential for design innovation. *International Journal of Engineering Education*, *30*(6A), 1405–1418. Retrieved from http://www.ijee.ie/

Fleenor, J. W., Smither, J. W., Atwater, L. E., Braddy, P. W., & Sturm, R. E. (2010). Self–other rating agreement in leadership: A review. *The Leadership Quarterly*, *21*(6), 1005–1034. https://doi.org/10.1016/j.leaqua.2010.10.006

Florman, S. C. (1997). Non-technical studies for engineers: The challenge of relevance. *European Journal of Engineering Education*, *22*(3), 249–258. https://doi.org/10.1080/03043799708923457

Flynn, E. A., Savage, G., Penti, M., Brown, C., & Watke, S. (1991). Gender and modes of collaboration in a chemical engineering design course. *Journal of Business and Technical Communication*, *5*(4), 444–462. https://doi.org/10.1177/1050651991005004006

Gibson, C. B., & Gibbs, J. L. (2006). Unpacking the concept of virtuality: The effects of geographic dispersion, electronic dependence, dynamic structure, and national diversity on team innovation. *Administrative Science Quarterly*, *51*(3), 451–495. https://doi.org/10.2189/asqu.51.3.451

Godwin, A., Kirn, A., & Rohde, J. (2017). Awareness without action: Student attitudes after engineering teaming experiences. *International Journal of Engineering Education*, *33*(6a), 1878–1891.

Grulke, E. A., Beert, D. C., & Lane, D. R. (2001). The effects of physical environment on engineering team performance: A case study. *Journal of Engineering Education*, *90*(3), 319–330. https://doi.org/10.1002/j.2168-9830.2001.tb00611.x

Guzzo, R. a., & Shea, G. P. (1992). Group performance and intergroup relations in organizations. In M. Dunnette & L. Hough (Eds.), *Handbook of industrial and organizational psychology, Vol. 3* (pp. 269–313). Palo Alto, CA: Consulting Psychologist Press.

Hackman, J. R. (1987). The design of work teams. In J. W. Lorsch (Ed.), *Handbook of organizational behavior* (pp. 315–342). Englewood Cliffs, NJ: Prentice Hall.

Hansen, R. S. (2006). Benefits and problems with student teams: Suggestions for improving team projects. *Journal of Education for Business*, *86*(1), 11–19. https://doi.org/10.3200/JOEB.82.1.11-19

Harrison, D. A., & Klein, K. J. (2007). What’s the difference? diversity constructs as separation, variety, or disparity in organizations. *Academy of Management Review*, *32*(4), 1199–1228. https://doi.org/10.5465/amr.2007.26586096

Hartman, H, & Hartman, M. (2006). Attitudes to group work: Gendered differences. In *American Society for Engineering Education*.

Hartman, Harriet, & Hartman, M. (2009). Do gender differences in undergraduate engineering orientations persist when major is controlled? *International Journal of Gender, Science and Technology*, *1*(1), 62–82.

Harvey, J.-F., Johnson, K. J., Roloff, K. S., & Edmondson, A. C. (2019). From orientation to behavior: The interplay between learning orientation, open-mindedness, and psychological safety in team learning. *Human Relations*, *72*(11), 1726–1751. https://doi.org/10.1177/0018726718817812

Hirsch, P. L., & Mckenna, A. F. (2008). Using reflection to promote teamwork understanding in engineering design education. *International Journal of Engineering Education*, *24*(2), 377–385.

Hirshfield, L., & Fowler, R. (2020). Developing and piloting a survey to assess dissatisfaction of women in student teams. In *ASEE Virtual Annual Conference Experience*.

Hobfoll, S. E. (1989). Conservation of resources: A new attempt at conceptualizing stress. *American Psychologist*, *44*(3), 513–524. https://doi.org/10.1037/0003-066X.44.3.513

Humphrey, S. E., Aime, F., Cushenbery, L., Hill, A. D., & Fairchild, J. (2017). Team conflict dynamics: Implications of a dyadic view of conflict for team performance. *Organizational Behavior and Human Decision Processes*, *142*, 58–70. https://doi.org/10.1016/j.obhdp.2017.08.002

Jakobsen, J. C., Gluud, C., Wetterslev, J., & Winkel, P. (2017). When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with flowcharts. *BMC Medical Research Methodology*, *17*(1), 162. https://doi.org/10.1186/s12874-017-0442-1

James, L. R. (1982). Aggregation bias in estimates of perceptual agreement. *Journal of Applied Psychology*, *67*(2), 219–229. https://doi.org/10.1037/0021-9010.67.2.219

James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology*, *69*(1), 85–98. https://doi.org/10.1037/0021-9010.69.1.85

Jaschik, S. (2005). What Larry Summers said. Retrieved from https://www.insidehighered.com/news/2005/02/18/what-larry-summers-said

Jassawalla, A., Sashittal, H., & Sashittal, A. (2009). Students' perceptions of social loafing: Its antecedents and consequences in undergraduate business classroom teams. *Academy of Management Learning & Education*, *8*(1), 42–54. https://doi.org/10.5465/amle.2009.37012178

Jehn, K. A. (1994). Enahancing effectiveness: An investigation of advantages and disadvantages of value based intragroup conflict, *5*(3), 223–238.

Jehn, K. A. (1995). A multimethod examination of the benefits and detriments of intragroup conflict, *40*(2), 256–282.

Jehn, K. A., & Bendersky, C. (2003). Intragroup conflict in organizations: A contingency perspective on the conflict-outcome relationship. *Research in Organizational Behavior*. https://doi.org/10.1016/S0191-3085(03)25005-X

Jehn, K. A., Greer, L., Levine, S., & Szulanski, G. (2008). The effects of conflict types, dimensions, and emergent states on group outcomes. *Group Decision and Negotiation*, *17*(6), 465–495. https://doi.org/10.1007/s10726-008-9107-0

Jehn, K. A., & Mannix, E. A. (2001). The dynamic nature of conflict: A longitudinal study of intragroup conflict and group performance. *Academy of Management Journal*, *44*(2), 238–251. https://doi.org/10.2307/3069453

Jimenez-Useche, I. C., Ohland, M. W., & Hoffmann, S. R. (2015). Multicultural dynamics in first-year engineering teams in the US. In *2015 ASEE Annual Conference and Exposition Proceedings*.

Johnson, D. R., & Creech, J. C. (1983). Ordinal measures in multiple indicator models: A simulation study of categorization error. *American Sociological Review*, *48*(3), 398. https://doi.org/10.2307/2095231

Johnson, D. W., Johnson, R. T., & Smith, K. (2007). The state of cooperative learning in postsecondary and professional settings. *Educational Psychology Review*, *19*(1), 15–29. https://doi.org/10.1007/s10648-006-9038-8

Joshi, A. (2014). By whom and when is women’s expertise recognized? The interactive effects of gender and education in science and engineering teams. *Administrative Science Quarterly*, *59*(2), 202–239. https://doi.org/10.1177/0001839214528331

Kahn, W. A. (1990). Psychological conditions of personal engagement and disengagement at work. *Academy of Management Journal*, *33*(4), 692–724. https://doi.org/10.5465/256287

Kaufman, D. B., Felder, R. M., & Fuller, H. (2000). Accounting for individual effort in cooperative learning teams. *Journal of Engineering Education*, *89*(2), 133–140. https://doi.org/10.1002/j.2168-9830.2000.tb00507.x

Kilic, D. (2016). An examination of using self-, peer-, and teacher-assessment in higher education: A case study in teacher education. *Higher Education Studies*, *6*(1), 136. https://doi.org/10.5539/hes.v6n1p136

Kirk-Brown, A., & Van Dijk, P. (2016). An examination of the role of psychological safety in the relationship between job resources, affective commitment and turnover intentions of Australian employees with chronic illness. *The International Journal of Human Resource Management*, *27*(14), 1626–1641. https://doi.org/10.1080/09585192.2015.1053964

Klein, Katherine J., Conn, A. B., Smith, D. B., & Sorra, J. S. (2001). Is everyone in agreement? An exploration of within-group agreement in employee perceptions of the work environment. *Journal of Applied Psychology*, *86*(1), 3–16. https://doi.org/10.1037/0021-9010.86.1.3

Kowalski, G. J., & Smyser, B. M. (2017). Success factors for international students in capstone design teams. *International Journal of Engineering Education*, *33*(5), 1432–1441.

Krasikova, D. V., & LeBreton, J. M. (2012). Just the two of us: Misalignment of theory and methods in examining dyadic phenomena. *Journal of Applied Psychology*, *97*(4), 739–757. https://doi.org/10.1037/a0027962

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one’s own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

Laeser, M., Moskal, B. M., Knecht, R., & Lasich, D. (2003). Engineering design: Examining the impact of gender and the team’s gender composition. *Journal of Engineering Education*, *92*(1), 49–56. https://doi.org/10.1002/j.2168-9830.2003.tb00737.x

Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria. *Organizational Research Methods*, *9*(2), 202–220. https://doi.org/10.1177/1094428105284919

Lang, J. D., Cruse, S., McVey, F. D., & McMasters, J. (1999). Industry expectations of new engineers: A survey to assist curriculum designers. *Journal of Engineering Education, 88*(1), 43–51. https://doi.org/10.1002/j.2168-9830.1999.tb00410.x

Lau, D. C., & Murnighan, J. K. (2005). Interactions within groups and subgroups: The effects of demographic faultlines. *The Academy of Management Journal*, *48*(4), 645–659. https://doi.org/10.2307/20159684

Layton, R. A., Loughry, M. L., Ohland, M. W., & Ricco, G. D. (2010). Design and validation of a web-based system for assigning members to teams using instructor-specified criteria. *Advances in Engineering Education*, *2*(1), 1–28.

Leaper, C., & Ayres, M. M. (2007). A meta-analytic review of gender variations in adults' language use: Talkativeness, affiliative speech, and assertive speech. *Personality and Social Psychology Review*, *11*(4), 328–363. https://doi.org/10.1177/1088868307302221

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods*, *11*(4), 815–852. https://doi.org/10.1177/1094428106296642

Lee, H. W., Choi, J. N., & Kim, S. (2018). Does gender diversity help teams constructively manage status conflict? An evolutionary perspective of status conflict, team psychological safety, and team creativity. *Organizational Behavior and Human Decision Processes*, *144*, 187–199. https://doi.org/10.1016/j.obhdp.2017.09.005

Lejk, M., & Wyvill, M. (2001). The effect of the inclusion of selfassessment with peer assessment of contributions to a group project: A quantitative study of secret and agreed assessments. *Assessment & Evaluation in Higher Education*, *26*(6), 551–561. https://doi.org/10.1080/02602930120093887

Leonardi, P. M., Jackson, M. H., & Diwan, A. (2009). The enactment-externalization dialectic: Rationalization and the persistence of counterproductive technology design practices in student engineering. *Academy of Management Journal*, *52*(2), 400–420.

Leroy, H., Dierynck, B., Anseel, F., Simons, T., Halbesleben, J. R. B., McCaughey, D., … Sels, L. (2012). Behavioral integrity for safety, priority of safety, psychological safety, and patient safety: A team-level study. *Journal of Applied Psychology, 97*(6), 1273–1281. https://doi.org/10.1037/a0030076

Lewis, P., Aldridge, D., & Swamidass, P. M. (1998). Assessing teaming skills acquisition on undergraduate project teams. *Journal of Engineering Education*, *87*(2), 149–155. https://doi.org/10.1002/j.2168-9830.1998.tb00335.x

Liang, J., Farh, C. I. C., & Farh, J.-L. (2012). Psychological antecedents of promotive and prohibitive voice: A two-wave examination. *Academy of Management Journal*, *55*(1), 71–92. https://doi.org/10.5465/amj.2010.0176

Liu, E. Z.-F., Lin, S. S. J., Chiu, C.-H., & Yuan, S.-M. (2001). Web-based peer review: the learner as both adapter and reviewer. *IEEE Transactions on Education*, *44*(3), 246–251. https://doi.org/10.1109/13.940995

Liu, S., Hu, J., Li, Y., Wang, Z., & Lin, X. (2014). Examining the cross-level relationship between shared leadership and learning in teams: Evidence from China. *The Leadership Quarterly*, *25*(2), 282–295. https://doi.org/10.1016/j.leaqua.2013.08.006

Lloyd, N. A., & Szymakowski, J. (2017). Student experiences in first-year engineering classrooms – exploring issues of gender in an Australian programme. *Australasian Journal of Engineering Education*, *22*(1), 23–31. https://doi.org/10.1080/22054952.2017.1365328

Loignon, A. C., Woehr, D. J., Thomas, J. S., Loughry, M. L., Ohland, M. W., & Ferguson, D. M. (2017). Facilitating peer evaluation in team contexts: The impact of frame-of-reference rater training. *Academy of Management Learning and Education*, *16*(4), 562–578. https://doi.org/10.5465/amle.2016.0163

Loughry, M. L., & Tosi, H. L. (2008). Performance implications of peer monitoring. *Organization Science*, *19*(6), 876–890. Retrieved from http://www.jstor.org/stable/25146226

Luechtefeld, R., Baca, D., & Watkins, S. E. (2008). Training for self-managed student teams. *International Journal of Engineering Education*, *24*(6), 1139–1147.

Luthar, H. K. (1996). Gender differences in evaluation of performance and leadership ability: Autocratic vs. democratic managers. *Sex Roles*, *35*(5–6), 337–361. https://doi.org/10.1007/BF01664773

Madjar, N., & Ortiz-Walters, R. (2009). Trust in supervisors and trust in customers: Their independent, relative, and joint effects on employee performance and creativity. *Human Performance*, *22*(2), 128–142. https://doi.org/10.1080/08959280902743501

Marcangelo, C., Cartney, P., & Barnes, C. (2010). The opportunities and challenges of self, peer and group assessment. In M. Hammick & C. Reid (Eds.), *Contemporary issues in assessment in health sciences and practice education* (pp. 1–20). London, UK: Higher Education Academy. Retrieved from https://eprints.mdx.ac.uk/id/eprint/5490

McCleary, L. (2002). Using multiple imputation for analysis of incomplete aata in clinical research. *Nursing Research*, *51*(5), 339–343. https://doi.org/10.1097/00006199-200209000-00012

McGrath, J. E. (1964). *Social psychology : A brief introduction*. *The Sociological Quarterly*. New York, NY: Rinehart and Winston.

Mead, P. F., Moore, D., Natishan, M., Schmidt, L., Brown, S. V., Lathan, C., … Mouring, S. (1999). Faculty and student views on engineering student team effectiveness. *Journal of Women and Minorities in Science and Engineering*, *5*(4).

Meadows, L. A., & Sekaquaptewa, D. (2013). The influence of gender stereotypes on role adoption in student teams. In *ASEE Annual Conference and Exposition, Conference Proceedings*.

Meier, R. L., Williams, M. R., & Humphreys, M. A. (2000). Refocusing our efforts: Assessing non-technical competency gaps. *Journal of Engineering Education*, *89*(3), 377–385. https://doi.org/10.1002/j.2168-9830.2000.tb00539.x

Menekse, M., Higashi, R., Schunn, C. D., & Baehr, E. (2017). The role of robotics teams' collaboration quality on team performance in a robotics tournament. *Journal of Engineering Education*, *106*(4), 564–584. https://doi.org/10.1002/jee.20178

Monaghan, C., Bizumic, B., Reynolds, K., Smithson, M., Johns-Boast, L., & van Rooy, D. (2015). Performance of student software development teams: The influence of personality and identifying as team members. *European Journal of Engineering Education*, *40*(1), 52–67. https://doi.org/10.1080/03043797.2014.914156

Moshavi, D., Brown, F. W., & Dodd, N. G. (2003). Leader self-awareness and its relationship to subordinate attitudes and performance. *Leadership & Organization Development Journal*, *24*(7), 407–418. https://doi.org/10.1108/01437730310498622

Myaskovsky, L., Unikel, E., & Dew, M. A. (2005). Effects of gender diversity on performance and interpersonal behavior in small work groups. *Sex Roles*, *52*(9–10), 645–657. https://doi.org/10.1007/s11199-005-3732-8

Natishan, M. E., Schmidt, L. C., & Mead, P. (2000). Student focus group results on student team performance issues. *Journal of Engineering Education*, *89*(3), 269–272. https://doi.org/10.1002/j.2168-9830.2000.tb00524.x

Nepal, K. P. (2016). Simplified framework for managing team learning in engineering subjects. *International Journal of Engineering Education*, *32*(3(A)), 1182–1193.

Nepal, K. P. (2018). A relative relevance approach to refine inconsistent peer- and self-assessment scores in teamwork assessment. *International Journal of Engineering Education*, *34*(4), 1289–1298.

Neumeyer, X., & Mckenna, A. F. (2011). Assessing Team Conflict in Student Design Teams. In ASME 2011 International Mechanical Engineering Congress and Exposition. https://doi.org/10.1115/IMECE2011-63933

Neumeyer, X., & Santos, S. C. (2020). The effect of team conflict on teamwork performance: An engineering education perspective. *International Journal of Engineering Education*, *36*(1(B)), 502–509.

Newman, A., Donohue, R., & Eva, N. (2017). Psychological safety: A systematic review of the literature. *Human Resource Management Review*, *27*(3), 521–535. https://doi.org/10.1016/j.hrmr.2017.01.001

Ngambeki, I., Habashi, M., Evangelou, D., Graziano, W., Sakka, D., & Corapci, F. (2012). Using profiles of Person-Thing Orientation to examine the underpresentation of women in engineering in thee cultural contexts. *International Journal of Engineering Education*, *28*(3), 621–632.

Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education*, *15*(5), 625–632. https://doi.org/10.1007/s10459-010-9222-y

Novoselich, B. J., & Knight, D. B. (2019). Relating shared leadership to capstone team effectiveness. *International Journal of Engineering Education*, *35*(6(B)), 1888–1906.

Oakley, B., Felder, R. M., Brent, R., & Elhajj, I. (2004). Turning student groups into effective teams. *Journal of Student Centered Learning*, *2*(1), 9–34.

Ohland, M. W., Brawner, C. E., Camacho, M. M., Layton, R. A., Long, R. A., Lord, S. M., & Wasburn, M. H. (2011). Race, gender, and measures of success in engineering education. *Journal of Engineering Education*, *100*(2), 225–252. https://doi.org/10.1002/j.2168-9830.2011.tb00012.x

Ohland, M. W., Loughry, M. L., Woehr, D. J., Bullard, L. G., Felder, R. M., Finelli, C. J., … Schmucker, D. G. (2012). The comprehensive assessment of team member effectiveness: Development of a behaviorally anchored rating scale for self- and peer evaluation. *Academy of Management Learning and Education*, *11*(4), 609–630. https://doi.org/10.5465/amle.2010.0177

Ostafichuk, P. M., Mattucci, S., D'Entremont, A., Shirzad, N., & Naylor, C. (2015). Self-perception differences based on gender and personality type in team projects. *Proceedings of the Canadian Engineering Education Association*. https://doi.org/10.24908/pceea.v0i0.5784

Ostafichuk, P. M., & Sibley, J. (2019). Self-bias and gender-bias in student peer evaluation:An expanded study. *Proceedings of the Canadian Engineering Education Association (CEEA)*. https://doi.org/10.24908/pceea.vi0.13864

Palanski, M. E., & Vogelgesang, G. R. (2011). Virtuous creativity: The effects of leader behavioural integrity on follower creative thinking and risk taking. *Canadian Journal of Administrative Sciences / Revue Canadienne Des Sciences de l'Administration*, *28*(3), 259–269. https://doi.org/10.1002/cjas.219

Parker, R., Sangelkar, S., Swenson, M., & Ford, J. D. (2019). Launching for success: A review of team formation for capstone design. *International Journal of Engineering Education*, *35*(6(B)), 1926–1936.

Passow, H. J. (2012). Which ABET competencies do engineering graduates find most important in their work? *Journal of Engineering Education*, *101*(1), 95–118. https://doi.org/10.1002/j.2168-9830.2012.tb00043.x

Passow, H. J., & Passow, C. H. (2017). What competencies should undergraduate engineering programs emphasize? A systematic review. *Journal of Engineering Education*, *106*(3), 475–526. https://doi.org/10.1002/jee.20171

Pawley, A., Dickerson, D., Ohland, M., & Zywicki, S. (2019). Identifying marginalization and allying tendencies to transform engineering relationships. *National Science Foundation Award 1936778*. Retrieved from https://www.nsf.gov/awardsearch/showAward?AWD_ID=1936778

Pawley, A. L., Schimpf, C., & Nelson, L. (2016). Gender in engineering education research: A content analysis of research in JEE, 1998–2012. *Journal of Engineering Education*, *105*(3), 508–528. https://doi.org/10.1002/jee.20128

Peeters, M. A. G., Rutte, C. G., Van Tuijl, H. F. J. M., & Reymen, I. M. M. J. (2006). The big five personality traits and individual satisfaction with the team. *Small Group Research*, *37*(2), 187–211. https://doi.org/10.1177/1046496405285458

Pelled, L. H., Eisenhardt, K. M., & Xin, K. R. (1999). Exploring the black box : An analysis of work group diversity , conflict , and performance. *Administrative Science Quarterly*, *44*(1), 1–28.

Rathert, C., Ishqaidef, G., & May, D. R. (2009). Improving work environments in health care. *Health Care Management Review*, *34*(4), 334–343. https://doi.org/10.1097/HMR.0b013e3181abce2b

Reason, J. T. (2000). Human error: Models and management. *British Medical Journal*, *320*, 768–770.

Reid, K., Reeping, D., & Spingola, E. (2018). A taxonomy for introduction to engineering courses. *The International Journal of Engineering Education*, *34*(1), 2–19.

Rosch, D., Collier, D., & Zehr, S. (2018). Self-vs.-teammate assessment of leadership competence: The effects of gender, leadership self-efficacy, and motivation to lead. *Journal of Leadership Education*, *13*(2), 96–124. https://doi.org/10.12806/v13/i2/r5

Rosch, D. M., & Collier, D. (2014). Incoming leadership-oriented differences between students in a leadership studies course and a team-based project course. *Journal of Leadership Education*, *12*(2), 103–121. https://doi.org/10.12806/v12/i2/103

Rosser, S. V. (1990). *Female-friendly science: Applying women's studies methods and theories to attract students*. New York, NY: Pergamon Press.

Rosser, S. V. (1998). Group work in science, engineering, and mathematics: Consequences of ignoring gender and race. *College Teaching*, *46*(3), 82–88. https://doi.org/10.1080/87567559809596243

Roussin, C. J., Larraz, E., Jamieson, K., & Maestre, J. M. (2018). Psychological safety, self-efficacy, and speaking up in interprofessional health care simulation. *Clinical Simulation in Nursing*, *17*, 38–46. https://doi.org/10.1016/j.ecns.2017.12.002

Rust, C., Price, M., & O'Donovan, B. (2003). Improving students' learning by developing their understanding of assessment criteria and processes. *Assessment & Evaluation in Higher Education*, *28*(2), 147–164. https://doi.org/10.1080/02602930301671

Sande, J. C. G. L. DE, & Godino-Llorente, J. I. (2014). Peer assessment and self-assessment: Effective learning tools in higher education. *International Journal of Engineering Education*, *30*(3), 711–721.

Schaubroeck, J., Lam, S. S. K., & Peng, A. C. (2011). Cognition-based and affect-based trust as mediators of leader behavior influences on team performance. *Journal of Applied Psychology*, *96*(4), 863–871. https://doi.org/10.1037/a0022625

Scott, G., & Yates, K. W. (2002). Using successful graduates to improve the quality of undergraduate engineering programmes. *European Journal of Engineering Education*, *27*(4), 363–378. https://doi.org/10.1080/03043790210166666

Shimazoe, J., & Aldrich, H. (2010). Group work can be gratifying: Understanding & overcoming resistance to cooperative learning. *College Teaching*, *58*(2), 52–57. https://doi.org/10.1080/87567550903418594

Siemsen, E., Roth, A. V., Balasubramanian, S., & Anand, G. (2009). The influence of psychological safety and confidence in knowledge on employee knowledge sharing. *Manufacturing & Service Operations Management*, *11*(3), 429–447. https://doi.org/10.1287/msom.1080.0233

Singh, B., & Winkel, D. E. (2012). Racial differences in helping behaviors: The role of respect, safety, and identification. *Journal of Business Ethics*, *106*(4), 467–477. https://doi.org/10.1007/s10551-011-1011-x

Singh, B., Winkel, D. E., & Selvarajan, T. T. (2013). Managing diversity at work: Does psychological safety hold the key to racial differences in employee performance? *Journal of Occupational and Organizational Psychology*, *86*(2), 242–263. https://doi.org/10.1111/joop.12015

Steiner, M., & Kanai, J. (2016). Creating effective multidisciplinary capstone project teams. *International Journal of Engineering Education*, *32*(2(A)), 625–639.

Sullivan, G. M., & Artino, A. R. (2013). Analyzing and interpreting data from Likert-type scales. *Journal of Graduate Medical Education*, *5*(4), 541–542. https://doi.org/10.4300/JGME-5-4-18

Takai, S., & Esterman, M. (2019). A review of team effectiveness models and possible instruments for measuring design-team inputs, processes, and outputs. *International Journal of Engineering Education*, *35*(6(A)), 1684–1697.

Tannenbaum, S. I., Beard, R. L., & Salas, E. (1992). Team building and its influence on team effectiveness: An examination of conceptual and empirical developments. In K. Kelley (Ed.), *Advances in psychology* (pp. 117–153). Amsterdam, The Netherlands: North-Holland.

Tax, S. S., & Brown, S. W. (1998). Recovering and learning fromservice failure. *Sloan Management Review*, *40*(1), 75–88.

Tenenberg, J. (2019). Factors affecting free riding on teams: Implications for engineering education. *International Journal of Engineering Education*, *35*(6(A)), 1703–1724.

Tonso, K. L. (1996). The impact of cultural norms on women. *Journal of Engineering Education*, *85*(3), 217–225. https://doi.org/10.1002/j.2168-9830.1996.tb00236.x

Tonso, K. L. (2006). Teams that work: Campus culture, engineer identity, and social interactions. *Journal of Engineering Education*, *95*(1), 25–37. https://doi.org/10.1002/j.2168-9830.2006.tb00875.x

Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, *68*(3), 249. https://doi.org/10.2307/1170598

Trevelyan, J. (2014). *The making of an expert engineer*. Boca Raton, FL: CRC Press.

Tucker, R., Abbasi, N., Thorpe, G., Ostwald, M., Williams, A., & Wallis, L. (2014). *Enhancing and assessing group and team learning in architecture and related design contexts*. Sydney, Australia: Office for Learning and Teaching, Department of Education. Retrieved from http://hdl.handle.net/10536/DRO/DU:30069902

Tynan, R. (2005). The effects of threat sensitivity and face giving on dyadic psychological safety and upward communication. *Journal of Applied Social Psychology*, *35*(2), 223–247. https://doi.org/10.1111/j.1559-1816.2005.tb02119.x

Ulusoy, N., Mölders, C., Fischer, S., Bayur, H., Deveci, S., Demiral, Y., & Rössler, W. (2016). A matter of psychological safety. *Journal of Cross-Cultural Psychology*, *47*(4), 626–645. https://doi.org/10.1177/0022022115626513

Vecchio, R. P., & Anderson, R. J. (2009). Agreement in self-other ratings of leader effectiveness: The role of demographics and personality. *International Journal of Selection and Assessment*, *17*(2), 165–179. https://doi.org/10.1111/j.1468-2389.2009.00460.x

Vegt, G. S., Emans, B. J. M., & Vliert, E. (2001). Patterns of interdependence in workteams: A two-level investigation of the relations with job and team satisfaction. *Personnel Psychology*, *54*(1), 51–69. https://doi.org/10.1111/j.1744-6570.2001.tb00085.x

Visser, B. A., Ashton, M. C., & Vernon, P. A. (2008). What makes you think you're so smart? *Journal of Individual Differences*, *29*(1), 35–44. https://doi.org/10.1027/1614-0001.29.1.35

Wageman, R. (1995). Interdependence and group effectiveness. *Administrative Science Quarterly*, *40*(1), 145–180. https://doi.org/10.2307/2393703

Walker, A. (2001). British psychology students' perceptions of group-work and peer assessment. *Psychology Learning & Teaching*, *1*(1), 28–36. https://doi.org/10.2304/plat.2001.1.1.28

Wang, H.-C., Hsieh, Y.-J., & Chen, W.-F. (2016). The effect of online peer assessment in engineering education: A quasi-experimental study. *International Journal of Engineering Education*, *32*(1(A)), 199–208.

Weber, K., & Custer, R. (2005). Gender-based preferences toward technology education content, activities, and instructional methods. *Journal of Technology Education*, *16*(2), 55–71. https://doi.org/10.21061/jte.v16i2.a.4

Whitman, L. E., Malzahn, D. E., Chaparro, B. S., Russell, M., Langrall, R., & Mohler, B. A. (2005). A comparison of group processes, performance, and satisfaction in face-to-face versus computer-mediated engineering student design teams. *Journal of Engineering Education*, *94*(3), 327–337. https://doi.org/10.1002/j.2168-9830.2005.tb00857.x

Wilkens, R., & London, M. (2006). Relationships between climate, process, and performance in continuous quality improvement groups. *Journal of Vocational Behavior*, *69*(3), 510–523. https://doi.org/10.1016/j.jvb.2006.05.005

Willey, K., & Gardner, A. (2010). Investigating the capacity of self and peer assessment activities to engage students and promote learning. *European Journal of Engineering Education*, *35*(4), 429–443. https://doi.org/10.1080/03043797.2010.490577

Windschitl, M. (1999). Using small-group discussions in science lectures: A study of two professors. *College Teaching*, *47*(1), 23–27. https://doi.org/10.1080/87567559909596075

Woehr, D. J., Loignon, A. C., Schmidt, P. B., Loughry, M. L., & Ohland, M. W. (2015). Justifying aggregation with consensus-based constructs. *Organizational Research Methods*, *18*(4), 704–737. https://doi.org/10.1177/1094428115582090

Wolfe, J., Powell, B. A., Schlisserman, S., & Kirshon, A. (2016). Teamwork in engineering undergraduate classes: What problems do students experience? In *ASEE Annual Conference & Exposition*.

Wolfe, J., & Powell, E. (2009). Biases in interpersonal communication: How engineering students perceive gender typical speech acts in teamwork. *Journal of Engineering Education*, *98*(1), 5–16. https://doi.org/10.1002/j.2168-9830.2009.tb01001.x

Zhang, Y., Fang, Y., Wei, K.-K., & Chen, H. (2010). Exploring the role of psychological safety in promoting the intention to continue sharing knowledge in virtual communities. *International Journal of Information Management*, *30*(5), 425–436. https://doi.org/10.1016/j.ijinfomgt.2010.02.003

Zhao, B., & Olivera, F. (2006). Error reporting in organizations. *Academy of Management Review*, *31*(4), 1012–1030. https://doi.org/10.5465/amr.2006.22528167

Zumbo, B. D., & Zimmerman, D. W. (1993). Is the selection of statistical methods governed by level of measurement? *Canadian Psychology/Psychologie Canadienne*, *34*(4), 390–400. https://doi.org/10.1037/h0078865