The Future of Engineering Education
2024 Annual Conference & Exposition
Oregon Convention Center
Portland, OR . June 23 - 26, 2024
ASEE
Paper ID #43142

Board 115: Examining Engineering Students' Gender and Racial Effects in College Course Team Peer Assessment: A Quantitative Intersectional Approach

Miss Xiaping Li, University of Michigan

Xiaping Li is a Ph.D. candidate in Engineering Education Research at the University of Michigan. Her research interests include faculty development and change, neurodiverse college student learning experiences and outcomes, international students in engineering, and cognitive sciences. She holds a B.S. in Hydrology and Water Resources Engineering and an M.S. in Geological Sciences.

Dr. Robin Fowler, University of Michigan

Robin Fowler is a Technical Communication lecturer and a Engineering Education researcher at the University of Michigan. Her teaching is primarily in team-based engineering courses, and her research focuses on equity in communication and collaboration as well as in group design decision making (judgment) under uncertainty. She is especially interested in how power relationships and rhetorical strategies affect group judgment in engineering design; one goal of this work is to to understand factors that inhibit full participation of students who identify with historically marginalized groups and investigate evidence-based strategies for mitigating these inequities. In addition, she is interested in technology and how specific affordances can change the ways we collaborate, learn, read, and write. Teaching engineering communication allows her to apply this work as she coaches students through collaboration, design thinking, and design communication. She is part of a team of faculty innovators who originated Tandem (tandem.ai.umich.edu), a tool designed to help facilitate equitable and inclusive teamwork environments.

Dr. Mark Mills, University of Michigan

Mark Mills (he/him) is a Data Scientist on the Research & Analytics team at University of Michigan's Center for Academic Innovation. He directs and supports analytics across CAI's portfolio of educational technologies. His experience is in prediction and classification of longitudinal and hierarchically cross-classified data structures such as students in courses measured over time.

Examining Engineering Students' Gender and Racial Effects in College Course Team Peer Assessment: A Quantitative Intersectional Approach

Abstract

Peer assessment is commonly employed in college courses embracing team-based learning, with a growing focus on the design's impact on student learning outcomes. Existing research highlights the influence of factors like gender and race, yet a literature gap persists in understanding how students' gender and race impact their interactions within small groups and further shape peer assessment in the context of college course teamwork. In this work-in-progress, we employ a quantitative intersectional approach to examine gender and racial effects on peer assessment among over 1,700 engineering college students at a large research-oriented university located in the Midwest. Our analysis indicates a shift in the dominant role of male students, with females playing a more prominent role, particularly among White and Asian students. Gender-based disparities in peer assessment are associated with how White raters evaluate Asian male teammates, highlighting potential biases and the marginalization of Asian males. Furthermore, our findings highlight the underprivileged status of Minoritized groups in engineering education, regardless of their gender. This study stresses the importance of considering gender and race in peer assessment design for evaluating team-based learning outcomes. Moreover, we advocate for the inclusion of group diversity effects in terms of gender and race in future research examining team-based learning and related factors such as designed interventions.

Introduction

Teamwork is a fundamental skill for college students, and team-based learning has been incorporated into engineering courses to effectively improve student academic achievements [1] - [3]. Peer assessment, a crucial method in evaluating students' team performance, is utilized in many team-based learning courses to provide valuable feedback on student learning and teamwork contributions [4], [5].

Although previous studies have acknowledged that individual factors such as gender, race, and motivation can influence student interactions and impact teamwork assessment, potentially introducing inequities and biases in peer assessment [5] - [8], the exploration of these factors in the context of engineering higher education is limited [4]. Alqassab and colleagues conducted a systematic review of 449 research papers on peer assessment design, revealing a mere 4.14 percent focus on engineering and related domains. Furthermore, within the reviewed papers, 28 studies investigated gender as a peer assessment moderator, and only four studies considered the impact of race and culture [4]. In addition, students' individual factors, such as gender and race, are intertwined, with their intersectional effects becoming a focal point in research addressing equity and social justice in higher education [9], but not yet in most peer assessment work.

In this project, we apply intersectionality as a critical theory and approach [10] to guide our examination to identify marginalized engineering students in college course teams, recognize the inequalities they potentially experience in teamwork and peer assessment, and improve their learning experiences and well-being. Following Else-Quest and Hyde's three essential elements for intersectional research, our study simultaneously examines multiple social categories (e.g., gender and race), delves into power dynamics and inequality rooted in interconnected social categories, and recognizes the fluidity of these categories and dynamics of power across contexts and over time [10].

Engineering is often a White, male space, which leads to power imbalances and inequalities [11], [12]. This issue is exacerbated for marginalized groups, especially when considering the intersectionality of gender and race, such as female Native Americans [11], [12]. Additionally, gender and race have been shown to be related to team dynamics and teamwork effectiveness [13], [14], further justifying the adoption of an intersectional approach.

Despite the prevalent use of qualitative methods in studying intersectionality, Else-Quest and Hyde advocate for the integration of quantitative methods (e.g., multilevel modeling) with intersectional approaches in empirical research [15]. An intersectional approach can explore additive effects (e.g., main effects), multiplicative effects (e.g., interaction effects), and intersectional effects [10].

Thus, the present study aims to bridge the literature gap by exploring how engineering students' gender and race, as well as their intersection, shape peer ratings in team-based learning courses, responding to the call for the need for intersectional research to enhance social justice in higher education. The investigation delves into the influence of raters' and targets' (i.e., those being rated) gender and race in peer assessment, seeking answers to the following research questions:

> RQ1: How do the gender and race of engineering college students correlate with ratings of teammates in course teamwork?
> RQ2: How do the gender and race of engineering college students correlate with the ratings targets receive from teammates in course teamwork?
> RQ3: How can we characterize the intersectional effects of race and gender in peer ratings within engineering student teamwork?

Methods

Participants

We conducted this project at a large research-oriented university located in the Midwest. In total, this study involves data from 1,722 engineering college students, within which 1,701 students (i.e., Target) were rated by their teammates, and 1,601 students (i.e., Rater) rated their teammates' performance. These students formed 507 teams. The initial sample size was larger than 1,722, but we did not include students with missing information on gender, race, or major. Participant demographic information (Table 1) was obtained from the university's learning analytics dataset. While the institution identifies our construct of interest as "gender," we note that the data we obtained is "sex" and our data is separated into two categories which we are using as a proxy for gender in this analysis. For our race indicator, we combined institutional codes of Black, Hispanic, Native American, and Hawaiian students as a single minoritized group as the frequency of these categories was low, following common quantitative practice. We recognize that our data and analytical choices are non-ideal, and choices of convenience based on institutional data available to us as well as historical patterns of inclusion and exclusion that affect who is well-represented in our dataset.

Table 1. Participant demographic information

	Rater			Target		
	Gender			**Gender**		
Race	Female	Male	**Total** (percent)	Female	Male	**Total** (percent)
White	234	571	805 (50.3%)	240	618	858 (50.4%)
Asian	161	362	523 (32.7%)	166	380	546 (32.1%)
Minoritized	101	172	273 (17.1%)	107	190	297 (17.5%)
Total (percent)	496 (31.0%)	1,105 (69.0%)	1,601	513 (30.2%)	1,188 (69.8%)	1,701

Data Collection

Teamwork peer ratings were collected using Tandem, an online instructional tool aimed at fostering equitable teamwork. This tool was designed to address teamwork challenges and identify unfair behaviors within teams, especially those affecting marginalized student populations [16]. Peer assessments comprised eight items on 9-point Likert scales (Table 2). Peer ratings were given from a student to each of their team members at the midterm and at the end of the term.

Table 2. Tandem peer rating items

Items	**Lower anchor**	**Upper anchor**
Peer Ideas	I didn't hear many ideas from $TeamMember.	$TeamMember offered up many ideas.
Peer Teacher	$TeamMember did not explain what they were doing on a task or actively share their skills and knowledge.	$TeamMember actively teaches others and shares their skills and knowledge.
Peer Listener	$TeamMember discouraged, dismissed, or didn't listen to other teammates.	$TeamMember encouraged new perspectives by listening to other teammates.
Peer Enacted	Our project didn't include many ideas from $TeamMember.	Many of $TeamMember's ideas were used in our project.
Peer Effort	$TeamMember didn't put in as much effort as they should have.	$TeamMember did more than their fair share of work for our assignments.
Peer Quality	$TeamMember's work often needed to be redone or wasn't good enough.	$TeamMember's work for our team was exceptional.
Peer Reliability	$TeamMember was often late, was distracted while we were collaborating, or was generally unreliable.	$TeamMember always showed up, responded to messages, and was generally reliable.
Peer Valuable	$TeamMember was still gaining the skills needed for our project.	The skills $TeamMember brought to the team are incredibly valuable.

Note: $TeamMember represents a team member's name in actual surveys.

Data Analysis

The data structure is nested and crossed as shown in Figure 1. Each student provides ratings for each team member across the eight items. Therefore, the ratings (level-1) are nested within students and items (level-2), with students and items being crossed, as each student responds to each item. This crossing at level-2 is further nested within teams (level-3) in courses (level-4).

We employed a four-level linear model where responses are nested in the crossing of students and items, which in turn are nested in teams within courses, using Stata/SE 18.0. Multilevel modeling can separately estimate the peer ratings variance existing in these levels (e.g., difference between students, teams, and courses) [17]. Peer ratings (Peer rating items stacked in Table 3) serves as the dependent variable, and the main factors include raters' and targets' gender and race.

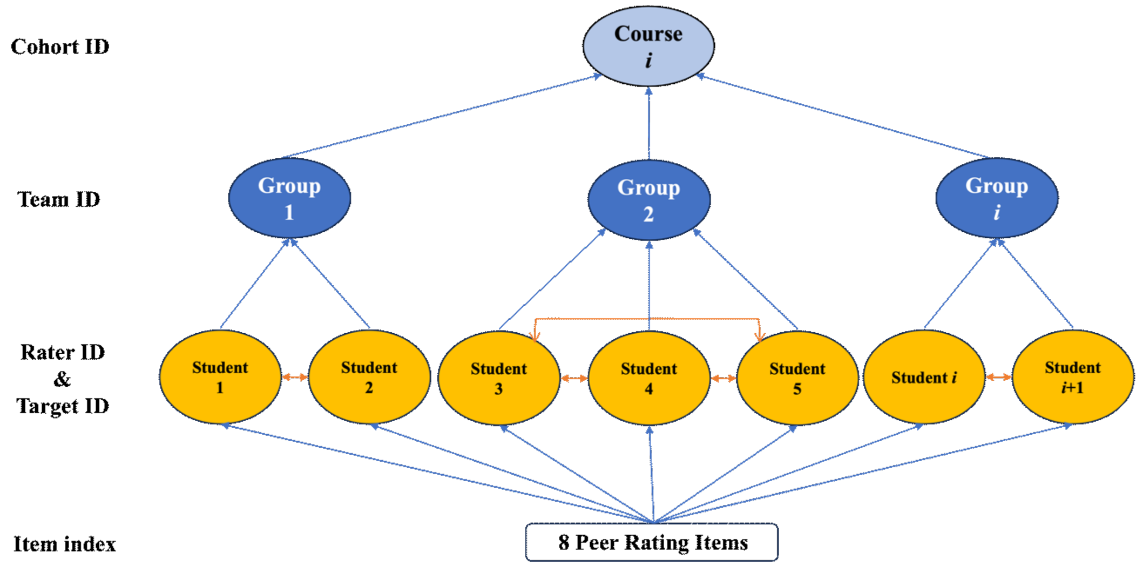


Figure 1. Data structure

Table 3. Descriptive statistics

Variables	n	Mean	Standard Deviation	Variance	Skewness	Kurtosis
Peer rating items						
Peer Ideas	10,063	7.34	1.65	2.72	-1.43	5.17
Peer Teacher	10,063	7.22	1.60	2.56	-1.18	4.62
Peer Listener	10,063	7.46	1.61	2.60	-1.40	5.04
Peer Enacted	10,063	7.22	1.65	2.72	-1.31	4.84
Peer Effort	10,063	7.11	1.70	2.91	-1.13	4.32
Peer Quality	10,063	7.55	1.50	2.25	-1.57	6.19
Peer Reliability	10,058	7.65	1.71	2.93	-1.72	5.93
Peer Valuable	10,062	7.57	1.47	2.16	-1.40	5.44
Peer rating items stacked	80,498	7.39	1.62	2.64	-1.38	5.11

Results and Discussion

The results of the multilevel model (e.g., fixed-effect and random-effect parameter estimates) are detailed in Appendix 1-1. In these descriptions, the reference level is set as female for gender and White for race. For instance, the reference group is female White raters rating female White targets for the 4-way interaction. The following subsections are arranged to answer the three research questions.

RQ1: How do the gender and race of engineering college students influence their ratings of teammates in course teamwork?

The top section of rater effects in Table 4 shows that there was no statistically significant association between the gender of raters and their evaluations of teammates. Despite a slight average difference of 0.03 higher ratings given by female raters compared to male raters, it was not statistically significant ($p = 0.54$).

In contrast, the analysis of marginal means highlights a statistically significant association between the racial identity of raters and the peer ratings they assigned. On average, White students assigned lower peer ratings by 0.16 ($p < 0.001$) and 0.22 ($p < 0.001$) compared to raters from Asian and Minoritized groups, respectively. Taken together, the findings suggest that, on average, students' race played a role in influencing their reported assessment of teammates, whereas their gender did not.

Table 4. Estimates for peer rating means and marginal effects

Independent variables	Mean	Std. err.	Marginal effects*	Std. err.	z	p	95% confidence interval Lower	Upper
Rater Gender								
Female	7.43	0.06						
Male	7.40	0.05	-0.03	0.05	-0.69	0.49	-0.13	0.06
Rater Race								
White	7.32	0.05						
Asian	7.47	0.06	0.14	0.05	2.74	0.01	0.04	0.25
Minoritized	7.56	0.07	0.24	0.07	3.65	<0.01	0.11	0.37
Target Gender								
Female	7.55	0.05						
Male	7.35	0.05	-0.20	0.04	-5.63	<0.01	-0.27	-0.13
Target Race								
White	7.50	0.05						
Asian	7.34	0.05	-0.16	0.04	-4.38	<0.01	-0.23	-0.09
Minoritized	7.28	0.06	-0.21	0.04	-4.75	<0.01	-0.30	-0.12

*Note: *Reference level for gender and race: Female for Gender and White for Race.*

In terms of the interaction between rater gender and rater race, although the fixed-effect coefficients (see Appendix 1-1) and the estimates of test commands (see Appendix 1-2) suggest statistically nonsignificant interactions, the marginal effects indicate variations in peer rating means among race and gender intersectional subgroups (see Figure 2 and Appendix 2-1). Specifically, for male raters, there were noticeable differences in how they rated their teammates across racial groups. On average, White male students assigned lower peer ratings by 0.17 ($p < 0.01$) and 0.26 ($p < 0.001$) compared to male raters from Asian and Minoritized groups, respectively. However, this pattern did not extend to female raters.

RQ2: How do the gender and race of engineering college students influence the ratings targets receive from teammates in course teamwork?

The lower section of Table 4 shows that predicted peer rating means are significantly associated with both the gender and race of targets. Female students received higher average peer ratings by 0.22 ($p < 0.001$) compared to male targets. Additionally, in comparison to their White teammates, students from Asian and Minoritized groups received lower ratings by an average of 0.15 ($p = 0.01$) and 0.23 ($p < 0.001$), respectively. Accordingly, the results indicate that, on average, female and White students received higher peer ratings from their teammates in the context of engineering student teamwork.

Similar to the findings for rater characteristics, although the fixed-effect coefficients and the estimates of test commands indicate statistically nonsignificant interactions between target gender and race, differences emerged when considering how targets were rated by their teammates across gender and racial groups (see Figure 3 and Appendix 2-2). Specifically, White and Asian female students received higher average peer ratings by 0.22 ($p < 0.001$) and 0.30 ($p = 0.001$) compared to their male counterparts, respectively. In contrast to White male targets, male students from the other two racial groups were assigned with lower average peer ratings by 0.18 ($p < 0.05$) and 0.20 ($p < 0.05$), respectively. In addition, female students from the Minoritized group were rated lower by an average 0.31 ($p < 0.01$), compared to their White female targets. While female students generally received higher peer ratings than their male teammates, this trend did not extend to female students from the Minoritized group, whose peer rating means were similar to their male counterparts.

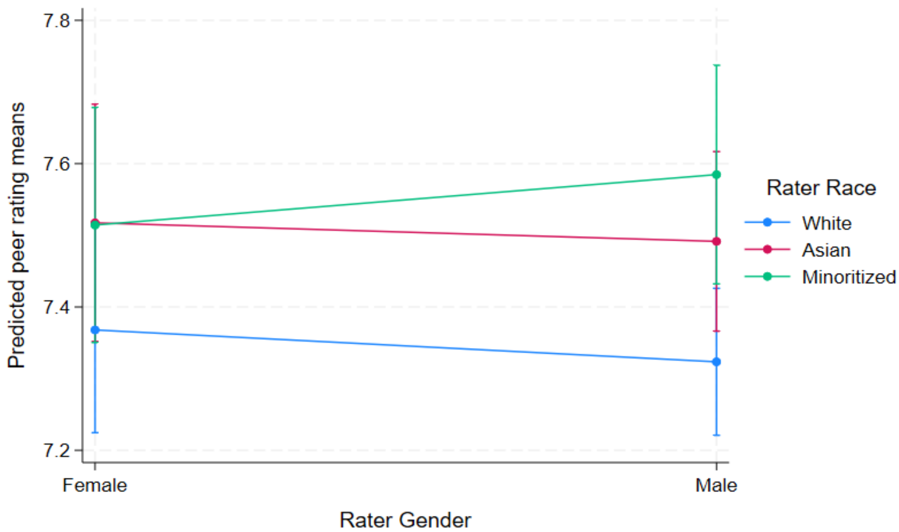


Figure 2. Mean peer rating assigned by rater gender and race
Note: Error bars represent 95% confidence intervals. Among male raters, White students (indicated by the right-hand blue point) assigned lower average ratings to their teammates compared to Asian raters (represented by the right-hand red point, p<0.01) and students from the Minoritized group (denoted by the right-hand green point, p<0.001).

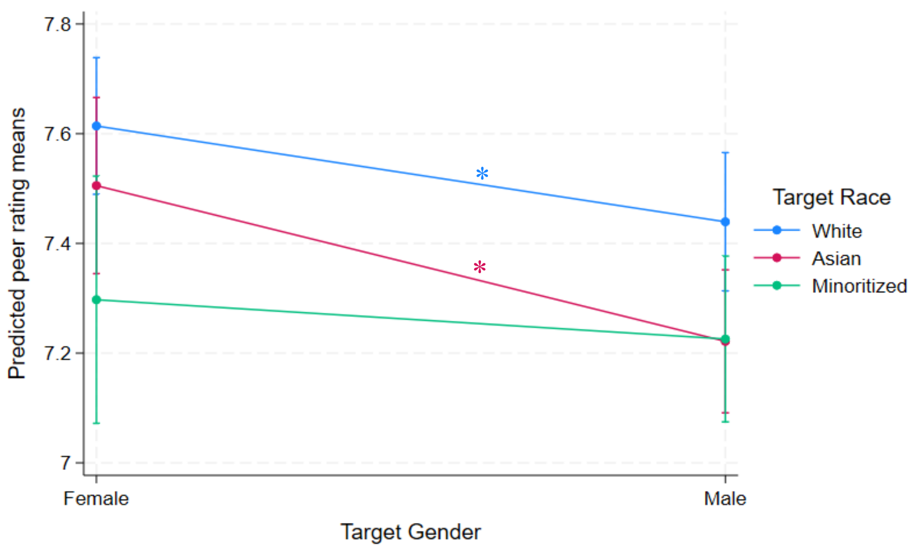
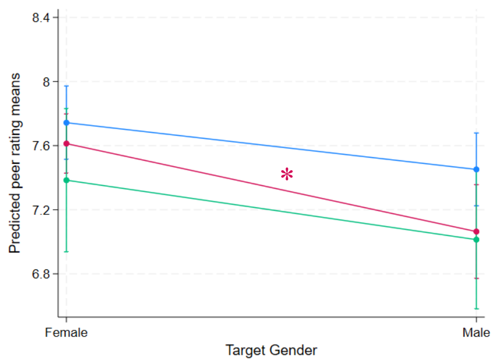
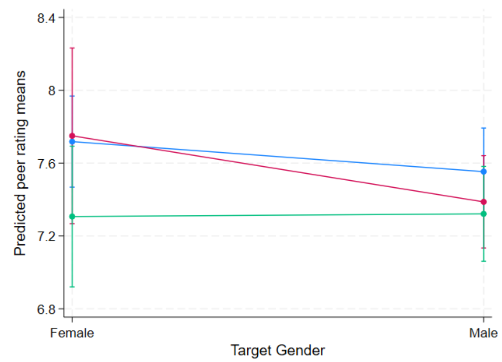


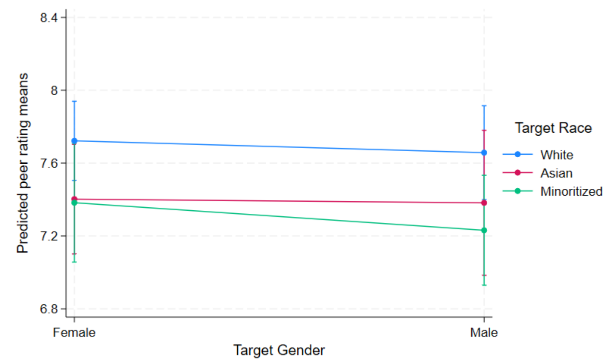
Figure 3. Mean peer rating assigned by target gender and race
Note: Error bars represent 95% confidence intervals. An asterisk (*) denotes a statistically significant difference in average peer ratings between female and male racial groups. Students from the Minoritized groups (represented by the green line) received lower average peer ratings compared to White students (indicated by the blue line), with statistical significance (p < 0.01). This trend was also observed when comparing male Asian targets (the right-hand red point) to male White targets (the right-hand blue point).



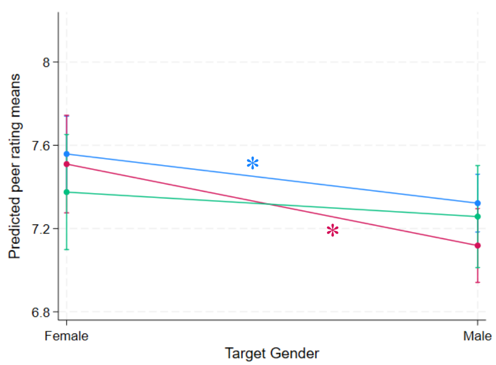
4.1. Mean peer rating assigned by White female raters
Note: White female students assigned their Asian male teammates (the right-hand red point) lower average ratings compared to White male targets (the right-hand blue point, $p<0.05$).



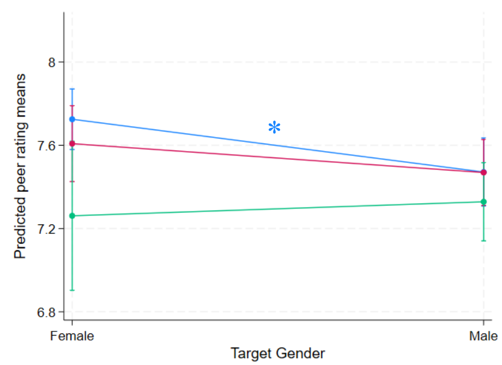
4.3. Mean peer rating assigned by Asian female raters



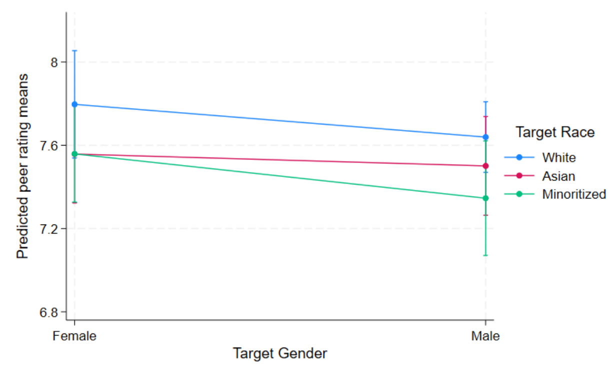
4.5. Mean peer rating assigned by female raters from Minoritized group
Note: Female students from Minoritized group assigned their Asian female teammates (the left-hand red point) lower average ratings compared to White female targets (the left-hand blue point, $p<0.05$).



4.2. Mean peer rating assigned by White male raters
Note: White male students assigned their Asian male teammates (the right-hand red point) lower average ratings compared to White male targets (the right-hand blue point, $p<0.05$).

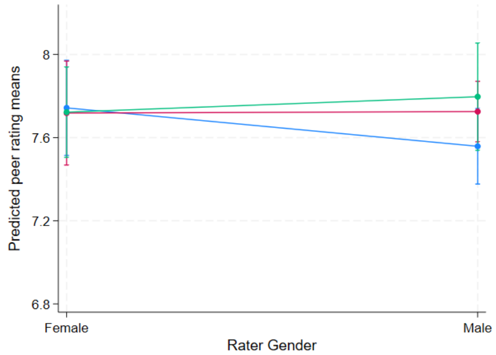


4.4. Mean peer rating assigned by Asian male raters
Note: Asian male students assigned their female teammates from Minoritized group (the left-hand green point) lower average ratings compared to White female targets (the left-hand blue point, $p<0.05$).

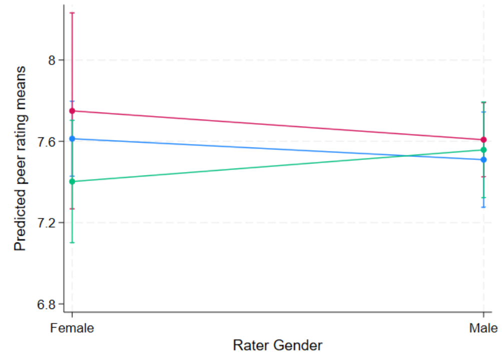


4.6. Mean peer rating assigned by male raters from Minoritized group
Note: Male students from Minoritized group assigned their male teammates from Minoritized group (the right-hand green point) lower average ratings compared to White male targets (the right-hand blue point, $p<0.05$).

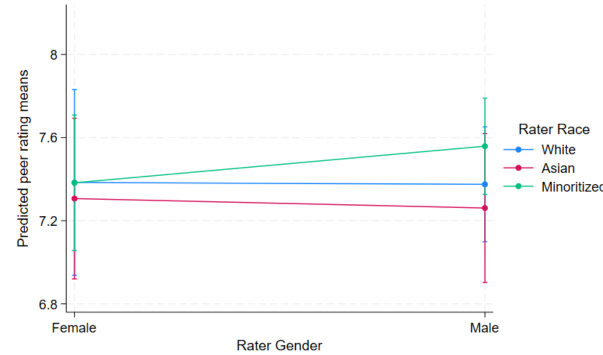
Figure 4. Mean peer rating assigned by gender and race subgroups of raters as a function of target gender and target race
Note: Error bars represent 95% confidence intervals. An asterisk (*) denotes a statistically significant difference in average peer ratings between female and male racial groups.



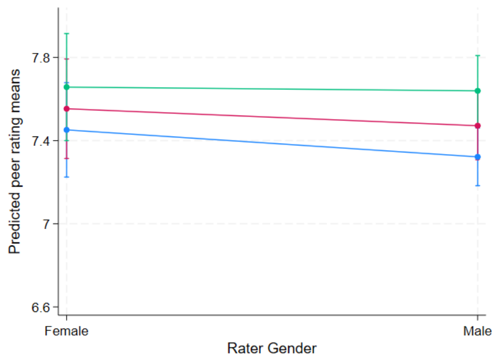
5.1. Mean peer rating assigned to <u>White female targets</u>



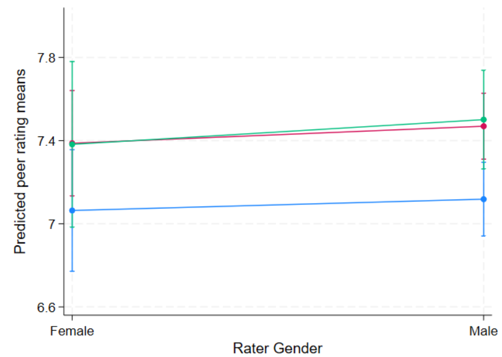
5.3. Mean peer rating assigned to <u>Asian female targets</u>



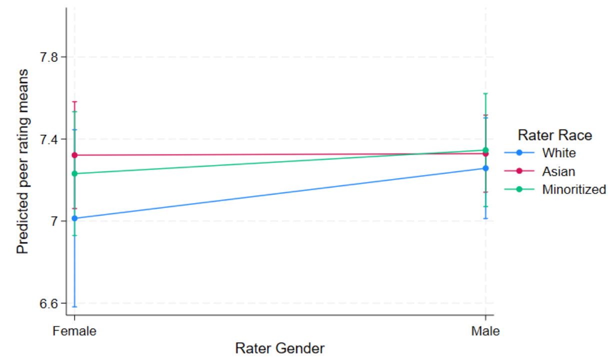
5.5. Mean peer rating assigned to <u>female targets from Minoritized group</u>



5.2. Mean peer rating assigned to <u>White male targets</u>
Note: Male students from Minoritized group (the right-hand green point) assigned their white male teammates higher average ratings compared to White male raters (the right-hand blue point, p<0.001).



5.4. Mean peer rating assigned to <u>Asian male targets</u>
Note: Asian female students (the left-hand red point) assigned their male Asian teammates higher average ratings compared to White female raters (the left-hand blue point, p<0.01). White male students assigned their male Asian teammates (the right-hand blue point) lower average ratings compared to male raters from other racial groups (the right-hand red and green point, p<0.01).



5.6. Mean peer rating assigned to <u>male targets from Minoritized group</u>

Figure 5. Mean peer rating assigned by gender and race subgroups of targets as a function of rater gender and rater race
Note: Error bars represent 95% confidence intervals.

RQ3: How can we characterize the intersectional effects of race and gender in peer ratings within engineering student teamwork?

We examined average peer ratings as a function of gender and race for each gender and race interactional subgroup of racers (Figure 4) and targets (Figure 5). Cell means for the interactions among raters' and targets' gender and race illustrate the intersectional effects of race and gender (see Appendix 3-1 and 3-2 for the complete information for the estimates).

The presence of predominantly significant *p* values, particularly evident in items related to White raters, suggests that the intersectional effects of gender and race primarily manifest within the group of White raters (see Figure 4.1 and 4.2). This may imply that White students assessed their peers differently based on the targets' race and gender, though it also reflects a larger sample for those cells. Our analysis revealed that the gender-based differences in peer ratings are predominantly associated with White raters. Notably, the most substantial disparities in predicted peer rating means between female and male targets were observed when White students evaluated their Asian teammates, suggesting that Asian male students underperformed or contributed less than Asian female students in course small group activities from the perspectives of their White male teammates. However, White students did not rate female and male teammates from the Minoritized group differently. In addition, both White and Asian male students rated their White male teammates lower than their White female teammates (see Figure 4.2 and 4.4), whereas students from the Minoritized group did not assign different scores to their teammates based on their gender (see Figure 4.5 and 4.6).

Upon further examination of the gender-based differences across racial groups, we observed that students from Asian and Minoritized groups were assessed lower compared to their White teammates. Both female and male White students assigned their Asian male teammates lower than their White male teammates (see Figure 4.1 and 4.2), while female students from the Minoritized group rated their Asian female teammates lower than their White female teammates (see Figure 4.5), on average. Moreover, female students from the Minoritized group were perceived to underperform their White female teammates in course teamwork by Asian male raters (see Figure 4.4), whereas Minoritized male raters perceived male students from the Minoritized group as underperforming compared to their White male teammates in course teamwork (see Figure 4.6), on average.

Figures 5.1 - 5.6 illustrate the impact of rater gender and race on peer rating means for each intersectional subgroup of targets. Predominantly significant *p* values present in peer rating means of White male targets (Figure 5.2) and Asian male targets (Figure 5.4). On average, male raters from the Minoritized group rated White male targets 0.32 ($p < 0.001$) higher compared to White male raters (Figure 5.2), whereas White male students assigned lower peer ratings to Asian male targets compared to Asian (mean = 0.36, $p = 0.001$) and Minoritized (mean = 0.38, $p < 0.01$) male raters, respectively (Figure 5.4). The findings may suggest that students from the Minoritized group valued the contributions of their White male teammates more than other racial groups did, while White students underestimated the performance of their Asian male teammates. Furthermore, within each female target subgroup (Figure 5.1, 5.3, and 5.5), although the Asian and Minoritized female targets rated each other slightly lower, there were no statistically significant differences in peer rating means assigned by raters based on their gender and race.

It is interesting that Asian female students did not differentiate in their peer ratings of their teammates based on targets' gender and race (Figure 4.3). In addition, their peer rating means assigned by teammates did not vary by raters' gender and race (Figure 5.3).

Conclusion

We applied a quantitative intersectional approach to examine the effects of engineering student gender and race in peer assessment in college course teamwork, given the specific items listed in Table 2. Our analysis indicates rater and target intersectional effects of gender and race (RQ1 and 2). For instance, White male students assigned lower peer ratings compared to raters from Asian and Minoritized groups. Also, White and Asian female students received higher average peer ratings than their male counterparts, and male students from other racial groups received lower average peer ratings compared to White male targets. In contrast, peer rating means of targets from the Minoritized group did not show gender-based differences. These findings may indicate a shift in the dominant role of male students in engineering, with female students taking a more prominent role in contributing to teamwork in the context of university course team-based learning, particularly among White and Asian students. This is consistent with some other work that finds similar associations, e.g., [7], suggesting that female students outperformance compared with male students may be partially attributable to their higher academic performance (e.g. course or cumulative GPA) and non-cognitive skills (e.g. communication and organization) [18], [19].

In addressing the characterization of intersectional effects of raters' and targets' race and gender in peer ratings (RQ3), our results further reveal significant simple interactions mainly among White raters, targets' gender, and targets' race. Specifically, gender-based differences in peer ratings in this study were predominantly associated with how White raters assessed their Asian male teammates, indicating the potential identity-based bias in college course team peer assessment and the potential marginalization of Asian male students in course teamwork activities. Furthermore, our results echo existing literature, highlighting the underprivileged status of the Minoritized group in engineering education, irrespective of their gender [12] - [14].

We do not find it surprising that different studies do and do not find group mean differences in peer assessment, given different contexts and different items. Our teamwork tool includes ratings of task-specific contributions to projects, as well, but because those differ across courses, it was impossible to investigate those in this large analysis. We would be unsurprised to find group mean differences across those items, though we note that the differences by subgroup as well as the directions of bias may show up differently.

Overall, this study contributes valuable insights into the complex dynamics of peer assessments in engineering college course teamwork, shedding light on the associations between peer ratings and a rater's and target's gender and race. Our findings stress the importance of considering gender and race in peer assessment design for evaluating team-based learning outcomes. Group mean differences are concerning for faculty who use peer assessments as part of a students' course assessment. Moreover, we advocate for the inclusion of group diversity effects in terms of gender and race in future research examining team-based learning and related factors such as designed interventions.

Acknowledgements

This work was supported by Enhancing Engineering Education Grants, University of Michigan and National Science Foundation under IUSE (Grant# 2120252).

References

[1] E. D. Prada, M. Mercedes, and M. Pino-Juste, "Teamwork skills in higher education: Is university training contributing to their mastery?," *Psicologia: Reflexao e Critica,* vol. 35, 2022.
[2] L. Riebe, A. Girardi, and C. Whitsed, "A systematic literature review of teamwork pedagogy in higher education," *Small Group Research*, vol. 47, no. 6, pp. 619-664, 2016.
[3] L. Springer, M. E. Stanne, and S. S. Donovan, "Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis," *Review of educational research*, vol. 69, no. 1, pp. 21-51, 1999.
[4] M. Alqassab, J. W. Strijbos, E. Panadero, J. F. Ruiz, M. Warrens, and J. To, "A systematic review of peer assessment design elements," *Educational Psychology Review*, vol. 35, no. 1, 2023.
[5] E. Panadero, M. Alqassab, J. F. Ruiz, and J. C. Ocampo, "A systematic review on peer assessment: intrapersonal and interpersonal factors," *Assessment & Evaluation in Higher Education*, vol. 48, no. 8, pp.1053-1075, Jan. 2023.
[6] M. Espey, "Gender and peer evaluations," *The Journal of Economic Education*, vol. 53, no. 1, pp. 1-10, 2022.
[7] R. Fowler, "Demographic effects on student-reported satisfaction with teams and teammates in a first-year, team-based, problem-based course," In *2016 ASEE Annual Conference & Exposition*, *New Orleans, LA, USA, June 26- 29, 2016*.
[8] J. Rhee, D. Parent, and A. Basu, "The influence of personality and ability on undergraduate teamwork and team performance," *SpringerPlus*, vol. 2, no. 1, pp. 1-14, 2013.
[9] J. C. Harris, and L. D. Patton, "Un/doing intersectionality through higher education research," *The Journal of higher education*, vol., 90, no. 3, pp. 347-372, Dec. 2019.
[10] N. M. Else-Quest, and J. S. Hyde, "Intersectionality in quantitative psychological research: I. Theoretical and epistemological issues," *Psychology of Women Quarterly*, vol. 40, no. 2, pp. 155-170, 2016.
[11] A. L. Pawley, "Learning from small numbers: Studying ruling relations that gender and race the structure of US engineering education," *Journal of Engineering Education*, vol. 108, no. 1, pp. 13-31, 2019.
[12] M. Ong, N. Jaumot-Pascual, and T. Ko. Lily, "Research literature on women of color in undergraduate engineering education: A systematic thematic synthesis," *Journal of Engineering Education*, vol. 109, no. 3, pp. 581-615, 2020.
[13] M. Borrego, K. Jennifer, L. D. McNair, and K. Beddoes, "Team effectiveness theory from industrial and organizational psychology applied to engineering student project teams: A research review," *Journal of Engineering Education*, vol. 102, no. 4, pp. 472-512, 2013.
[14] B. Beigpourian, and M. W. Ohland, "A systematized review: Gender and race in teamwork in undergraduate engineering classrooms," In *2019 ASEE Annual Conference & Exposition, Tampa, FL, USA, June 15-19, 2019*.
[15] N. M. Else-Quest, and J. S. Hyde, "Intersectionality in quantitative psychological research: II. Methods and techniques," *Psychology of Women Quarterly*, vol. 40, no. 3, pp. 319-336, 2016.
[16] R. Fowler, L. K. Alford, S. Sheffield, C. Hayward, T. S. Henderson, and R. L. Matz, "Supporting Equitable Team Experiences Using Tandem, an Online Assessment and Learning Tool," In *2021 ASEE Virtual Annual Conference Content Access*, *July 26, 2021*.

[17] J. L. Peugh, "A practical guide to multilevel modeling," *Journal of school psychology*, vol. 48, no. 1, pp. 85-112, 2010.

[18] M. Espey, "Gender and peer evaluations," *The Journal of Economic Education,* vol. 53, no. 1, pp. 1-10, 2022.

[19] B. A. Jacob, "Where the boys aren't: Non-cognitive skills, returns to school and the gender gap in higher education," *Economics of Education review,* vol. 21, no. 6, pp. 589-598, 2002.

Appendix 1-1. The Results of the Multilevel Linear Model

```
Mixed-effects regression                        Number of obs =   80,498

        Grouping information
```

Group variable	No. of groups	Observations per group Minimum	Average	Maximum
Cohort_id	42	32	1,916.6	7,608
team_id	507	8	158.8	800
Rater_stud~d	1,758	8	45.8	160
Rateeevalu~d	5,299	8	15.2	32

```
                                                Wald chi2(35) = 11878.24
Log pseudolikelihood = -130214.88               Prob > chi2   =   0.0000

                                   (Std. err. adjusted for 42 clusters in Cohort_id)
```

Peer_rating_values	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
Rater_Gender_CD						
Male	-.1847374	.1276119	-1.45	0.148	-.4348521	.0653773
Rater_Race_CD						
Asian	-.0249294	.1810888	-0.14	0.891	-.3798569	.3299981
Others	-.0206426	.139448	-0.15	0.882	-.2939556	.2526704
Rater_Gender_CD#Rater_Race_CD						
Male#Asian	.191838	.2008073	0.96	0.339	-.2017372	.5854131
Male#Others	.2590999	.1997887	1.30	0.195	-.1324787	.6506785
Ratee_Gender_CD						
Male	-.2915193	.1491957	-1.95	0.051	-.5839375	.0008989
Rater_Gender_CD#Ratee_Gender_CD						
Male#Male	.0550334	.1875735	0.29	0.769	-.3126039	.4226707
Rater_Race_CD#Ratee_Gender_CD						
Asian#Male	.1266752	.1674132	0.76	0.449	-.2014487	.4547991
Others#Male	.2266137	.1623882	1.40	0.163	-.0916613	.5448888
Rater_Gender_CD#Rater_Race_CD#Ratee_Gender_CD						
Male#Asian#Male	-.1438143	.2325111	-0.62	0.536	-.5995277	.3118991
Male#Others#Male	-.1474552	.2060687	-0.72	0.474	-.5513425	.2564321
Ratee_Race_CD						
Asian	-.1302863	.151786	-0.86	0.391	-.4277815	.1672088
Others	-.3588239	.2225897	-1.61	0.107	-.7950917	.0774439
Rater_Gender_CD#Ratee_Race_CD						
Male#Asian	.0817048	.1635224	0.50	0.617	-.2387932	.4022028
Male#Others	.1757373	.1999776	0.88	0.380	-.2162115	.5676861
Rater_Race_CD#Ratee_Race_CD						
Asian#Asian	.1620234	.3459604	0.47	0.640	-.5160465	.8400933
Asian#Others	-.0525665	.3030459	-0.17	0.862	-.6465254	.5413925
Others#Asian	-.1899876	.1723793	-1.10	0.270	-.5278448	.1478695
Others#Others	.0189568	.3179275	0.06	0.952	-.6041696	.6420832
Rater_Gender_CD#Rater_Race_CD#Ratee_Race_CD						
Male#Asian#Asian	-.2307316	.3715209	-0.62	0.535	-.9588993	.497436
Male#Asian#Others	-.228261	.3682644	-0.62	0.535	-.950046	.4935241
Male#Others#Asian	-.0002061	.2178476	-0.00	0.999	-.4271796	.4267674
Male#Others#Others	-.0739592	.2977703	-0.25	0.804	-.6575782	.5096598

Ratee_Gender_CD#Ratee_Race_CD						
Male#Asian	-.2574889	.2027474	-1.27	0.204	-.6548666	.1398888
Male#Others	-.0793104	.2948988	-0.27	0.788	-.6573014	.4986805
Rater_Gender_CD#Ratee_Gender_CD#Ratee_Race_CD						
Male#Male#Asian	.1025257	.2475018	0.41	0.679	-.3825688	.5876203
Male#Male#Others	.1980287	.2643962	0.75	0.454	-.3201783	.7162357
Rater_Race_CD#Ratee_Gender_CD#Ratee_Race_CD						
Asian#Male#Asian	.0597346	.3580736	0.17	0.868	-.6420768	.761546
Asian#Male#Others	.2587587	.3667476	0.71	0.480	-.4600534	.9775707
Others#Male#Asian	.3022378	.2948412	1.03	0.305	-.2756403	.8801158
Others#Male#Others	-.0070794	.3224072	-0.02	0.982	-.638986	.6248271
Rater_Gender_CD#Rater_Race_CD#Ratee_Gender_CD# Ratee_Race_CD						
Male#Asian#Male#Asian	.2101202	.4228333	0.50	0.619	-.6186179	1.038858
Male#Asian#Male#Others	-.0565626	.3784591	-0.15	0.881	-.7983287	.6852036
Male#Others#Male#Asian	-.0469387	.3731292	-0.13	0.900	-.7782584	.684381
Male#Others#Male#Others	-.1671774	.3384256	-0.49	0.621	-.8304794	.4961246
_cons	7.743056	.1167122	66.34	0.000	7.514304	7.971807

Random-effects parameters	Estimate	Robust std. err.	[95% conf.	interval]
Cohort_id: Independent				
var(1.Intervention_ID)	2.20e-08	2.77e-07	4.52e-19	1073.046
var(_cons)	.0475837	.0609151	.0038706	.5849754
team_id: Identity				
var(_cons)	.3553752	.6051137	.0126272	10.00155
Rater_stud~d: Unstructured				
var(2.Time_id)	.5768921	.0338343	.5142475	.6471679
var(_cons)	.6598425	.2111006	.3524674	1.235269
cov(2.Time_id,_cons)	-.284126	.0490943	-.3803492	-.1879029
Rateeevalu~d: Identity				
var(R.peer_rating_index)	.8959557	.0641917	.7785765	1.031031
var(Residual)	.7181952	.0338081	.6548974	.7876109

. estat ic

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	80,498	.	-130214.9	44	260517.8	260926.8

Note: BIC uses N = number of observations. See **[R] IC note**.

Appendix 1-2. Estimates of the Interactions between Predictors

```
. testparm i.Rater_Gender_CD#i.Rater_Race_CD#i.Ratee_Gender_CD#i.Ratee_Race_CD

 ( 1)  [Peer_rating_values]2.Rater_Gender_CD#2.Rater_Race_CD#2.Ratee_Gender_CD#2.Ratee_Race_CD = 0
 ( 2)  [Peer_rating_values]2.Rater_Gender_CD#2.Rater_Race_CD#2.Ratee_Gender_CD#3.Ratee_Race_CD = 0
 ( 3)  [Peer_rating_values]2.Rater_Gender_CD#3.Rater_Race_CD#2.Ratee_Gender_CD#2.Ratee_Race_CD = 0
 ( 4)  [Peer_rating_values]2.Rater_Gender_CD#3.Rater_Race_CD#2.Ratee_Gender_CD#3.Ratee_Race_CD = 0

           chi2(  4) =    0.61
         Prob > chi2 =    0.9621

. testparm i.Rater_Gender_CD#i.Rater_Race_CD

 ( 1)  [Peer_rating_values]2.Rater_Gender_CD#2.Rater_Race_CD = 0
 ( 2)  [Peer_rating_values]2.Rater_Gender_CD#3.Rater_Race_CD = 0

           chi2(  2) =    1.92
         Prob > chi2 =    0.3823

. testparm i.Ratee_Gender_CD#i.Ratee_Race_CD

 ( 1)  [Peer_rating_values]2.Ratee_Gender_CD#2.Ratee_Race_CD = 0
 ( 2)  [Peer_rating_values]2.Ratee_Gender_CD#3.Ratee_Race_CD = 0

           chi2(  2) =    1.62
         Prob > chi2 =    0.4444

. testparm i.Rater_Race_CD#i.Ratee_Gender_CD

 ( 1)  [Peer_rating_values]2.Rater_Race_CD#2.Ratee_Gender_CD = 0
 ( 2)  [Peer_rating_values]3.Rater_Race_CD#2.Ratee_Gender_CD = 0

           chi2(  2) =    1.95
         Prob > chi2 =    0.3776

. testparm i.Rater_Gender_CD#i.Ratee_Gender_CD

 ( 1)  [Peer_rating_values]2.Rater_Gender_CD#2.Ratee_Gender_CD = 0

           chi2(  1) =    0.09
         Prob > chi2 =    0.7692

. testparm i.Rater_Gender_CD#i.Rater_Race_CD#i.Ratee_Gender_CD

 ( 1)  [Peer_rating_values]2.Rater_Gender_CD#2.Rater_Race_CD#2.Ratee_Gender_CD = 0
 ( 2)  [Peer_rating_values]2.Rater_Gender_CD#3.Rater_Race_CD#2.Ratee_Gender_CD = 0

           chi2(  2) =    0.58
         Prob > chi2 =    0.7499

. testparm i.Rater_Race_CD#i.Ratee_Gender_CD#i.Ratee_Race_CD

 ( 1)  [Peer_rating_values]2.Rater_Race_CD#2.Ratee_Gender_CD#2.Ratee_Race_CD = 0
 ( 2)  [Peer_rating_values]2.Rater_Race_CD#2.Ratee_Gender_CD#3.Ratee_Race_CD = 0
 ( 3)  [Peer_rating_values]3.Rater_Race_CD#2.Ratee_Gender_CD#2.Ratee_Race_CD = 0
 ( 4)  [Peer_rating_values]3.Rater_Race_CD#2.Ratee_Gender_CD#3.Ratee_Race_CD = 0

           chi2(  4) =    1.85
         Prob > chi2 =    0.7633

. testparm i.Ratee_Gender_CD#i.Ratee_Race_CD

 ( 1)  [Peer_rating_values]2.Ratee_Gender_CD#2.Ratee_Race_CD = 0
 ( 2)  [Peer_rating_values]2.Ratee_Gender_CD#3.Ratee_Race_CD = 0

           chi2(  2) =    1.62
         Prob > chi2 =    0.4444
```

Appendix 2-1. Marginal means and effects for Rater's gender and race

```
. margins  i.Rater_Gender_CD#i.Rater_Race_CD

Predictive margins                                     Number of obs = 80,498
Model VCE: Robust

Expression: Linear prediction, fixed portion, predict()
```

	Margin	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
Rater_Gender_CD#Rater_Race_CD						
Female#White	7.36796	.0731781	100.69	0.000	7.224534	7.511387
Female#Asian	7.517398	.0844908	88.97	0.000	7.351799	7.682997
Female#Others	7.514376	.0836622	89.82	0.000	7.350402	7.678351
Male#White	7.323471	.0522992	140.03	0.000	7.220966	7.425975
Male#Asian	7.491536	.0639165	117.21	0.000	7.366262	7.61681
Male#Others	7.584724	.0778278	97.46	0.000	7.432185	7.737264

```
margins  i.Rater_Gender_CD, dydx(i.Rater_Race_CD)

Average marginal effects                               Number of obs = 80,498
Model VCE: Robust

Expression: Linear prediction, fixed portion, predict()
dy/dx wrt:  2.Rater_Race_CD 3.Rater_Race_CD
```

	dy/dx	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
1.Rater_Race_CD	(base outcome)					
2.Rater_Race_CD						
Rater_Gender_CD						
Female	.1494383	.0759275	1.97	0.049	.0006231	.2982535
Male	.168065	.056517	2.97	0.003	.0572936	.2788363
3.Rater_Race_CD						
Rater_Gender_CD						
Female	.1464164	.0968879	1.51	0.131	-.0434804	.3363131
Male	.2612538	.0658307	3.97	0.000	.132228	.3902795

```
Note: dy/dx for factor levels is the discrete change from the base level.

. margins i.Rater_Race_CD, dydx(i.Rater_Gender_CD)

Average marginal effects                               Number of obs = 80,498
Model VCE: Robust

Expression: Linear prediction, fixed portion, predict()
dy/dx wrt:  2.Rater_Gender_CD
```

	dy/dx	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
1.Rater_Gender_CD	(base outcome)					
2.Rater_Gender_CD						
Rater_Race_CD						
White	-.0444894	.0521265	-0.85	0.393	-.1466555	.0576768
Asian	-.0258627	.0873654	-0.30	0.767	-.1970958	.1453704
Others	.070348	.0870792	0.81	0.419	-.1003242	.2410202

```
Note: dy/dx for factor levels is the discrete change from the base level.
```

Appendix 2-2. Marginal means and effects for Target's gender and race

```
margins  i.Ratee_Gender_CD#i.Ratee_Race_CD

Predictive margins                                 Number of obs = 80,498
Model VCE: Robust

Expression: Linear prediction, fixed portion, predict()
```

	Margin	Delta-method std. err.	z	P>|z|	[95% conf. interval]	
Ratee_Gender_CD#Ratee_Race_CD						
Female#White	7.676818	.0603126	127.28	0.000	7.558607	7.795028
Female#Asian	7.571222	.0807453	93.77	0.000	7.412964	7.72948
Female#Others	7.362748	.1181057	62.34	0.000	7.131265	7.594231
Male#White	7.454643	.0652395	114.27	0.000	7.326776	7.58251
Male#Asian	7.275169	.0682879	106.54	0.000	7.141327	7.409011
Male#Others	7.250598	.0831969	87.15	0.000	7.087535	7.413661

```
. margins i.Ratee_Race_CD, dydx(i.Ratee_Gender_CD)

Average marginal effects                           Number of obs = 80,498
Model VCE: Robust

Expression: Linear prediction, fixed portion, predict()
dy/dx wrt:  2.Ratee_Gender_CD
```

	dy/dx	Delta-method std. err.	z	P>|z|	[95% conf. interval]	
1.Ratee_Gender_CD	(base outcome)					
2.Ratee_Gender_CD						
Ratee_Race_CD						
White	**-.2221749**	**.0537443**	**-4.13**	**0.000**	**-.3275117**	**-.1168381**
Asian	**-.2960534**	**.0927471**	**-3.19**	**0.001**	**-.4778343**	**-.1142724**
Others	**-.1121498**	**.1399576**	**-0.80**	**0.423**	**-.3864616**	**.162162**

Note: dy/dx for factor levels is the discrete change from the base level.

```
. margins  i.Ratee_Gender_CD, dydx(i.Ratee_Race_CD)

Average marginal effects                           Number of obs = 80,498
Model VCE: Robust

Expression: Linear prediction, fixed portion, predict()
dy/dx wrt:  2.Ratee_Race_CD 3.Ratee_Race_CD
```

	dy/dx	Delta-method std. err.	z	P>|z|	[95% conf. interval]	
1.Ratee_Race_CD	(base outcome)					
2.Ratee_Race_CD						
Ratee_Gender_CD						
Female	-.1055956	.0737477	-1.43	0.152	-.2501384	.0389472
Male	-.1794741	.081208	-2.21	0.027	-.3386388	-.0203093
3.Ratee_Race_CD						
Ratee_Gender_CD						
Female	-.3140696	.1115869	-2.81	0.005	-.5327758	-.0953634
Male	-.2040445	.0811499	-2.51	0.012	-.3630954	-.0449936

Note: dy/dx for factor levels is the discrete change from the base level.

Appendix 3-1. Marginal effects

```
margins i.Ratee_Race_CD, dydx(i.Ratee_Gender_CD) at (Rater_Gender_CD=(1 2) Rater_Race_CD=(1 2 3))

Conditional marginal effects                          Number of obs = 80,498
Model VCE: Robust

Expression: Linear prediction, fixed portion, predict()
dy/dx wrt:  2.Ratee_Gender_CD
1._at: Rater_Gender_CD = 1
       Rater_Race_CD   = 1
2._at: Rater_Gender_CD = 1
       Rater_Race_CD   = 2
3._at: Rater_Gender_CD = 1
       Rater_Race_CD   = 3
4._at: Rater_Gender_CD = 2
       Rater_Race_CD   = 1
5._at: Rater_Gender_CD = 2
       Rater_Race_CD   = 2
6._at: Rater_Gender_CD = 2
       Rater_Race_CD   = 3
```

Race: 1=White, 2=Asian, 3=the Minoritized group
Gender: 1=Female and 2=Male

	dy/dx	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
1.Ratee_Gender_CD	(base outcome)					
2.Ratee_Gender_CD						
_at#Ratee_Race_CD						
1#White	-.2915193	.1491957	-1.95	0.051	-.5839375	.0008989
1#Asian	-.5490081	.1207358	-4.55	0.000	-.7856459	-.3123704
1#Others	-.3708297	.2797173	-1.33	0.185	-.9190655	.1774061
2#White	-.1648441	.1049392	-1.57	0.116	-.3705212	.0408331
2#Asian	-.3625983	.2907051	-1.25	0.212	-.9323699	.2071732
2#Others	.0146042	.2250272	0.06	0.948	-.426441	.4556493
3#White	-.0649055	.1030177	-0.63	0.529	-.2668166	.1370055
3#Asian	-.0201566	.2177329	-0.09	0.926	-.4469053	.406592
3#Others	-.1512954	.2044468	-0.74	0.459	-.5520039	.249413
4#White	-.2364858	.1013398	-2.33	0.020	-.4351083	-.0378634
4#Asian	-.391449	.1591864	-2.46	0.014	-.7034486	-.0794493
4#Others	-.1177676	.2098865	-0.56	0.575	-.5291375	.2936024
5#White	-.2536249	.0840957	-3.02	0.003	-.4184494	-.0888005
5#Asian	-.1387333	.083301	-1.67	0.096	-.3020002	.0245336
5#Others	.0672894	.1789209	0.38	0.707	-.2833891	.4179679
6#White	-.1573273	.1041755	-1.51	0.131	-.3615075	.046853
6#Asian	-.0569913	.138349	-0.41	0.680	-.3281503	.2141676
6#Others	-.2128659	.1734677	-1.23	0.220	-.5528564	.1271246

Note: dy/dx for factor levels is the discrete change from the base level.

```
. margins i.Ratee_Gender_CD, dydx(i.Ratee_Race_CD) at (Rater_Gender_CD=(1 2) Rater_Race_CD=(1 2 3))

Conditional marginal effects                          Number of obs = 80,498
Model VCE: Robust

Expression: Linear prediction, fixed portion, predict()
dy/dx wrt:  2.Ratee_Race_CD 3.Ratee_Race_CD
1._at: Rater_Gender_CD = 1
       Rater_Race_CD   = 1
2._at: Rater_Gender_CD = 1
       Rater_Race_CD   = 2
3._at: Rater_Gender_CD = 1
       Rater_Race_CD   = 3
4._at: Rater_Gender_CD = 2
       Rater_Race_CD   = 1
5._at: Rater_Gender_CD = 2
       Rater_Race_CD   = 2
6._at: Rater_Gender_CD = 2
       Rater_Race_CD   = 3
```

	dy/dx	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
1.Ratee_Race_CD	(base outcome)					
2.Ratee_Race_CD						
_at#Ratee_Gender_CD						
1#Female	-.1302863	.151786	-0.86	0.391	-.4277815	.1672088
1#Male	-.3877752	.1904141	-2.04	0.042	-.76098	-.0145705
2#Female	.0317371	.2826269	0.11	0.911	-.5222015	.5856757
2#Male	-.1660172	.1526024	-1.09	0.277	-.4651125	.133078
3#Female	-.320274	.1478748	-2.17	0.030	-.6101032	-.0304447
3#Male	-.2755251	.2548096	-1.08	0.280	-.7749428	.2238926
4#Female	-.0485816	.1142253	-0.43	0.671	-.2724591	.175296
4#Male	-.2035447	.1016285	-2.00	0.045	-.4027328	-.0043566
5#Female	-.1172898	.0846315	-1.39	0.166	-.2831644	.0485848
5#Male	-.0023982	.0917753	-0.03	0.979	-.1822745	.1774782
6#Female	-.2387753	.1192604	-2.00	0.045	-.4725215	-.0050292
6#Male	-.1384394	.1386436	-1.00	0.318	-.4101759	.133297
3.Ratee_Race_CD						
_at#Ratee_Gender_CD						
1#Female	-.3588239	.2225897	-1.61	0.107	-.7950917	.0774439
1#Male	-.4381344	.26028	-1.68	0.092	-.9482738	.0720051
2#Female	-.4113904	.2305092	-1.78	0.074	-.8631801	.0403994
2#Male	-.2319422	.1406341	-1.65	0.099	-.50758	.0436956
3#Female	-.3398671	.1992335	-1.71	0.088	-.7303576	.0506234
3#Male	-.426257	.2345526	-1.82	0.069	-.8859716	.0334576
4#Female	-.1830866	.1228738	-1.49	0.136	-.4239148	.0577416
4#Male	-.0643683	.1266759	-0.51	0.611	-.3126486	.183912
5#Female	-.463914	.1891225	-2.45	0.014	-.8345873	-.0932408
5#Male	-.1429997	.0995654	-1.44	0.151	-.3381442	.0521448
6#Female	-.238089	.1207809	-1.97	0.049	-.4748152	-.0013627
6#Male	-.2936276	.1285438	-2.28	0.022	-.5455688	-.0416863

Note: dy/dx for factor levels is the discrete change from the base level.

Appendix 3-2. Marginal effects

```
. margins i.Rater_Race_CD, dydx(i.Rater_Gender_CD ) at (Ratee_Gender_CD=(1 2) Ratee_Race_CD=(1 2 3))

Conditional marginal effects                          Number of obs = 80,498
Model VCE: Robust

Expression: Linear prediction, fixed portion, predict()
dy/dx wrt:  2.Rater_Gender_CD
1._at: Ratee_Gender_CD = 1
       Ratee_Race_CD   = 1
2._at: Ratee_Gender_CD = 1
       Ratee_Race_CD   = 2
3._at: Ratee_Gender_CD = 1
       Ratee_Race_CD   = 3
4._at: Ratee_Gender_CD = 2
       Ratee_Race_CD   = 1
5._at: Ratee_Gender_CD = 2
       Ratee_Race_CD   = 2
6._at: Ratee_Gender_CD = 2
       Ratee_Race_CD   = 3
```

	dy/dx	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
1.Rater_Gender_CD	(base outcome)					
2.Rater_Gender_CD						
_at#Rater_Race_CD						
1#White	-.1847374	.1276119	-1.45	0.148	-.4348521	.0653773
1#Asian	.0071006	.1337487	0.05	0.958	-.2550421	.2692433
1#Others	.0743625	.1279017	0.58	0.561	-.1763203	.3250454
2#White	-.1030326	.1171909	-0.88	0.379	-.3327225	.1266573
2#Asian	-.1419263	.2561357	-0.55	0.580	-.643943	.3600904
2#Others	.1558612	.1763456	0.88	0.377	-.1897698	.5014922
3#White	-.0090001	.1838465	-0.05	0.961	-.3693325	.3513324
3#Asian	-.0454231	.2463344	-0.18	0.854	-.5282296	.4373834
3#Others	.1761407	.17157	1.03	0.305	-.1601303	.5124117
4#White	-.129704	.1015442	-1.28	0.201	-.3287268	.0693189
4#Asian	-.0816803	.1323512	-0.62	0.537	-.3410839	.1777232
4#Others	-.0180592	.1553344	-0.12	0.907	-.3225089	.2863906
5#White	.0545265	.1486936	0.37	0.714	-.2369076	.3459607
5#Asian	.0819387	.1019518	0.80	0.422	-.1178832	.2817607
5#Others	.1190265	.217654	0.55	0.584	-.3075674	.5456204
6#White	.2440621	.2082079	1.17	0.241	-.1640179	.6521421
6#Asian	.0072622	.163532	0.04	0.965	-.3132548	.3277791
6#Others	.1145702	.2159483	0.53	0.596	-.3086806	.5378211

Note: dy/dx for factor levels is the discrete change from the base level.

```
margins i.Rater_Gender_CD, dydx(i.Rater_Race_CD) at (Ratee_Gender_CD=(1 2) Ratee_Race_CD=(1 2 3))

Conditional marginal effects                              Number of obs = 80,498
Model VCE: Robust

Expression: Linear prediction, fixed portion, predict()
dy/dx wrt:  2.Rater_Race_CD 3.Rater_Race_CD
1._at: Ratee_Gender_CD = 1
       Ratee_Race_CD   = 1
2._at: Ratee_Gender_CD = 1
       Ratee_Race_CD   = 2
3._at: Ratee_Gender_CD = 1
       Ratee_Race_CD   = 3
4._at: Ratee_Gender_CD = 2
       Ratee_Race_CD   = 1
5._at: Ratee_Gender_CD = 2
       Ratee_Race_CD   = 2
6._at: Ratee_Gender_CD = 2
       Ratee_Race_CD   = 3
```

	dy/dx	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
1.Rater_Race_CD	(base outcome)					
2.Rater_Race_CD						
_at#Rater_Gender_CD						
1#Female	-.0249294	.1810888	-0.14	0.891	-.3798569	.3299981
1#Male	.1669086	.1091166	1.53	0.126	-.0469561	.3807732
2#Female	.137094	.2485571	0.55	0.581	-.3500689	.624257
2#Male	.0982003	.1190788	0.82	0.410	-.1351898	.3315905
3#Female	-.0774959	.2789499	-0.28	0.781	-.6242276	.4692359
3#Male	-.1139189	.1663971	-0.68	0.494	-.4400513	.2122135
4#Female	.1017458	.1427766	0.71	0.476	-.1780911	.3815828
4#Male	.1497694	.0830406	1.80	0.071	-.0129872	.3125261
5#Female	.3235038	.1412701	2.29	0.022	.0466195	.6003882
5#Male	.350916	.1028745	3.41	0.001	.1492857	.5525463
6#Female	.307938	.2415519	1.27	0.202	-.165495	.781371
6#Male	.0711381	.1395451	0.51	0.610	-.2023653	.3446415
3.Rater_Race_CD						
_at#Rater_Gender_CD						
1#Female	-.0206426	.139448	-0.15	0.882	-.2939556	.2526704
1#Male	.2384573	.151001	1.58	0.114	-.0574992	.5344139
2#Female	-.2106302	.1595848	-1.32	0.187	-.5234107	.1021502
2#Male	.0482636	.136165	0.35	0.723	-.2186149	.315142
3#Female	-.0016858	.2840683	-0.01	0.995	-.5584494	.5550778
3#Male	.1834549	.1692295	1.08	0.278	-.1482288	.5151387
4#Female	.2059711	.1141398	1.80	0.071	-.0177388	.429681
4#Male	.3176159	.0768879	4.13	0.000	.1669184	.4683134
5#Female	.3182212	.2584803	1.23	0.218	-.1883909	.8248334
5#Male	.3827212	.132882	2.88	0.004	.1222772	.6431652
6#Female	.2178485	.1779513	1.22	0.221	-.1309297	.5666266
6#Male	.0883566	.1764051	0.50	0.616	-.2573911	.4341044

Note: dy/dx for factor levels is the discrete change from the base level.