Investigating the Impact of College Students' Personal Characteristics on Peer Assessment: A Multilevel Linear Modeling 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 encompass 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.

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.

Investigating the Impact of College Students' Personal Characteristics on Peer Assessment: A Multilevel Linear Modeling Approach

Abstract

Peer assessment, an essential method in team-based learning, offers valuable feedback. While previous research has explored correlations between peer ratings and personal characteristics like gender, academic performance, and personality traits, there's still a gap in comprehensively understanding how these factors influence peer ratings in college teamwork. To address this, our study employs multilevel linear modeling to investigate the relationships between these factors and peer rating scores in the context of college course teamwork. We used Tandem to collect peer rating data from 5,322 college students at a Midwest research university spanning the period from 2019 to 2023. Our analysis reveals statistically significant associations between students' peer rating values and their personal factors. Female students, students with higher GPAs, or those preferring working alone were more likely assigned higher peer ratings, while those rating themselves higher in extraversion and task control tended to receive lower ratings. In addition, the multiple-way interactions among personal characteristics suggest that academic performance is more influential among these factors analyzed. These results underscore the importance of considering personal factors in peer assessment design for team-based learning outcomes and future research in educational interventions.

Introduction

The development of effective collaboration within a team is acknowledged as an essential skill for college students, with proven benefits for their learning [1], [2]. Recognizing its significance, various academic disciplines have integrated teamwork into their curricula, necessitating the assessment of its effectiveness [3]. Peer assessment, a crucial assessment method commonly employed in team-based learning courses, provides valuable feedback and enhances student learning outcomes [4]. As a specific method of peer assessment, peer rating entails team members assigning ratings based on predefined performance criteria using diverse rating scales [5].

However, peer rating values may not fully or accurately capture the contributions individuals make to their teams and the skills they develop during group tasks, as peer ratings may be influenced by various other factors such as personal characteristics [6] - [7], group diversity [8] - [9], peer assessment design and structure [10], and assessment bias [5]. Existing research has underscored correlations between peer performance and ratings and personal characteristics, including gender [11] - [13], academic performance [6], personality traits [7], [14], and group preferences [15]. For instance, students with higher GPAs than their teammates are more likely to receive elevated peer rating values, indicating a positive correlation between academic performance and peer assessments [6]. In a study by Watson and colleagues (2010) involving 287 college students enrolled in a management course, multilevel linear modeling was employed to investigate individual characteristics. The results suggested that students displaying a proactive attitude and a tendency to take on more tasks received higher evaluations from their peers [14]. Their study also indicated no significant relationship between gender and peer ratings, which contrasts with findings in other studies [13], [16]. Some studies have suggested that female students receive higher marks from their teammates [6], [13], while others have identified greater peer ratings for male students [16] - [17].

Collectively, these prior studies paint a nuanced picture, highlighting a gap in our understanding of how these factors distinctly impact peer ratings within the context of college course teamwork. This knowledge deficit as to if and how personal characteristics enter into rating the performance of peers hinders our ability to provide instructors, students, and researchers with evidence-based insights into the effectiveness of team-based pedagogy and into the interpretation of peer assessments. To enhance the existing literature, our study employs multilevel linear modeling to investigate the relationships between targets (i.e., students being rated by their teammates)' personal characteristics (e.g., performance and demographic) factors and peer rating scores in the context of college course teamwork. This study attempts to answer the following research questions:

> RQ1. What is the relationship between the peer ratings college students receive and their demographic (gender, personality traits, and group preferences) and performance characteristics (cumulative GPA)?
> RQ2. How do interactions among these factors influence the college students' peer rating values?

Methods

Participants

The present study included data from 5,322 college students (2,292 female and 3,031 male) at a Midwest research university during the period from 2019 to 2023 (Table 1). These students were team members of 1,572 teams across 58 courses (where a course is a unique combination of course and term) spanning a wide range of disciplines: Engineering (n = 1,797), Sciences (n = 325), Arts and Sciences (n = 1,522), and Business (n = 2,003). Participant gender, major, and cumulative GPA information (up to enrollment in the team-based course) were obtained from the university's learning analytics dataset.

Table 1. Student demographics and group preferences.

Demographics and Group preferences	*n*	*percent*
*Sex**		
Female	2,291	43%
Male	3,031	56%
Major		
Engineering	1,797	34%
Arts and Sciences	1522	28%
Business	2,003	38%
Group preferences		
Work alone	1,263	24%
Work with one partner	1,765	33%
Work in a group	2,294	43%

Note: While gender is our characteristic of interest, "sex" is the data we were able to capture from the university dataset.

Data Collection

The teamwork survey data were collected using Tandem, a digital instructional tool designed to foster equitable teamwork. One of Tandem's missions is to identify unfair team behaviors and address issues of teamwork, especially as they affect marginalized student populations [18].

Participants reported their personality traits and group preferences at the course commencement, receiving evaluations from teammates at mid-term and end of term from 2019 to 2023. Personality trait items, such as extraversion and task control, were rated on 7-point Likert scales. Peer assessment included eight items with 9-point Likert scales, where students were rated individually by their teammates (Appendix 1). Group preference was categorized into three options: Work alone, work with one partner, and work in a group (Table 1).

Data Analysis

Figure 1 visualizes the nested and crossed data structure. Students rated and were rated by each of their team members on each of the eight items. Thus, ratings (level-1) are nested with students and items (level-2), where students and items are crossed given that each student responds to each item. The crossing at level-2 is in turn nested within teams (level-3) in courses (level-4). This nested and crossed data structure poses a challenge to the independence assumption (i.e., units of sampling are independent from one another) required for traditional statistical analyses like least-squares analysis of variance [19]. The inherent violation of independence due to nesting highlights the need for employing multilevel modeling, as traditional analytical models typically specify only one or two sources of variance and therefore introduce an elevated risk of Type I errors and biased parameter estimates [19]. In contrast, multilevel modeling allows for the specification of as many random effects as there are sources of variance, which is more reliable in accounting for unobserved heterogeneity, such as individual differences and group dynamics not directly measured or included in this study [20].

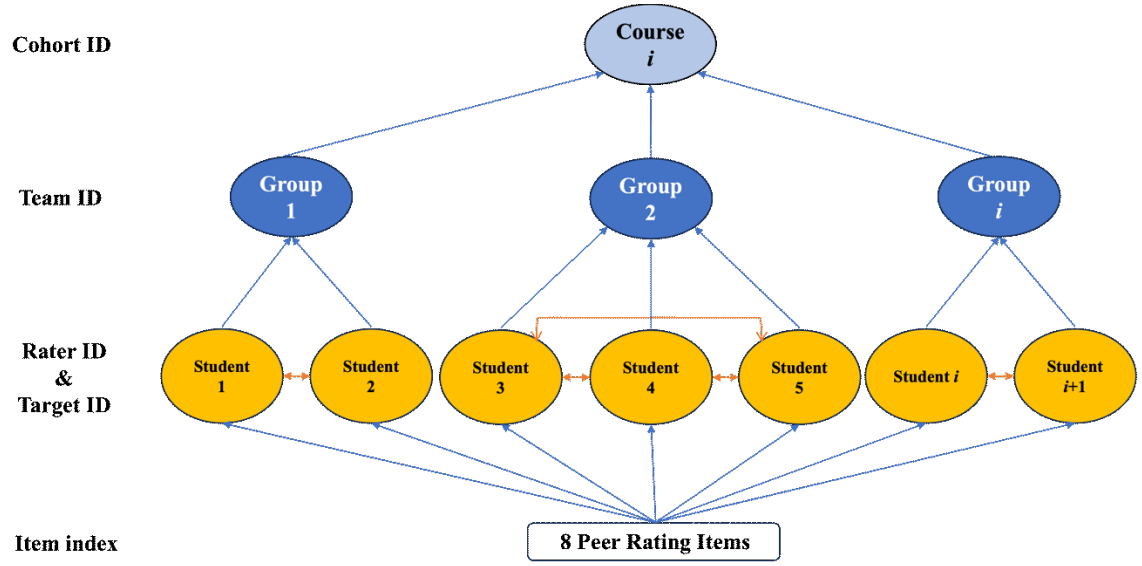

Figure 1. Data structure

Accordingly, 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. Peer rating items stacked together serve as the dependent variable, and the main factors include gender, cumulative GPA, personality traits (extraversion and task control), and group preference. Additionally, interactions between these factors were considered.

Results

Descriptive Statistics

The descriptive statistics provide insights into the participants' characteristics and perceptions in the study. Cumulative GPA, a measure of academic performance, shows a mean of 3.63 (SD = 0.350) out of 4.00, indicating that participants generally achieved high levels of achievement. Personality traits such as Extraversion and Task control, which were rated on a 7-point scale, reflect the participants' tendencies in group settings. The mean of 4.52 (SD = 1.418) for Extraversion indicates a propensity to actively contribute in groups, while the mean of 3.69 (SD = 1.442) for Task control suggests a balanced approach to task delegation. The mean of 7.60 on a 9-point scale (SD = 1.52) indicates positive perceptions of team members' contributions.

Table 3. Descriptive statistics.

Variables	*n*	Mean	Standard Deviation	Skewness	Kurtosis
Cumulative GPA	5,322	3.63	0.35	-2.47	16.88
Personality traits					
Extraversion	5,322	4.52	1.42	-0.36	2.20
Task Control	5,322	3.69	1.44	0.12	2.15
Peer rating items stacked*	373,816	7.60	1.52	-1.53	5.76

*Note: *All eight peer rating items were aggregated as the independent variable.*

Results of Multilevel Linear Model

The complete results of multilevel models are presented in Appendix 2. The fixed-effects estimates indicate the impact of predictors and the interactions between them on peer rating values. Furthermore, the random effects refer to the variability of peer rating values at different levels, accounting for differences between individual peer rating items, students, teams, and courses that are not explained by the fixed effects in the model.

Table 4. Main effects

					95% confidence interval	
Independent variables	Estimates*	Standard error	z	*p*	Lower	Upper
Cumulative GPA	0.694	0.009	77.820	<0.001	0.677	0.712
Rater Gender						
Male	-0.171	0.006	-30.930	<0.001	-0.182	-0.160
Work Preference						
Work with one partner	-0.008	0.008	-1.090	0.274	-0.023	0.007
Work in a group	-0.048	0.007	-6.510	<0.001	-0.062	-0.033
Personality traits						
Extraversion	0.053	0.002	27.050	<0.001	0.049	0.057
Task Control	0.017	0.002	8.420	<0.001	0.013	0.020

*Note: *Reference level for gender and group preference: Female for Gender and Work alone for Group preference.*

Table 4 indicates the main effects of the predictors. The positive estimate for Cummulative GPA (0.694, $p < 0.001$) suggests that higher academic performance was associated with a higher peer rating given by teammates, holding other factors constant. Similarly, students' personality traits (such as Extraversion or Task control) were positively associated with their peer ratings assigned

by teammates, indicating that individuals exhibiting higher levels of extraversion or exercising more control over tasks were associated with a higher mean peer rating. However, the negative coefficient (-0.197, $p < 0.001$; see Appendix 2) for the interaction within personality traits suggests that, on average, individuals exhibiting higher levels of Extraversion and simultaneously exercising more control over tasks were associated with a lower mean peer rating. In addition, the negative estimate for gender suggests that male students were assigned 0.171 ($p < 0.001$) lower peer rating means compared with their female teammates. Students with different preferences for group working were assigned different peer rating means, with students who preferred working alone having slightly higher peer rating means.

Most coefficients representing two-way and multiple-way interactions are statistically significant, with the exception of the two-way interaction between gender and cumulative GPA, the two-way interaction between gender and extraversion, and the three-way interaction among the three factors. The negative coefficient for the interaction between gender and group preference indicated that male students preferring working in a group were assigned the lowest peer rating values compared to their teammates (Figure 2), on average. Similarly, the negative coefficients (-0.315 and -0.366) for the interaction among gender, group preference, and personality traits (Figure 3.1 and 3.2) suggest a persistence of the pattern in Figure 2. However, the coefficients of a three-way interaction among gender, group preference and cumulative GPA are positive. This implies that, although the pattern shown in Figure 2 is extended by accounting for the effect of cumulative GPA in the three-way interaction (Figure 4), students with higher cumulative GPA were more likely to have higher means regardless of their gender and group preference (Figure 4.1 versus 4.2).

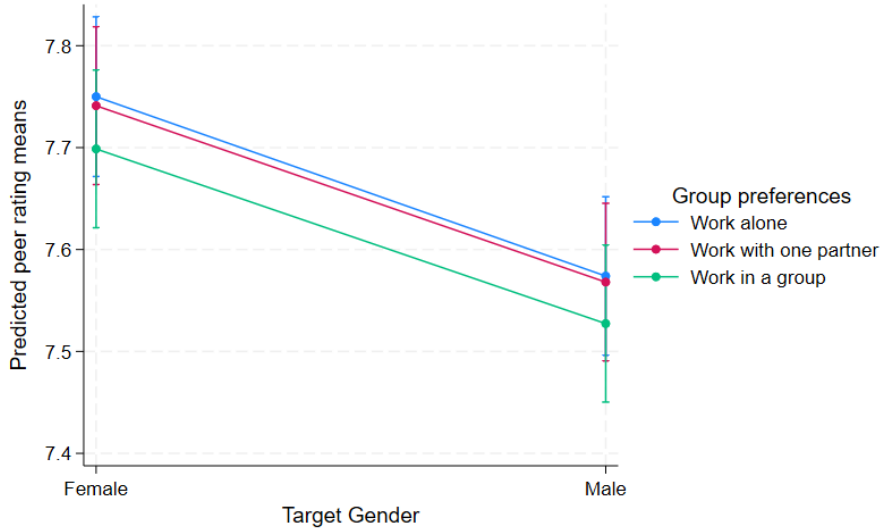

Figure 2. Peer rating means by a function of target gender and group preferences

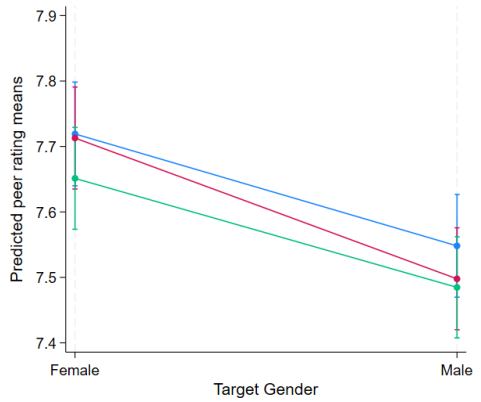

Figure 3.1. Peer rating means at Extraversion of 4 and Task control of 3

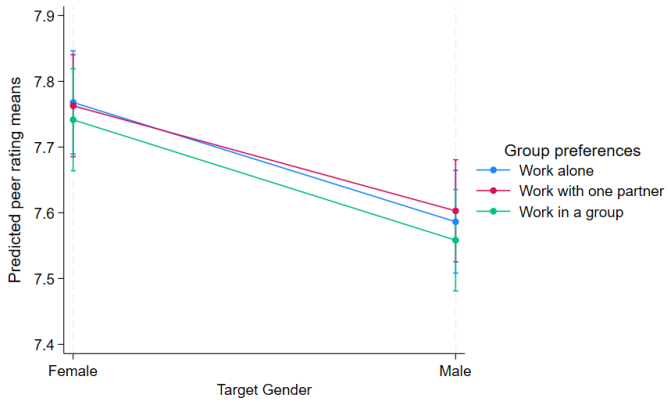

Figure 3.2. Peer rating means at Extraversion of 5 and Task control of 4

Figure 3. Four-way interaction among target gender, group preferences, and personality traits on peer rating means

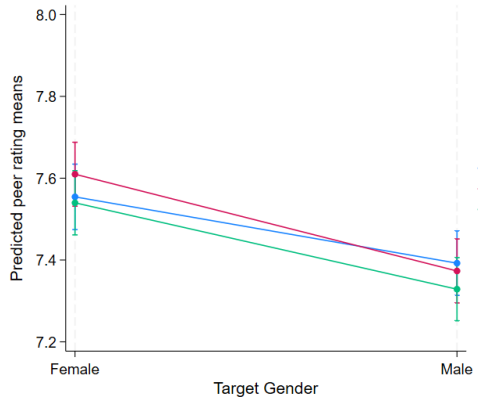

Figure 4.1. Peer rating means at cumulative GPA of 3.4

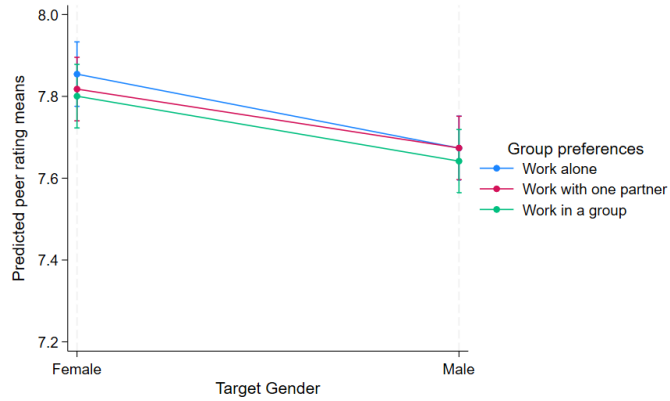

Figure 4.2. Peer rating means at cumulative GPA of 3.8

Figure 4. Three-way interaction among target gender, group preferences, and cumulative GPA

While the interaction between Cumulative GPA and Group preference, reflected by the negative coefficients (-0.709 for working with one partner and -2.002 for working in a group), introduced a distinctive pattern. For students who preferred working in a group, having a higher Cumulative GPA was paradoxically linked to a lower mean peer rating. This may suggest that, within the context of group work preference, academic performance alone did not consistently predict higher peer assessments. Moreover, this effect was influenced by the interplay with personality traits, as evidenced by interaction terms with estimates equal to -0.059 ($p < 0.001$) and -0.181 ($p < 0.001$).

Interestingly, the three-way interactions involving personality traits and group preference (0.223 for students with a preference for working with one partner and 0.650 for those liking working in a group) were positively associated with the peer rating mean value. This implied that when students preferred working with others, actively speaking up, and taking on task control, they were more likely to receive a higher mean peer assessment.

Overall, our analysis revealed statistically significant influences of academic performance and demographic factors. Students with higher cummulative GPAs were assigned with greater average peer rating values compared to their peers. Students who rated themselves higher in extraversion and preference for task control were more likely to receive lower ratings, on average. In addition, on average, students who indicated a preference for working in groups received lower peer rating values than those who preferred individual work or working with a partner.

Summary

Our results emphasize the importance of considering the effects of gender, academic performance, group preference, and personality traits in the design of peer assessment for evaluating team-based learning outcomes. Moreover, when conducting research aimed at examining team-based learning outcomes and related factors (e.g., designed interventions), we should also account for the effects of personal factors typically overlooked, such as the varying experiences associated with gender, academic performance, personality traits, and group preferences.

Based on the preliminary findings of this work-in-progress study, future research may delve into investigating the influence of personal characteristics on peer assessment among college engineering students as well as disaggregating peer assessment items. Importantly, in our dataset, in comparison to students from other majors (such as Business), a higher percentage of engineering students are male, exhibit preferences for working in a group, possess lower cumulative GPAs, and tend to self-score lower in personality traits (See Appendix 3).

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] A. Planas-Lladó, L. Feliu, F. Castro, R. M. Fraguell, G. Arbat, J. Pujol, J. J. Suñol, and P. Daunis-i-Estadella, "Using peer assessment to evaluate teamwork from a multidisciplinary perspective," *Assessment & Evaluation in Higher Education*, vol. 43, no. 1, pp. 14-30, 2018.

[4] D. Weaver, and A. Esposto, "Peer assessment as a method of improving student engagement," *Assessment & Evaluation in Higher Education*, vol. 37, no. 7, pp. 805-816, Nov. 2012.

[5] J. S. Kane, and E. E. Lawler, "Methods of peer assessment," *Psychological bulletin*, vol. 85, no. 3, pp. 555- 586, 1978.

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

[7] 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.

[8] N. A. Van Gennip, M. S. Segers, and H. H. Tillema, "Peer assessment as a collaborative learning activity: The role of interpersonal variables and conceptions," *Learning and Instruction*, vol. 20, no. 4, pp. 280-290, 2010.

[9] L. Wittie, J. Bennett, C. Merrill, J. Graham, and T. Schwab. "Bias in First-Year Engineering Student Peer Evaluations," In *2021 ASEE Virtual Annual Conference Content Access*, July 26, 2021.

[10] E. V. Rubin, and A. Edwards, "The performance of performance appraisal systems: understanding the linkage between appraisal structure and appraisal discrimination complaints," *The International Journal of Human Resource Management*, vol. 31, no. 15, pp. 1938-1957, 2020.

[11] N. Falchikov, and D. Magin. "Detecting gender bias in peer marking of students' group process work," *Assessment & Evaluation in Higher Education*, vol. 22, no. 4, pp. 385-396, 1997.

[12] 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*.

[13] R. Tucker, "Sex does not matter: gender bias and gender differences in peer assessments of contributions to group work," *Assessment & Evaluation in Higher Education*, vol. 39, no. 3, pp. 293-309, 2014.

[14] W. E., Watson, A. BarNir, and R. Pavur, "Elements influencing peer evaluation: An examination of individual characteristics, academic performance, and collaborative processes," *Journal of Applied Social Psychology*, vol. 40, no. 12, pp. 2995-3019, 2010.

[15] M. Bächtold, P. Roca, and K.D. Checchi, "Students' beliefs and attitudes towards cooperative learning, and their relationship to motivation and approach to learning," *Studies in Higher Education*, vol. 48, no. 1, pp. 100-112, 2023.

[16] R. E. Bryan, A. J. Krych, S. W. Carmichael, T. R. Viggiano, and W. Pawlina, "Assessing professionalism in early medical education: experience with peer evaluation and self-

evaluation in the gross anatomy course," *Annals-Academy of Medicine Singapore*, vol. 34, no. 8, pp. 486-491, 2005.
[17] J. Ghorpade, and J. R. Lackritz. "Peer evaluation in the classroom: A check for sex and race/ethnicity effects," *Journal of Education for business*, vol. 76, no. 5, pp. 274-281, 2001.
[18] 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.
[19] J. L. Peugh, "A practical guide to multilevel modeling," *Journal of school psychology*, vol. 48, no. 1, pp. 85-112, 2010.
[20] C. A. Bester, and C. Hansen, "Identification of marginal effects in a nonparametric correlated random effects model," *Journal of Business & Economic Statistics*, vol. 27, no. 2, pp. 235-250, 2009.

Appendix

Appendix 1. Tandem teamwork survey items.

Survey items	Lower anchor	Upper anchor
Personality traits		
Extraversion	In groups, I tend to listen more than speak.	I often speak up in groups.
Task Control	I think it's good to share work, even if my team might finish tasks differently than me.	I'd rather pick up extra work so I know it's done right.
Peer assessment		
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.

Appendix 2. Original output of the multilevel model

Mixed-effects ML regression — Number of obs = 373,816

Grouping information

Group variable	No. of groups	Observations per group Minimum	Average	Maximum
Cohort_id	58	48	6,445.1	37,984
team_id	1,572	8	237.8	768
Rater_stud~d	6,901	3	54.2	160
Rateeevalu~d	24,449	1	15.3	32

Log likelihood = -575827.27 — Wald chi2(47) = 10774.88; Prob > chi2 = 0.0000

Peer_rating_values	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
Ratee_Gender_CD						
Male	2.006821	1.154017	1.74	0.082	-.2550112	4.268653
Ratee_cumulative_GPA	1.219484	.2549835	4.78	0.000	.7197253	1.719242
Ratee_Gender_CD# c.Ratee_cumulative_GPA						
Male	-.5615223	.3187367	-1.76	0.078	-1.186235	.0631902
Ratee_GroupPreference						
Work with one partner	2.70843	1.187805	2.28	0.023	.3803758	5.036485
Work in a group	6.962192	1.201771	5.79	0.000	4.606764	9.317619
Ratee_Gender_CD# Ratee_GroupPreference						
Male#Work with one partner	-3.592622	1.497348	-2.40	0.016	-6.527369	-.657875
Male#Work in a group	-3.906728	1.47578	-2.65	0.008	-6.799204	-1.014252
Ratee_GroupPreference# c.Ratee_cumulative_GPA						
Work with one partner	-.7091773	.3283251	-2.16	0.031	-1.352683	-.0656719
Work in a group	-2.002145	.33254	-6.02	0.000	-2.653911	-1.350379
Ratee_Gender_CD# Ratee_GroupPreference# c.Ratee_cumulative_GPA						
Male#Work with one partner	.7789944	.4136964	1.88	0.060	-.0318356	1.589824
Male#Work in a group	1.123975	.4079599	2.76	0.006	.3243887	1.923562
Ratee_Extraversion	.6256894	.1963033	3.19	0.001	.240942	1.010437
Ratee_Gender_CD# c.Ratee_Extraversion						
Male	-.369159	.2502368	-1.48	0.140	-.8596141	.1212961
c.Ratee_cumulative_GPA# c.Ratee_Extraversion	-.1585259	.0539752	-2.94	0.003	-.2643154	-.0527364
Ratee_Gender_CD# c.Ratee_cumulative_GPA# c.Ratee_Extraversion						
Male	.0883078	.0686806	1.29	0.199	-.0463037	.2229193
Ratee_GroupPreference# c.Ratee_Extraversion						
Work with one partner	-.6325722	.2522156	-2.51	0.012	-1.126906	-.1382387
Work in a group	-1.666221	.2542164	-6.55	0.000	-2.164476	-1.167966
Ratee_Gender_CD# Ratee_GroupPreference# c.Ratee_Extraversion						
Male#Work with one partner	.2560187	.3235888	0.79	0.429	-.3782036	.890241
Male#Work in a group	.9190266	.3173343	2.90	0.004	.2970628	1.54099
Ratee_GroupPreference# c.Ratee_cumulative_GPA# c.Ratee_Extraversion						
Work with one partner	.1611303	.0694444	2.32	0.020	.0250217	.2972388
Work in a group	.4667516	.0699719	6.67	0.000	.3296091	.6038941

Ratee_Gender_CD# Ratee_GroupPreference# c.Ratee_cumulative_GPA# c.Ratee_Extraversion						
Male#Work with one partner	-.0296482	.0890238	-0.33	0.739	-.2041316	.1448352
Male#Work in a group	-.2562142	.0872419	-2.94	0.003	-.4272051	-.0852233
Ratee_Control	.6490467	.1981253	3.28	0.001	.2607282	1.037365
Ratee_Gender_CD#c.Ratee_Control						
Male	-.8790897	.2586322	-3.40	0.001	-1.385999	-.3721799
c.Ratee_cumulative_GPA# c.Ratee_Control	-.1857369	.0547473	-3.39	0.001	-.2930397	-.0784341
Ratee_Gender_CD# c.Ratee_cumulative_GPA# c.Ratee_Control						
Male	.2416703	.0710547	3.40	0.001	.1024057	.3809348
Ratee_GroupPreference# c.Ratee_Control						
Work with one partner	-.7534808	.2704119	-2.79	0.005	-1.283478	-.2234832
Work in a group	-2.743265	.2870784	-9.56	0.000	-3.305928	-2.180602
Ratee_Gender_CD# Ratee_GroupPreference# c.Ratee_Control						
Male#Work with one partner	1.833288	.3579163	5.12	0.000	1.131785	2.534791
Male#Work in a group	1.437802	.3586161	4.01	0.000	.7349272	2.140677
Ratee_GroupPreference# c.Ratee_cumulative_GPA# c.Ratee_Control						
Work with one partner	.2033605	.0743823	2.73	0.006	.057574	.349147
Work in a group	.773634	.0791491	9.77	0.000	.6185046	.9287634
Ratee_Gender_CD# Ratee_GroupPreference# c.Ratee_cumulative_GPA# c.Ratee_Control						
Male#Work with one partner	-.4590122	.0980931	-4.68	0.000	-.6512712	-.2667532
Male#Work in a group	-.4197089	.098668	-4.25	0.000	-.6130945	-.2263232
c.Ratee_Extraversion# c.Ratee_Control	-.1972938	.04084	-4.83	0.000	-.2773387	-.1172488
Ratee_Gender_CD# c.Ratee_Extraversion# c.Ratee_Control						
Male	.171232	.055286	3.10	0.002	.0628735	.2795905
c.Ratee_cumulative_GPA# c.Ratee_Extraversion# c.Ratee_Control	.0550867	.0112089	4.91	0.000	.0331176	.0770558
Ratee_Gender_CD# c.Ratee_cumulative_GPA# c.Ratee_Extraversion# c.Ratee_Control						
Male	-.045767	.0150695	-3.04	0.002	-.0753026	-.0162314
Ratee_GroupPreference# c.Ratee_Extraversion# c.Ratee_Control						
Work with one partner	.222548	.0556673	4.00	0.000	.1134421	.331654
Work in a group	.6501011	.0594815	10.93	0.000	.5335196	.7666827
Ratee_Gender_CD# Ratee_GroupPreference# c.Ratee_Extraversion# c.Ratee_Control						
Male#Work with one partner	-.3146503	.0757769	-4.15	0.000	-.4631703	-.1661304
Male#Work in a group	-.3663546	.0758203	-4.83	0.000	-.5149596	-.2177496
Ratee_GroupPreference# c.Ratee_cumulative_GPA# c.Ratee_Extraversion# c.Ratee_Control						
Work with one partner	-.0590121	.0152675	-3.87	0.000	-.0889358	-.0290883
Work in a group	-.1807012	.0163112	-11.08	0.000	-.2126705	-.1487319

Ratee_Gender_CD# Ratee_GroupPreference# c.Ratee_cumulative_GPA# c.Ratee_Extraversion# c.Ratee_Control						
Male#Work with one partner	.0779742	.0206831	3.77	0.000	.037436	.1185124
Male#Work in a group	.1039187	.0207324	5.01	0.000	.063284	.1445534
_cons	3.118683	.9223068	3.38	0.001	1.310995	4.926371

Random-effects parameters	Estimate	Std. err.	[95% conf.	interval]
Cohort_id: Identity				
var(_cons)	.0612484	.0171842	.0353408	.1061481
team_id: Identity				
var(_cons)	.2109408	.0131136	.1867428	.2382745
Rater_stud~d: Unstructured				
var(2.Time_id)	.6185632	.0122444	.5950242	.6430334
var(_cons)	.6738184	.0140367	.6468611	.701899
cov(2.Time_id,_cons)	-.3098408	.0105079	-.330436	-.2892456
Rateeevalu~d: Identity				
var(R.Peer_rating_Index)	.6782167	.0036001	.6711973	.6853096
var(Residual)	.6599921	.0022512	.6555947	.6644191

LR test vs. linear model: chi2(6) = 2.1e+05 Prob > chi2 = 0.0000

. estat ic

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	373,816	.	-575827.3	55	1151765	1152360

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

Appendix 3.

```
. tabulate Ratee_first_major_category Ratee_Gender_CD, row
```

Key
frequency *row percentage*

Ratee First Major Category	Ratee Gender		
	Female	Male	Total
Engineering	556	1,241	1,797
	30.94	69.06	100.00
Sciences	185	140	325
	56.92	43.08	100.00
Multidisciplinary	624	407	1,031
	60.52	39.48	100.00
Arts	83	83	166
	50.00	50.00	100.00
Business	843	1,160	2,003
	42.09	57.91	100.00
Total	2,291	3,031	5,322
	43.05	56.95	100.00

```
. tabulate Ratee_first_major_category Ratee_GroupPreference , row
```

Key
frequency *row percentage*

Ratee First Major Category	Ratee Group Preference			
	Work alon	Work with	Work in a	Total
Engineering	391	561	845	1,797
	21.76	31.22	47.02	100.00
Sciences	101	92	132	325
	31.08	28.31	40.62	100.00
Multidisciplinary	249	335	447	1,031
	24.15	32.49	43.36	100.00
Arts	57	46	63	166
	34.34	27.71	37.95	100.00
Business	465	731	807	2,003
	23.22	36.50	40.29	100.00
Total	1,263	1,765	2,294	5,322
	23.73	33.16	43.10	100.00

```
. regress Ratee_Extraversion i.Ratee_major
```

Source	SS	df	MS		
Model	119.896717	3	39.9655725	Number of obs	= 5,322
Residual	10582.9075	5,318	1.99001645	F(3, 5318)	= 20.08
				Prob > F	= 0.0000
Total	10702.8042	5,321	2.01142721	R-squared	= 0.0112
				Adj R-squared	= 0.0106
				Root MSE	= 1.4107

Ratee_Extraversion	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
Ratee_major						
Science	-.132746	.0850326	-1.56	0.119	-.2994447	.0339527
Arts and Multidisciplinary	.0816412	.05263	1.55	0.121	-.0215351	.1848175
Business	.3092418	.0458359	6.75	0.000	.2193846	.3990989
_cons	4.391208	.0332778	131.96	0.000	4.325969	4.456446

```
. regress Ratee_Control i.Ratee_major
```

Source	SS	df	MS		
Model	44.2510819	3	14.7503606	Number of obs	= 5,322
Residual	11014.4607	5,318	2.07116598	F(3, 5318)	= 7.12
				Prob > F	= 0.0001
Total	11058.7118	5,321	2.07831456	R-squared	= 0.0040
				Adj R-squared	= 0.0034
				Root MSE	= 1.4392

Ratee_Control	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
Ratee_major						
Science	-.001166	.086749	-0.01	0.989	-.1712296	.1688976
Arts and Multidisciplinary	-.0203653	.0536923	-0.38	0.704	-.1256243	.0848937
Business	.1800913	.0467611	3.85	0.000	.0884204	.2717622
_cons	3.622705	.0339495	106.71	0.000	3.55615	3.689259

```
. regress Ratee_cumulative_GPA i.Ratee_major
```

Source	SS	df	MS		
Model	8.95094163	3	2.98364721	Number of obs	= 5,322
Residual	642.366048	5,318	.120790908	F(3, 5318)	= 24.70
				Prob > F	= 0.0000
Total	651.31699	5,321	.122404997	R-squared	= 0.0137
				Adj R-squared	= 0.0132
				Root MSE	= .34755

Ratee_cumulative_GPA	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
Ratee_major						
Science	.0464357	.0209495	2.22	0.027	.005366	.0875053
Arts and Multidisciplinary	.0844204	.0129665	6.51	0.000	.0590008	.10984
Business	.0900291	.0112926	7.97	0.000	.067891	.1121673
_cons	3.576872	.0081987	436.27	0.000	3.560799	3.592945