The Future of Engineering Education
2024 Annual Conference & Exposition
Oregon Convention Center
Portland, OR . June 23 - 26, 2024
ASEE
Paper ID #42675

Analyzing Patterns of Pre-Semester Concerns in First-Year Engineering Students

Mr. Jeong Hin Chin, University of Michigan

Jeong Hin Chin is a Master's student at the University of Michigan School of Information. Mr. Chin graduated with a triple major in Honors Data Science, Honors Asian Studies (Chinese), and Statistics from the College of Literature, Science, and the Arts (LSA) at the University of Michigan. He is interested in deep learning, machine learning, unsupervised clustering methods, and human-computer interaction in various disciplines such as Public Health, Psychometric, Engineering Education, Epidemiology, and Humanities.

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.

Christopher Brooks, University of Michigan

Analyzing Patterns of Pre-Semester Concerns in First-Year Engineering Students

Abstract

This complete research paper investigates the relationship between pre-semester concerns and the experiences of students participating in team-based pedagogical environments within first-year engineering courses at a large public university using data from an institutional database and Tandem, a team support tool. The analysis focuses on various demographic factors, including gender, ethnicity, and nationality, as well as seasonal variation. Key findings reveal that issues such as ideas *being heard* by others and taking up a *higher share of the workload to make up* for my teammates are among students' primary concerns. The study suggests implementing various solutions to enhance communication skills, improve teamwork outcomes, and support positive team dynamics.

Introduction

The introduction of Team-based learning (TBL) in the 1980s marked a significant shift in addressing the challenges of large class settings in educational environments [1], [2]. Originally a business school innovation, TBL has now permeated various disciplines including engineering, medicine, and social sciences globally. Some courses, such as first-year engineering, may combine TBL with project-based learning (PBL) to introduce students to common engineering themes such as design, sustainability, and ethics. Despite its wide-ranging benefits, TBL's effectiveness can be inequitable for a variety of reasons, including free riders, imbalances in task allocation, and more broad communication issues [3], [4]. Thus, the application of teamwork assessment and support tools, such as CATME and Tandem, is essential in monitoring and improving student teamwork experience through insight obtained from various analyses [5], [6].

This study leverages data from an institutional database and Tandem. Tandem is a tool designed for assessing team dynamics and providing formative feedback. Since Tandem's first implementation in 2019, it has collected responses from more than 13,000 students. For this study, only data from the "beginning-of-term" survey (BoT) in Tandem and the responses to one question in the Midterm survey were studied. In this study, we explored patterns and types of pre-semester concerns reported by first-year engineering students, investigating the following four research questions (RQs):

- RQ1: What are the patterns and perceptions of teamwork in TBL settings, analyzed through the lenses of gender, ethnicity, nationality, and semester-wise variations?
- RQ2: How do team experiences differ among students with positive versus negative overall team experiences, again considering factors such as gender, ethnicity, nationality, and season?
- RQ3: Among students with negative team experiences, what distinct clusters can be identified, and what are their characteristics?
- RQ4: What recurring factors or variables significantly influence these experiences, and what strategies can be implemented to address them?

This study focuses on analyses of survey data to uncover patterns and trends within team experiences related to pre-semester concerns. This method enables detailed exploration of the multifaceted aspects of team dynamics within team-based pedagogies settings, particularly as they relate to students' initial apprehensions and expectations.

Data

This research centers on first-year engineering students at a large, research-focused public university operating on a semester system. The primary data source for this study is derived from student responses to the Tandem surveys. These responses were gathered from participants in 15 distinct sections of a foundational engineering design course which was run by different faculty at different times over the Fall of 2019 to the Fall of 2023 period, covering a total of nine semesters. The unique circumstances of the COVID-19 pandemic meant that the courses in the Fall of 2020 and Spring of 2021 (Winter semester in the university's registration) were either entirely online or blended. Despite these variations, all course iterations maintained their emphasis on team-based pedagogies. For our analysis, we specifically examined data from the BoT survey and the responses to one question in the Midterm survey. The BoT survey, implemented at the semester's start, captured initial student attitudes and apprehensions about teamwork (refer to Fig. 1).

Beginning-of-Term (BoT) Survey

The BoT survey is administered at the outset of the semester before students interact with their designated team members in the course. This survey probes into individual traits that are considered significant in the context of teamwork, as identified in existing literature [5], [6]. These traits include but are not limited to, personality types, past experiences in team settings, and preferences related to teamwork. The survey items are grounded in well-established scales from teamwork literature. However, for brevity and based on feedback from initial user testing by Tandem's developers, these items are often condensed into single or occasionally double-barreled questions [7].

This analysis includes 1103 responses from the BoT survey. We focused on eight specific questions from the BoT. Of these questions, five utilized a seven-point Likert scale, allowing students to express their levels of *extraversion*, *procrastination*, sense of *belongingness* in the class, *desire to pick up more work* than others, and willingness to *speak up*. The other three questions, *experience* with group work, *attitude* toward group work, and *preference* for working with groups, required students to select a single response from the provided options. These particular questions were selected for their relevance in representing key aspects of a student's personality. A final question asked students to select from a list their concerns for teamwork before students interact with their designated team members in the course. Figure 1 illustrates the eight survey questions and the corresponding response options that were included in the analysis while Figure 2 shows the different concerns that the students can choose from in the survey.

Where would you place yourself on the following scales? [7 stops on the scale]			
[Extraversion]	In groups, I tend to listen more than speak.	←→	I often speak up in groups.
[Procrastination]	I usually do work close to a deadline.	←→	I get working on a project as soon as it is assigned.
[BT_Belongingness]	I expect to fit right into $Course.	←→	I expect to feel pretty out of place in $Course.
[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.
[SpeakUp]	I'd rather hold back ideas or preferences if my group stays happy.	←→	It's easy for me to speak up about my ideas or preferences even if it disrupts my group.

Where would you place yourself on the following scales? [categorical (single answer)]				
	Not at all	Once or Twice	Several Times	Many Times
[BT_PastGroups] Working with a team				

Where would you place yourself on the following scales? [scategorical (single answer)]					
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
[BT_PastPositive] My past teamwork experiences were generally positive.					

Where would you place yourself on the following scales? [categorical (single answer)]			
	[alone] Work alone	[partner] Work with one partner	[group] Work in a group
[GroupPreference] If given an option, I'd prefer to			

Fig. 1. Survey questions and answer choices asked in BoT. "$Course" is replaced by text describing the course (or sometimes, non-course context).

Midterm Survey

The Midterm (MT) Survey, a tool for both peer and self-evaluation, is typically administered at one or more key junctures during the academic term. In the context of this study, the MT Survey is conducted only once per semester. Our focus is particularly on a single variable from this survey, which is how students rate their experiences on their team until the point of the survey. The five choices for this question are "Positive", "Somewhat positive", "Neutral", "Somewhat Negative", and "Negative" (refer to Fig. 3).

What concerns you about the teamwork in $Course ? Select all that apply. *

- [ ] I think it will be hard to find time to work together. [Logistics]
- [ ] I'm nervous about getting to know new people. [Social]
- [ ] I think I'd do better work on my own. [BetterSolo]
- [ ] I don't want my grade to be affected by someone else. [GradeHarm]
- [ ] I don't want my performance to hurt my teammates' grades. [PeerHarm]
- [ ] I'm worried I'll have to do more than my share of work to make up for my teammates. [ExtraWork]
- [ ] I'm worried about my ideas being heard by the group. [BeingHeard]
- [ ] I worry some team members won't be able to do the quality of work I'd like. [WorkQuality]
- [ ] I don't want others to think I'm not smart if I make a mistake. [SelfDoubt]
- [ ] I don't have any concerns.

Fig. 2. The concerns for teamwork that students can choose from in the BoT survey. "$Course" is replaced by text describing the course (or sometimes, non-course context). The concerns are categorized and encoded with specific variable names (in brackets) to facilitate data analysis.

[MT_Overall] How would you describe your experience on your team so far?

Very positive
Somewhat positive
Neither positive nor negative
Somewhat negative
Very negative

Fig. 3. The MT_Overall variable and its answer choices in the midterm survey.

Methods

Statistical Analysis

To evaluate the data collected through the BoT surveys, we implemented a series of statistical analyses aimed at addressing the sub-questions outlined in the Introduction section. Descriptive statistics were used to provide an overview of the demographics and baseline characteristics of the study population. Cross-tabulations and frequency distributions were employed to summarize the categorical data, which will be discussed in the Results section. The investigation into team experiences considered multiple variables from both the Tandem data and an institutional enrollment dataset. As Tandem data is primarily indicated on a Likert scale with numerical values assigned to it, we will treat them as ordinal variables. Nonetheless, while calculating the statistics for each variable, we will convert these variables into numerical values and treat them as continuous [8]. Institutional data included items treated categorically, such as gender (gender is our characteristic of interest though institutional data provides sex), ethnicity, nationality, and which semester the course was taken. To analyze these variables, both t-tests and chi-squared tests were conducted to detect significant differences in the distribution of survey responses among various student groups through multiple pairwise comparisons.

Before delving into the results, it is essential to note the limitations of the data and the methods used. While the large sample size and the comprehensive nature of the surveys provide robust data for analysis, the observational design of the study limits the ability to infer causality from the correlations observed. Our institutional database provides “sex” as a categorical variable even though we are interested in gender. Ethnicity and nationality in the dataset are captured from the U.S.-centric lens, and are not fully inclusive of the range of ethnic identities individuals may associate with.

Hierarchical Cluster Analysis

For a more refined analysis of our dataset, especially to discern patterns among students with less favorable team experiences, we employed hierarchical clustering with the Ward.D2 method. The dendrogram presented in the Results section is a visual outcome of this analytical approach. Ward.D2 is an agglomerative clustering method that minimizes the total within-cluster variance [9]. At each step, it finds the pair of clusters that leads to a minimum increase in total within-cluster variance after merging. The choice of Ward.D2 was deliberate; by evaluating squared Euclidean distances between clusters, Ward.D2 helps to ensure that the groups formed are cohesive internally, yet well-separated from each other.

In addition, Ward’s method looks for clusters in multivariate Euclidean space [9], which is particularly pertinent to educational data as it often involves multiple interrelated variables that represent complex student behaviors and characteristics [10]. By considering the squared Euclidean distance between points, this method effectively captures the true 'distance' or dissimilarity between different student responses. This is especially useful in an educational setting where multidimensional data points—comprising students' attitudes, skills, and experiences—need to be grouped into coherent clusters that can inform targeted educational interventions and support mechanisms.

Statistical Software

Data manipulation, visualization, and statistical analyses were performed using R statistical software. The 'ggplot2' and 'plotly' packages were utilized for generating the figures that visually summarize the survey data, while 'dplyr' and 'tidyr' were used for data wrangling. The 'corrplot' package facilitated the correlation matrix visualization, and the 'hclust' package was used for hierarchical clustering.

Results

RQ1: What are the patterns and perceptions of teamwork in TBL settings, analyzed through the lenses of gender, nationality, ethnicity, and season-wise variations?

Gender

The analysis for Sub-question 1, which considers student concerns through the lenses of gender, ethnicity, nationality, and season, indicates distinct patterns. The results for the various Chi-squared tests can be found in the Appendix. Figure 4 shows that, compared to male students, female students reported higher levels of concern about *self-doubting* themselves when they make a mistake (SelfDoubt), ideas *being heard* by others (BeingHeard), and their *grades being hurt* by others (GradeHarm). This trend highlights communication and power dynamics in team settings that require further consideration. The pattern observed among female students in engineering teams aligns with findings from other studies, which report that female students often struggle to have their ideas acknowledged [3]. This phenomenon underscores the need for strategies to enhance inclusivity and equity in team collaborations within engineering education.

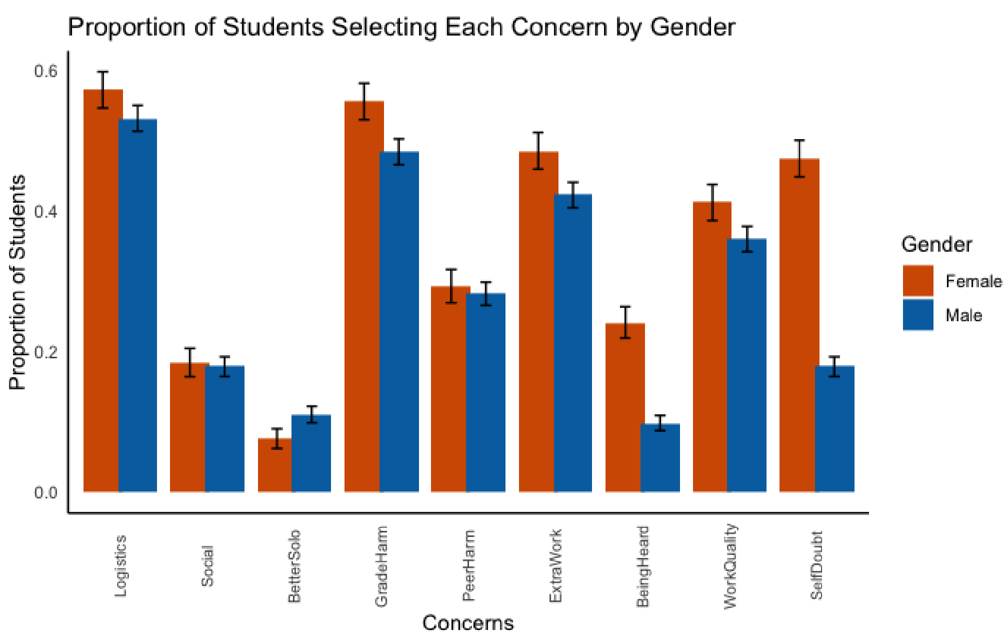

Figure 4. The proportion of students (with standard error) selecting each concern by gender.

Nationality

In assessing the association between students' domestic/minority status and a range of concern variables, the Chi-Square test revealed that the concern about *getting to know new people* (Social) has a p-value less than or equal to 0.05 (Table A2). This suggests that there are significant differences in the distribution of responses among international, domestic-minority, and domestic-non-minority students for this variable. However, upon further investigation using the Holm method for posthoc analysis [11], which adjusts for multiple comparisons to control for Type I error, the results did not maintain statistical significance within individual group comparisons. Thus, while the initial Chi-Square test pointed to a potential difference in social concerns across student groups, the post-hoc analysis did not find significant differences when comparing these groups individually as shown in Table 1. No other differences were significant. This underscores the complexity of the relationships between student demographics and team concerns, necessitating a nuanced approach to interpreting and addressing these issues in educational settings.

Table 1: Post-Hoc Analysis of Chi-Squared Test on the *Social* Concerns by Nationality with Holm Correction Method.

	Social		
Dimension	**Value**	**0**	**1**
International	Residuals	-2.10	2.10
	p values	0.22	0.22
Minority	Residuals	-0.77	0.77
	p values	1.00	1.00
Non-Minority	Residuals	1.94	-1.94
	p values	0.32	0.32

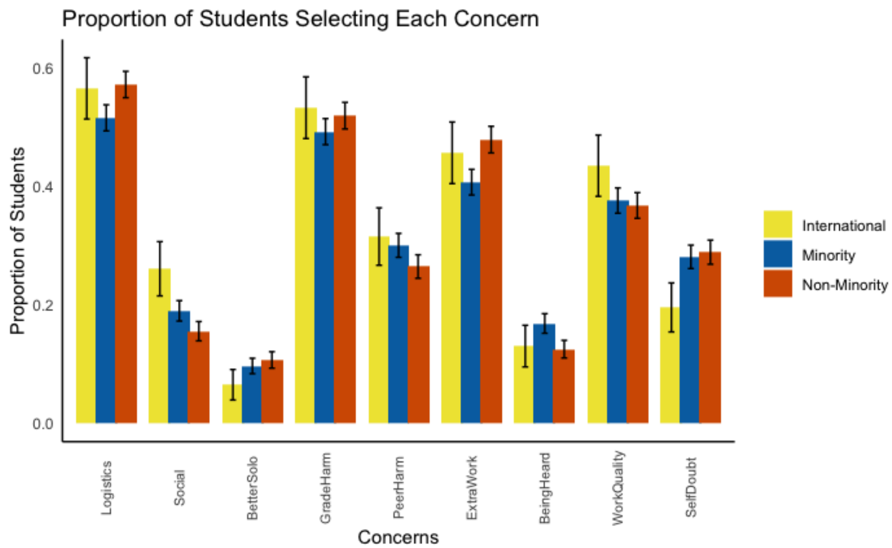

Figure 5. The proportion of students (with standard error) selecting each concern by nationality.

Race/Ethnicity

In assessing the association between students' race/ethnicity status and a range of concern variables, a Chi-squared test was conducted. The Chi-squared test (Table A3) highlighted teammates' *unsatisfactory quality of work* (WorkQuality), a *higher share of the workload to make up* for teammates (ExtraWork), and *self-doubting* themselves when they make a mistake (SelfDoubt) as significant areas of concern. To check the robustness of the result, a post hoc analysis using the Holm method was applied to further investigate the association of these concerns with students' ethnic backgrounds. The results of the post hoc analysis can be found in Table 2. This detailed analysis revealed that Black students demonstrated significantly fewer concerns regarding taking up a *higher share of the workload to make up* for my teammates (ExtraWork), with a Holm-adjusted p-value of 0.04. Similarly, Asian students exhibited a markedly lower level of concern in *self-doubting* themselves when they made a mistake (SelfDoubt), with a Holm-adjusted p-value of 0.03.

Table 2: Post-Hoc Analysis of Chi-Squared Test on the *work quality, extra work,* and *self-doubt* Concerns by Ethnicity with Holm Correction Method.

	Work Quality			**Extra Work**			**Self Doubt**		
Dimension	**Value**	**0**	**1**	**Value**	**0**	**1**	**Value**	**0**	**1**
2 or More	Residuals	-1.49	1.49	Residuals	0.38	-0.38	Residuals	-1.21	1.21
	p values	1.00	1.00	p values	1.00	1.00	p values	1.00	1.00
Asian	Residuals	-2.62	2.62	Residuals	-0.03	0.03	Residuals	3.03	-3.03
	p values	0.11	0.11	p values	1.00	1.00	p values	**0.03**	**0.03**
Black	Residuals	1.49	-1.49	Residuals	2.94	-2.94	Residuals	-2.41	2.41
	p values	1.00	1.00	p values	**0.04**	**0.04**	p values	0.19	0.19
Hispanic	Residuals	2.74	2.74	Residuals	1.60	-1.60	Residuals	-1.49	1.49
	p values	0.07	0.07	p values	1.00	1.00	p values	1.00	1.00
Not Indic	Residuals	1.08	-1.08	Residuals	-0.25	0.25	Residuals	1.67	-1.67
	p values	1.00	1.00	p values	1.00	1.00	p values	1.00	1.00
White	Residuals	0.47	-0.47	Residuals	-2.01	2.01	Residuals	-1.63	1.63
	p values	1.00	1.00	p values	0.53	0.53	p values	1.00	1.00

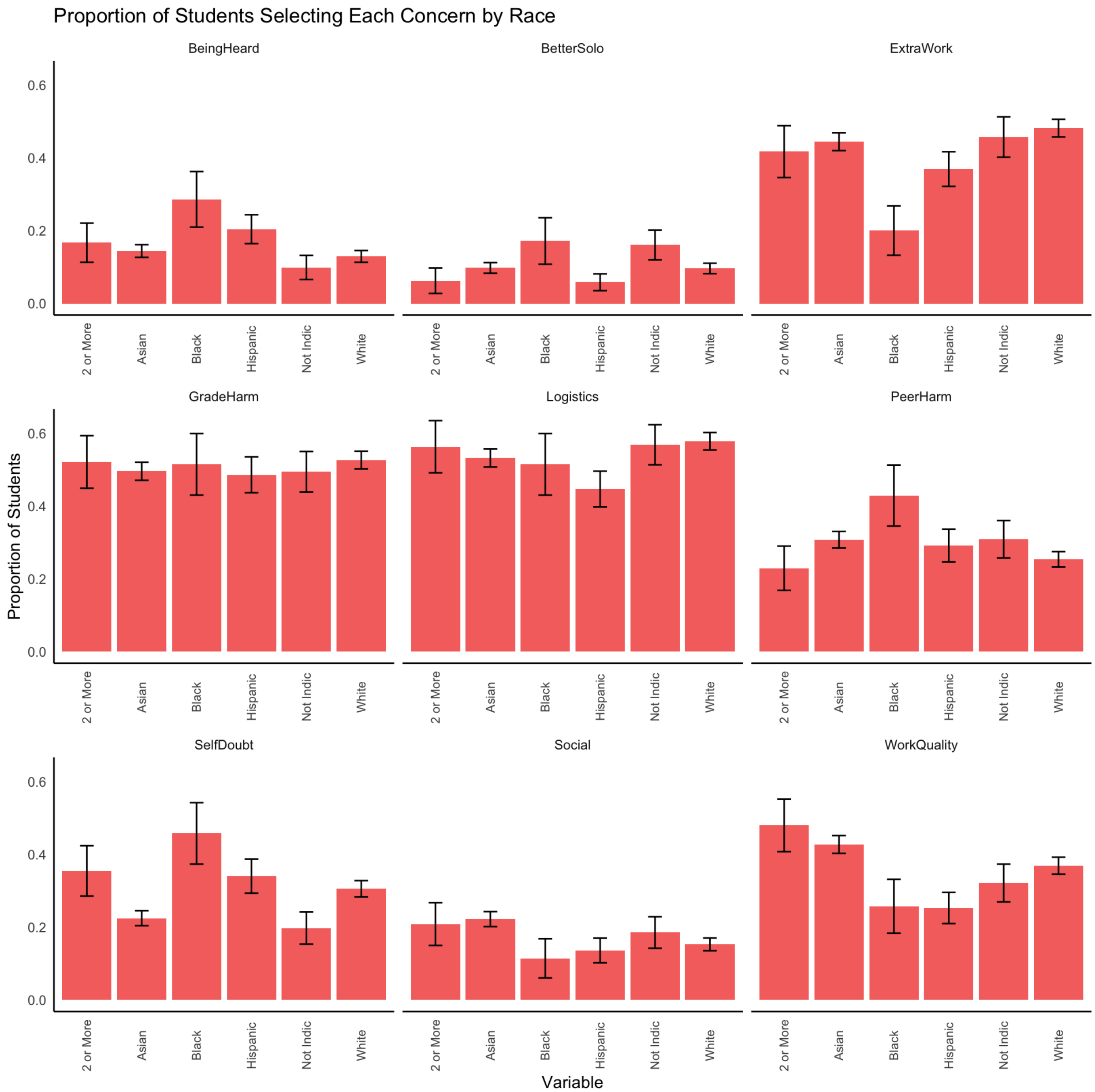

Figure 6. The proportion of students (with standard error) selecting each concern by race.

Semester of enrollment

Lastly, a Chi-squared test was conducted to assess the association between the semester of student enrollment and a range of concern variables. The test (Table A4) results indicated significant variations in students' concerns about taking a *higher share of the workload to make up* for teammates (ExtraWork) between the fall and winter seasons. This pattern suggests a potential link between students' workload perceptions and the academic calendar. Given that the data originated from first-year engineering courses at a research-focused public university, it is plausible that the increased concern in the fall season stemmed from freshmen being unfamiliar with the university's demanding workload. As these students transition into the winter season, their heightened awareness and experience of the academic rigor in engineering courses likely contribute to their increased perception of taking a *higher share of the workload to make up* for teammates (ExtraWork) as a significant concern. This observation suggests the need for faculty members to offer additional support, particularly in the fall season, to help incoming freshmen adjust to the challenging workload typical of engineering courses so that students can adapt more effectively to the rigorous academic environment and improve their overall educational experience.

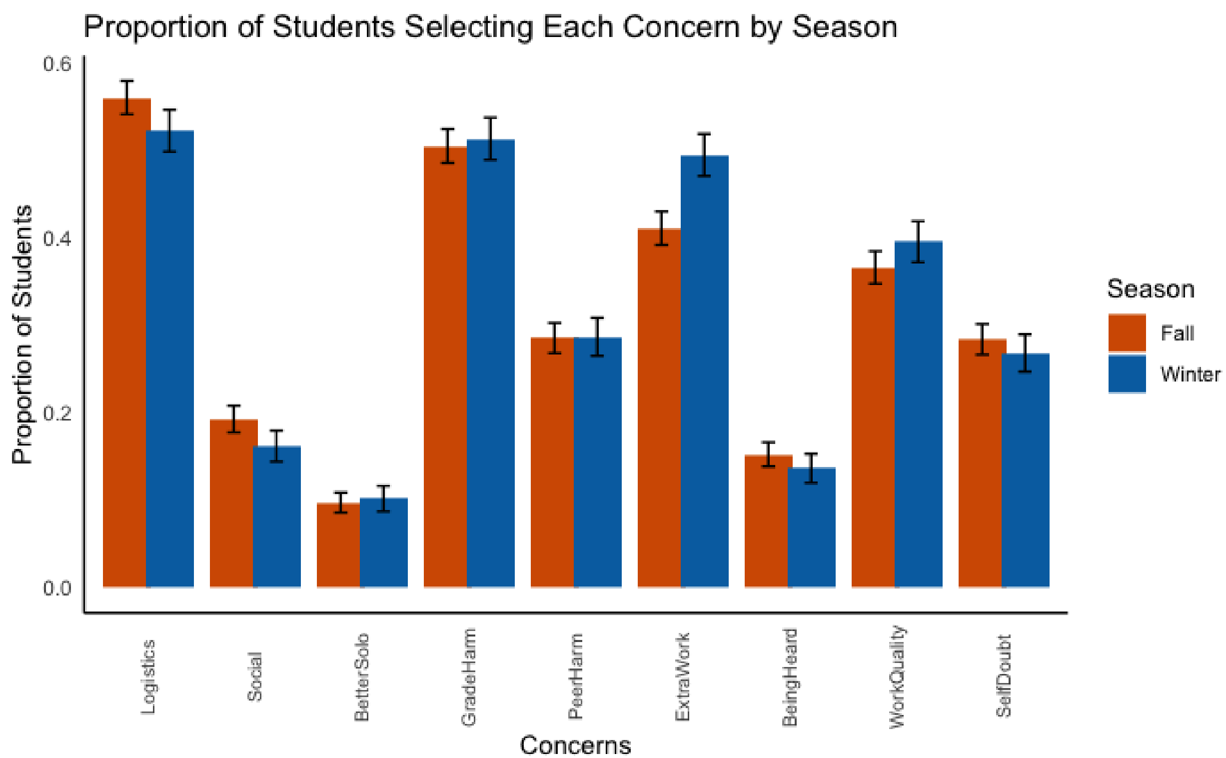

Figure 7. The proportion of students (with standard error) selecting each concern by season.

RQ2: How do these team experiences differ among students with positive versus negative overall team experiences, again considering factors such as gender, ethnicity, nationality, and season?

Figure 8 suggests that most students generally report that, before this specific course, they have had positive experiences within team-based learning settings, transcending gender, ethnicity, nationality, and season differences. For those with negative experiences, these factors do not significantly distinguish one group from another. This observation is supported by Chi-Square tests, which indicate that the distribution of negative experiences does not significantly differ across the various demographic categories examined. Such results may point to the effectiveness of the teamwork environment in providing a uniformly positive experience or may suggest that factors not captured by the data are influencing student perceptions. Further analysis could explore additional variables or qualitative feedback to understand the nuances behind the negative experiences.

Figure 8. Distribution of students with positive versus negative overall midterm team experiences across factors such as gender, ethnicity, nationality, and season.

RQ3: Among students with negative team experiences, what distinct clusters can be identified, and what are their characteristics?

Upon analyzing the hierarchical clustering dendrogram, we observed several distinct clusters among students reporting negative team experiences as shown in Figure 9. These clusters represent subgroups of students with shared characteristics in their perception of the team's dynamics. The dendrogram branches out into major clusters, differentiated by color, illustrating the levels at which groups of students are combined. In the context of this study, the height of the branches suggests the degree of difference between clusters – shorter branches represent closely related observations, while longer branches indicate greater dissimilarity. Cluster 1, 2, 3, and 4 have 36, 40, 6, and 11 students respectively. Additional information for each cluster can be found in Tables A5 to A7 in the Appendix. The cophenetic correlation coefficient was calculated to evaluate the quality of the clustering [12]. For the current clustering algorithm, the correlation coefficient is approximately 0.59, indicating a moderate to good fit between the cluster model and the original data.

The hierarchical clustering analysis reveals distinct patterns in team experience concerns among students, categorized into four clusters. Each cluster highlights specific variables with varying mean distances, indicative of their prominence and the consistency of responses within the cluster (within-cluster distance) compared to other clusters (between-cluster distance). From the plots, we are interested in variables that have a smaller within-cluster distance than the between-cluster distance and the confidence interval bar does not overlap as this indicates well-defined subgroup characteristics. In Clusters 1, 3, and 4, one of the variables that do not have an overlapping confidence interval bar is the concern of ideas *being heard* by others (BeingHeard). This variable has a mean distance of 0, suggesting that all students in the cluster do not think that this variable is a concern. On the other hand, only Cluster 2's BeingHeard variable mean distance is not 0, suggesting that researchers can dig deeper into the description of Cluster 2 to learn more about its composition (see Tables A5 to A7).

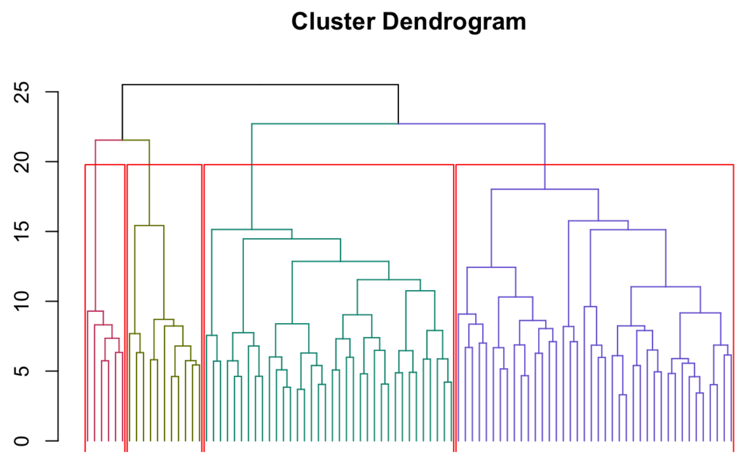

Figure 9. Hierarchical Clustering Dendrogram displaying the grouping of students.

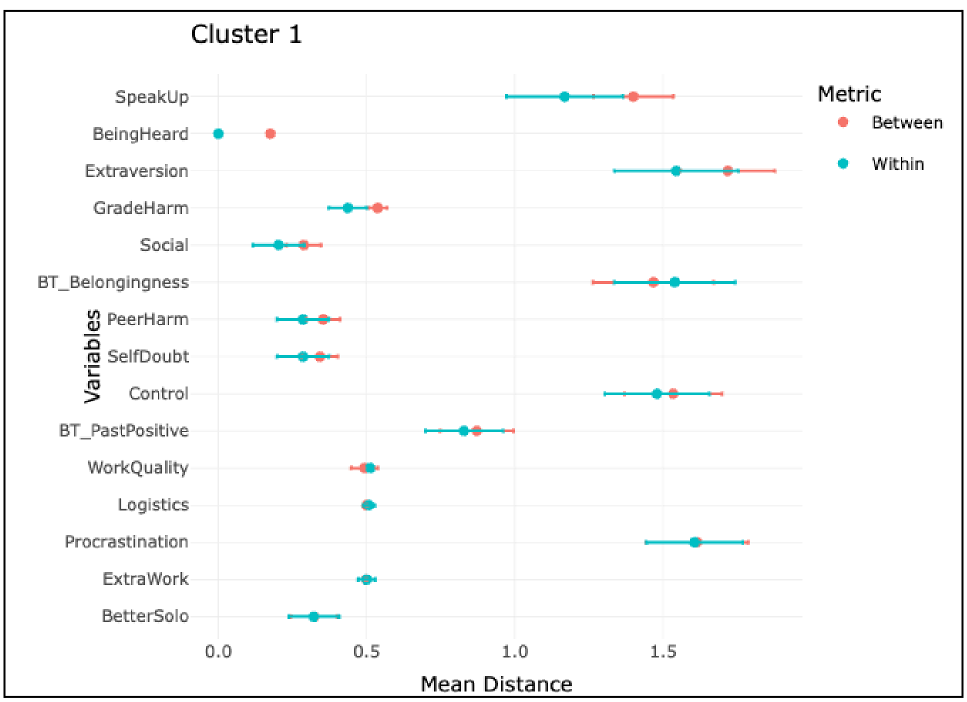

Figure 10. Cluster analysis showing mean distances for key variables "within" and "between" Clusters 1. For each variable, the "within" mean distance (blue) measures how closely each data is from other data of the same cluster (Cluster 1). The "between" mean distance (red) measures the distance of data in Cluster 1 from all other clusters.

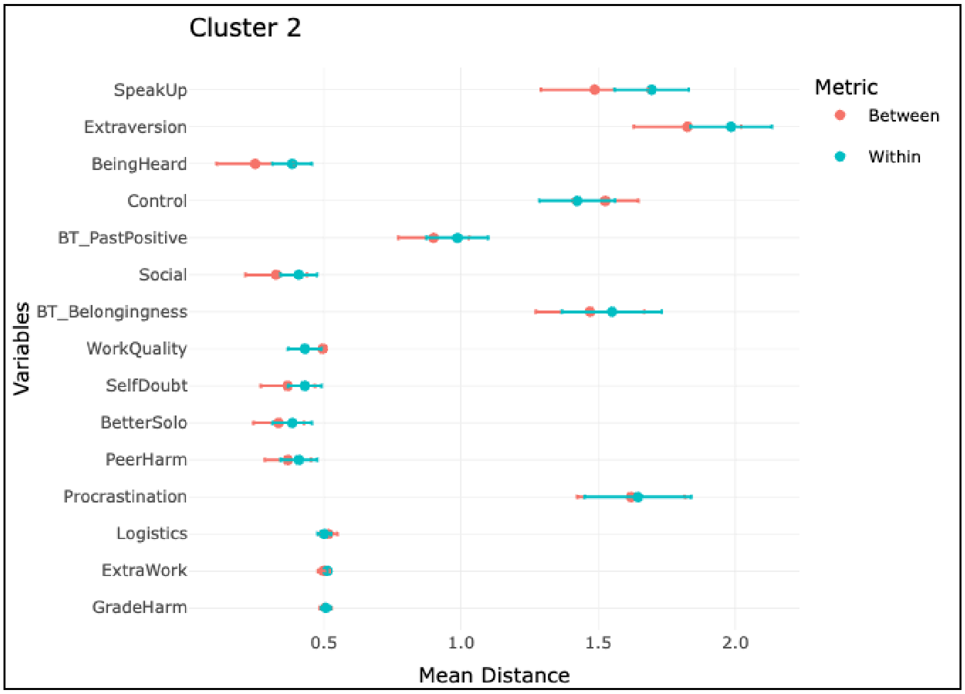

Figure 11. Cluster analysis showing mean distances for key variables "within" and "between" Clusters 2. For each variable, the "within" mean distance (blue) measures how closely each data is from other data of the same cluster (Cluster 2). The "between" mean distance (red) measures the distance of data in Cluster 2 from all other clusters.

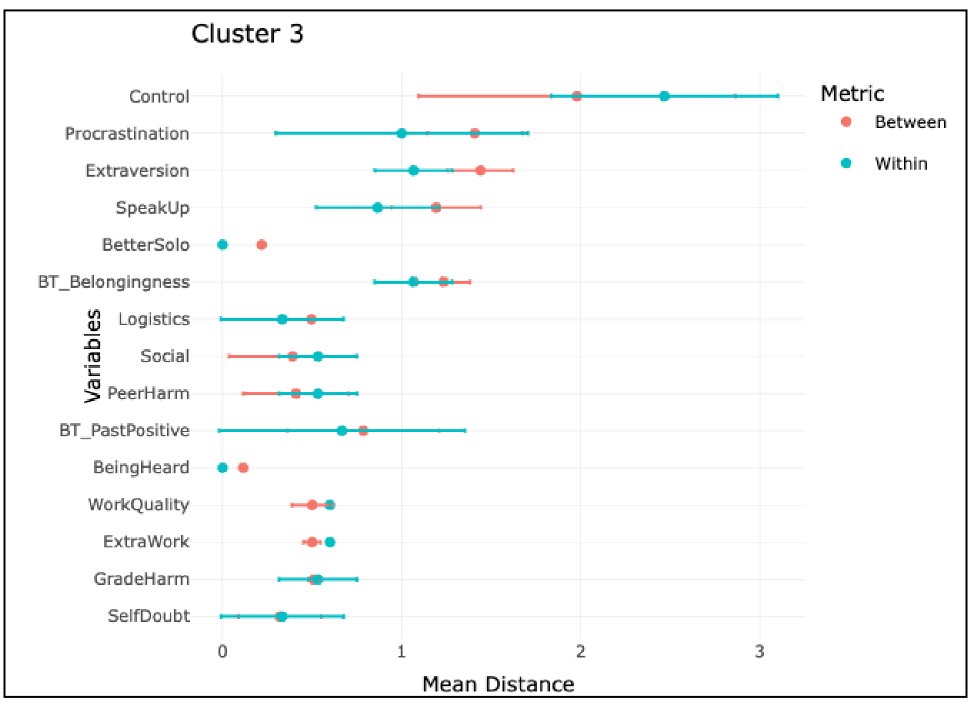

Figure 12. Cluster analysis showing mean distances for key variables "within" and "between" Clusters 3. For each variable, the "within" mean distance (blue) measures how closely each data is from other data of the same cluster (Cluster 3). The "between" mean distance (red) measures the distance of data in Cluster 3 from all other clusters.

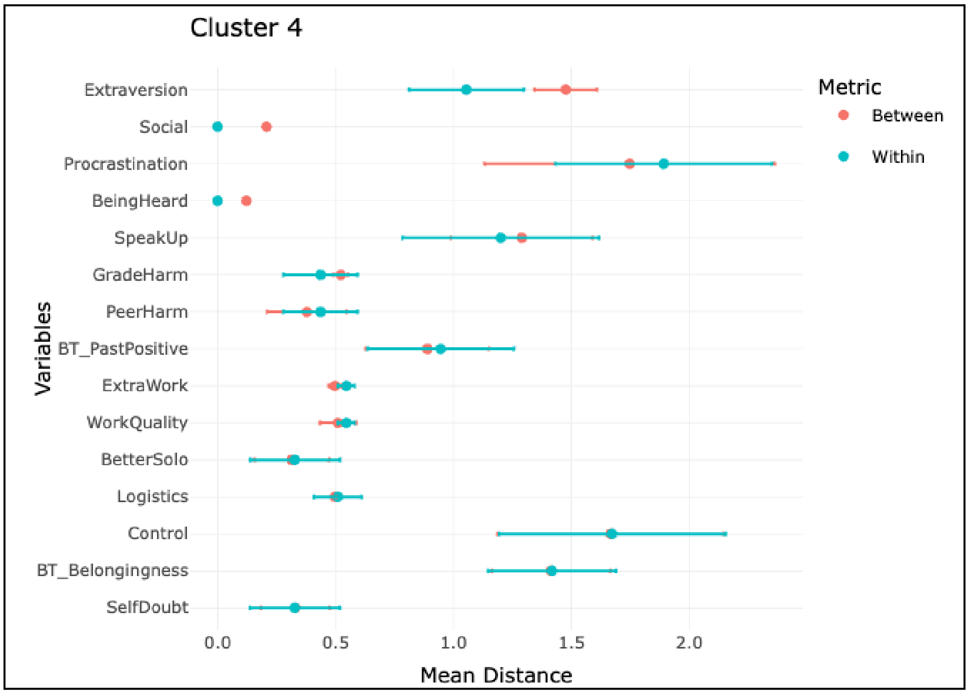

Figure 13. Cluster analysis showing mean distances for key variables "within" and "between" Clusters 4. For each variable, the "within" mean distance (blue) measures how closely each data is from other data of the same cluster (Cluster 4). The "between" mean distance (red) measures the distance of data in Cluster 4 from all other clusters.

RQ4: What recurring factors or variables significantly influence these experiences, and what strategies can be implemented to address them?

The recurring theme of concern about ideas *being heard* by others suggests that communication issues are a significant factor influencing negative experiences in team settings. The homogeneity within clusters regarding this concern highlights the need for strategies that enhance communication skills and ensure that all team members have equal opportunities to contribute. To address these communication challenges, strategies such as structured team meetings, clear role assignments, and active listening exercises could be implemented. For instance, incorporating regular feedback sessions where all members are encouraged to voice their opinions may help in ensuring that everyone is heard. Furthermore, training sessions on effective communication could be beneficial, especially for those identified within the *extraversion* variable, who may be more reserved and less likely to speak up in group settings.

The presence of better off *working alone* as a variable points to some students' preference for individual work, suggesting that team-based projects may not align with their natural working style. To accommodate such preferences, a blend of individual and group tasks could be introduced, allowing students to showcase their strengths in both domains. Concern about *getting to know new people* indicates the importance of social bonds and a supportive team environment. Team-building activities and social events could be organized to foster camaraderie and trust among team members. Lastly, since the *extraversion* personality appears as a variable, personality differences need to be considered in team compositions. When a team is less dissatisfied and the dynamic is harmonious, the team functions better and has better performance. Therefore, faculty members and instructors can consider forming balanced and equitable teams through clustering algorithms after analyzing the students' personalities and traits [5], [6]. By recognizing and addressing these key factors, faculty members and instructors can tailor their first-year engineering courses to better suit the varied needs of students, ultimately enhancing the teams' learning experiences and outcomes.

Conclusion and Future Directions

This study has highlighted the varied and complex nature of team experiences in first-year engineering courses across gender, ethnicity, nationality, and season. Notably, compared to male students, female students reported higher levels of concern about *self-doubting* themselves when they make a mistake (SelfDoubt), ideas *being heard* by others (BeingHeard), and their *grades being hurt* by others (GradeHarm). When analyzed across different races, *Black* students demonstrated significantly fewer concerns regarding taking up a *higher share of the workload to make up* for my teammates (ExtraWork) (Holm-adjusted p-value of 0.04). Similarly, *Asian* students exhibited a markedly lower level of concern in *self-doubting* themselves when they made a mistake (SelfDoubt), with a Holm-adjusted p-value of 0.03. When analyzed across seasons, our analysis indicated significant variations in students' concerns about taking a *higher share of the workload to make up* for teammates (ExtraWork) between the Fall and Winter seasons. Since the data originated from first-year engineering courses at a research-focused public university, it is plausible that the increased concern in the fall season stemmed from freshmen being unfamiliar with the university's workload. This evidence calls for increased support for students as they adapt to the academic rigor of their programs.

Future research should further investigate the root causes of communication difficulties within various types of groups and develop targeted interventions to facilitate better team dialogue. Since the data originated from first-year engineering courses, one plausible solution would be to start the intervention in high school. Additionally, longitudinal studies may provide deeper insights into how team experiences evolve throughout a student's academic career and how early interventions can have long-term benefits. Lastly, examining the intersectionality of students' identities and experiences will enhance our understanding of the multidimensional nature of team dynamics. Such research could be useful for Tandem so that it can be tailored to the unique strengths and preferences of individual students.

Acknowledgment

The authors express their sincere gratitude to the University of Michigan Center for Academic Innovation for providing the essential data and funding from the Teresa Noel Urban Blaurock Student Research Award to support the first author.

References

[1] L. K. Michaelsen, W. Watson, J. P. Cragin, and L. Dee Fink, “Team Learning: a Potential Solution To the Problems of Large Classes,” *Exch. Organ. Behav. Teach. J.*, vol. 7, no. 1, pp. 13–22, Jan. 1982, doi: 10.1177/105256298200700103.

[2] V. Najdanovic-Visak, “Team-based learning for first year engineering students,” *Educ. Chem. Eng.*, vol. 18, pp. 26–34, Jan. 2017, doi: 10.1016/j.ece.2016.09.001.

[3] R. R. Fowler and M. P. Su, “Gendered Risks of Team-Based Learning: A Model of Inequitable Task Allocation in Project-Based Learning,” *IEEE Trans. Educ.*, vol. 61, no. 4, pp. 312–318, Nov. 2018, doi: 10.1109/TE.2018.2816010.

[4] J. H. Chin, Y. Gao, H. Li, M. P. Su, and R. Fowler, “Predicting Team Project Score: It’s More about Team Harmony and Less about Individual Performance,” in *2020 ASEE Virtual Annual Conference Content Access*, 2020.

[5] J. H. Chin, H. Li, and R. Fowler, “Proof of Concept: An Algorithm for Consideration of Students’ Personalities in Team Formation,” in *2021 ASEE Virtual Annual Conference Content Access*, 2021.

[6] J. H. Chin, J. Ouyang, G. Xu, R. Fowler, and R. L. Matz, “Predicting Team Function Using Bayesian and Cognitive Diagnostic Modeling Approaches,” in *2023 ASEE Annual Conference & Exposition*, 2023.

[7] 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,” presented at the 2021 ASEE Virtual Annual Conference Content Access, Jul. 2021. Accessed: Oct. 29, 2023. [Online]. Available: https://peer.asee.org/supporting-equitable-team-experiences-using-tandem-an-online-assessment-and-learning-tool

[8] A. Robitzsch, “Why Ordinal Variables Can (Almost) Always Be Treated as Continuous Variables: Clarifying Assumptions of Robust Continuous and Ordinal Factor Analysis Estimation Methods,” *Front. Educ.*, vol. 5, 2020, Accessed: Feb. 08, 2024. [Online]. Available: https://www.frontiersin.org/articles/10.3389/feduc.2020.589965

[9] F. Murtagh and P. Legendre, “Ward’s Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward’s Criterion?,” *J. Classif.*, vol. 31, no. 3, pp. 274–295, Oct. 2014, doi: 10.1007/s00357-014-9161-z.

[10] T. L. Quy, G. Friege, and E. Ntoutsi, “A review of clustering models in educational data science towards fairness-aware learning,” 2023, pp. 43–94. doi: 10.1007/978-981-99-0026-8_2.

[11] S. Holm, “A Simple Sequentially Rejective Multiple Test Procedure,” *Scand. J. Stat.*, vol. 6, no. 2, pp. 65–70, 1979.

[12] S. Saraçli, N. Doğan, and İ. Doğan, “Comparison of hierarchical cluster analysis methods by cophenetic correlation,” *J. Inequalities Appl.*, vol. 2013, no. 1, p. 203, Apr. 2013, doi: 10.1186/1029-242X-2013-203.

Appendix

Table A1: Pearson's Chi-squared test results with Yates' continuity correction for student concerns (by Gender).

Pearson's Chi-squared test with Yates' continuity correction (Gender)			
Variables	X-squared	df	p-value
Logistics	1.46	1	0.23
GradeHarm	4.80	1	**0.03**
PeerHarm	0.09	1	0.76
WorkQuality	2.64	1	0.10
ExtraWork	3.67	1	0.06
BeingHeard	39.20	1	**0.00**
Social	0.02	1	0.88
SelfDoubt	105.70	1	**0.00**
BetterSolo	2.90	1	0.09

Table A2: Pearson's Chi-squared test results with Yates' continuity correction for student concerns (by Nationality).

Pearson's Chi-squared test with Yates' continuity correction (Nationality)			
Variables	X-squared	df	p-value
Logistics	3.39	2	0.18
GradeHarm	0.98	2	0.61
PeerHarm	2.01	2	0.37
WorkQuality	1.49	2	0.47
ExtraWork	5.35	2	0.07
BeingHeard	4.00	2	0.14
Social	6.41	2	**0.04**
SelfDoubt	3.43	2	0.18
BetterSolo	1.57	2	0.46

Table A3: Pearson's Chi-squared test results with Yates' continuity correction for student concerns (by Race).

Pearson's Chi-squared test with Yates' continuity correction (Race)			
Variables	X-squared	df	p-value
Logistics	6.49	5	0.26
GradeHarm	1.13	5	0.95
PeerHarm	7.58	5	0.18
WorkQuality	16.62	5	**0.01**
ExtraWork	13.38	5	**0.02**
BeingHeard	10.85	5	0.05
Social	9.70	5	0.08
SelfDoubt	19.02	5	**0.00**
BetterSolo	8.19	5	0.15

Table A4: Pearson's Chi-squared test results with Yates' continuity correction for student concerns (by Season).

Pearson's Chi-squared test with Yates' continuity correction (Season)			
Variables	X-squared	df	p-value
Logistics	1.36	1	0.24
GradeHarm	0.04	1	0.84
PeerHarm	0.00	1	1.00
WorkQuality	0.84	1	0.36
ExtraWork	7.14	1	**0.01**
BeingHeard	0.42	1	0.52
Social	1.49	1	0.22
SelfDoubt	0.25	1	0.62
BetterSolo	0.02	1	0.88

Table A5: Description of each cluster (BoT survey questions on students' traits).

Cluster	Control	SpeakUp	Extraversion	BT_PastPositive	Procrastination	BT_Belongingness	Past Teamwork Experience			Preference		
							Several Times	Many Times	Once Or Twice	Partner	Group	Alone
1	4.17	4.72	4.64	3.56	4.22	3.28	0.36	0.56	0.08	0.50	0.19	0.31
2	4.05	4.25	4.00	3.60	4.53	3.38	0.25	0.75	0.00	0.40	0.38	0.23
3	3.17	4.17	4.00	3.67	4.83	3.00	0.50	0.50	0.00	0.83	0.17	0.00
4	3.27	4.45	4.64	3.82	3.73	2.91	0.09	0.91	0.00	0.36	0.45	0.18

Table A6: Description of each cluster (BoT survey questions on pre-semester concerns).

Cluster	Logistics	GradeHarm	PeerHarm	WorkQuality	ExtraWork	BeingHeard	Social	SelfDoubt	BetterSolo
1	0.56	0.69	0.17	0.47	0.58	0.00	0.11	0.17	0.19
2	0.43	0.45	0.28	0.30	0.50	0.25	0.28	0.30	0.25
3	0.83	0.33	0.33	0.50	0.50	0.00	0.33	0.17	0.00
4	0.64	0.27	0.27	0.55	0.55	0.00	0.00	0.18	0.18

Table A7: Description of each cluster (Institutional database demographic information).

Cluster	International	Minority	Non-Minority	Non-Underrepresented Minority	Underrepresented Minority	Native English Speaker	Ethnicity (2 or More)	Asian	Black	Hispanic	Not Indic	White	Female	Male	Fall	Winter
1	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.19	0.81	0.36	0.64	0.67	0.33
2	0.00	0.85	0.15	1.00	0.00	0.10	0.08	0.78	0.00	0.00	0.03	0.13	0.30	0.70	0.58	0.43
3	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.67	0.00	0.17	0.17	0.00	0.50	0.50	0.50	0.50
4	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	0.27	0.73	0.00	0.00	0.36	0.64	0.45	0.55