Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025
Paper ID #47978

Collaborative Learning in Engineering: Analyzing the Effects of Group Formation on Student Outcomes

Major Brett Ryan Krueger, United States Military Academy

MAJ Brett Krueger, Instructor of Environmental Engineering, USMA; brett.krueger@westpoint.edu

MAJ Krueger is an Infantry Officer and Instructor in the Department of Geography and Environmental Engineering at West Point. Brett currently teaches courses in Environmental Engineering Technologies as well as Environmental Engineering in Developing Communities. He holds an MS in Civil and Environmental Engineering from Stanford University and a BS from the United States Military Academy in Environmental Engineering. Before joining the faculty at West Point, he served as an Infantry platoon leader and company commander in airborne and mechanized units with operational assignments in Poland, Germany, Ukraine, and Korea.

William B Vass, United States Military Academy

William Vass is an assistant professor in the Department of Geography and Environmental Engineering at the United States Military Academy. His PhD research involved air sampling to assess potential airborne virus exposure risks. Specifically, his work included virus aerosolization and collection, with a focus on the enrichment of viruses in samples while conserving their viability. He also conducted air sampling to quantify virus presence in occupational and residential settings and thereby better understand potential human health risks.

Matthew Baideme, United States Military Academy

Matthew Baideme is an active duty officer in the United States Army. He received his B.S. from the United States Military Academy (2002), M.S. from Stanford University (2012), and Ph.D. from Columbia University (2019). He teaches courses in environmental engineering at the United States Military Academy, with research and teaching interests focused on engineered biological treatment systems, microbial nitrogen cycling, and microbial biochemical degradation pathways.

Collaborative Learning in Engineering: Analyzing the Effects of Group Formation on Student Outcomes

Abstract

Learning and problem-solving in the fields of environmental engineering and sustainability necessitate the formation of student groups in academic settings. Group performance varies. This study examined how different group formation methods affect individual student performance in a collegiate collaborative learning environment. Groups were divided randomly, by student selection, or by using the Comprehensive Assessment of Team Member Effectiveness (CATME) System for Management, Assessment, Research, Training, Education, and Remediation (SMARTER) Teamwork tool by Purdue University. This study analyzed academic outcomes to determine how student group formation affected individual performance. The study found no statistically significant differences in individual performance between group formation methods. However, scores in all group types significantly improved between pre- and post-quizzes. That improvement indicated that collaborative learning positively impacted overall student performance. These findings suggest uncertainty about the extent to which educators can choose group formation methods without significantly affecting individual learning outcomes. This discussion also addresses the importance of other factors, such as group dynamics and instructional methods, in enhancing collaborative learning. Future research should explore additional group formation variables and unmeasured factors influencing group and individual performance.

Keywords: engineering education, collaborative learning, design project

1. Introduction

Collaborative learning has become a cornerstone of modern educational approaches, fostering student engagement and enhancing individual and collective performance. The question of how different group formation strategies impact student performance remains critical for optimizing collaborative learning environments. In a previous study, team learning in an academic course has been found to not significantly increase the individual performance levels of students [1]. This study investigated the effects of three distinct group formation methods, random assignment, self-selection, and algorithmic grouping via the CATME tool, on individual student performance [2], [3]. A previous study examined the impact of random and self-selection group formation on performance, finding no consistent correlation between group grades and individual performance [4]. The CATME tool effectively forms teams with compatible schedules, which positively affects group dynamics [5]. Previous research also indicated that pairing students of heterogeneous abilities, a capability of the CATME tool, can positively affect the learning of lesser-ability students [6]. While heterogeneous sorting can be beneficial to learning, these group formation methods have been found to negatively affect group dynamics through social loafing as well as resentment from higher-ability students [7]–[9]. Additionally, self-selected groups have a tendency to form in-groups with their friends or choose individuals that look like them, resulting in a lack of diversity [10]. Understanding the relationship between group formation methods and student performance can inform andragogical practices and help educators maximize the benefits of collaborative learning. The objective of this research was to identify how group formation methods influence student outcomes, measured by exam grades, engineering design project (EDP) scores, and specific lesson objective performance assessments. This study produced empirical insights into optimal group formation practices through a controlled experiment with undergraduate students enrolled in a collaborative learning-based course. The findings contribute to the broader discourse on effective teaching strategies and support data-driven decision-making in educational settings.

2. Methods

The study design involved the enrollment of 196 undergraduate students taking a senior design course for non-engineers. As part of the undergraduate curriculum, students whose academic major is not an engineering field must take three engineering courses, which can include this senior design course. Due to the low risk to the subjects, the Human Research Protections Program determined this study to be Institutional Research Board Exempt (CA-2025-10).

Students were taught by four (4) different instructors who collaborated with each other and oversaw 11 total cohort sections. 30% of the course point allocation was allocated to group assessments. Instructor teaching experience ranged from zero to seven years. Cohorts were randomly assigned a grouping methodology among three group formation methods (Table 1). All groups in the same cohort were formed using the same method. Students grouped randomly were assigned using the Microsoft Excel random function. Students chose their partners in the self-select group with no limiting criteria. CATME groups were sorted using the CATME tool, a web-based platform designed to create well-functioning academic teams. The tool attempted to optimize groupings based on student schedules and performance in their most recent engineering course, Environmental Engineering Technologies. The tool prioritized the alignment of student availability within groups and the equitable distribution of higher and lower performers between groups. One instructor had only a CATME cohort because that person taught only one section of the course. All other instructors had cohorts of each group formation method.

Table 1. Summary of the Number of Students (Groups) in Data. Random, Self-Select, and CATME refer to the group formation method employed by each of the four instructors.

Instructor	Random	Self-Select	CATME	Total
Instructor A	17 (6)	19 (7)	18 (6)	54 (19)
Instructor B	-	-	19 (7)	19 (7)
Instructor C	31 (11)	19 (7)	18 (6)	68 (24)
Instructor D	19 (7)	17 (6)	19 (7)	55 (20)
Total	67 (24)	55 (20)	74 (26)	196 (70)
Students Removed from Data	8	5	7	20
Number of Students in Data	**59**	**50**	**67**	**176**

The experiment began with all students individually completing a pre-quiz to establish baseline knowledge and performance related to specific learning objectives. Those objectives pertained to the EDP, which was designed to evaluate student comprehension of engineering economics. Briefly, EDP desired learning outcomes were that students could (1) understand the concept of the time value of money, (2) construct a cash flow diagram (CFD) to model payments and receipts over time, (3) estimate project costs, and (4) solve a cash flow series for net present worth equivalence. In the three weeks following the pre-quiz, students collaborated within their assigned groups on the EDP. The project consisted of a series of designs addressing water, sanitation, and health scenarios in low- and middle-income countries. Two questions directly assessed the same engineering economic learning objectives from the pre-and post-quizzes. Performance on those questions was tallied as an EDP sub-score and used to assess student performance for this study. After completing the group work, all students individually took a post-quiz to assess learning outcomes and the impact of their group experience. Individual performance was measured using pre- and post-quiz scores, while group performance was assessed through scores on the engineering design project sub-scores.

We conducted statistical analyses using data collected from pre- and post-quizzes and group project sub-scores to evaluate the impact of group formation methods on student performance. Some students (20) could not take either the pre-quiz or post-quiz on the designated dates and were removed from the dataset. Paired tests were used to determine whether improvement over the course of the study was significant while accounting for repeated measures of individuals. Individual performance scores on the pre- and post-quizzes were analyzed within the three group formation categories. Wilcoxon signed rank tests were used to determine the significance of the changes. Data were further divided according to students' exam performance to explore potential interactions between group formation and individual performance capacity. Exam score categories were based on two test scores and the final exam score. "High" performers were those who scored on average 87-100% (B+ - A+), while "Medium" performers scored 77-86.99% (C+ - B), and "Low" performers scored <77% (C and below). These group cut points helped generate similarly-sized groups. These analyses were undertaken to help explain how each group formation method influenced individual learning trajectories, complementing the broader comparison of score changes across groups.

Multivariate linear models were utilized to help explain the most influential factors on changes in students' pre- and post-quiz scores due to complexities and uncertainties inherently associated with this type of study. The response variable in these models was the change in individual quiz scores (post-quiz minus pre-quiz). Predictors were the group formation method (Random, Self Select, CATME), instructor (A, B, C, D), and exam performance level (High, Medium, Low). The analysis aimed to identify the relative effects of these factors on performance changes and explore potential interactions between them. Five models were constructed and compared using the Akaike Information Criterion (AIC) to evaluate model fit and complexity. The selected model incorporated the main effects and interactions between the group formation method and instructor. The Low exam performers in Self Select groups taught by Instructor A were held as the reference group within the model. This model provided insights into the relationships between the predictors and the outcome variable while accounting for variability across groups and instructors.

3. Results and Discussion

Shapiro-Wilk tests indicated that score change data met the assumption of normality across the three group formation methods. However, Levene's test revealed heteroskedasticity. Given these results, non-parametric tests were used in our analyses. A Kruskal-Wallis test was used to compare changes in scores across the three group formation methods. This test provided insight into whether the groups had statistically significant differences in performance changes. Performance changes were calculated as the difference between post-quiz and pre-quiz scores. Additionally, medians and interquartile ranges were calculated to summarize the central tendency and variability within each group. A paired analysis was conducted to examine the individual performance changes within each group formation method. This analysis focused on paired pre- and post-quiz scores of each participant, allowing for a direct assessment of learning improvements within each group. A Kruskal-Wallis test comparing changes in quiz scores across the three group formation methods showed no significant effect ($p = 0.35$) from the group formation method on the median scores within each group type. The median change in performance scores for random, self-select, and CATME groups were 2.38, 2.5, and 2.25, respectively (Figure 1). These results can be explained by qualitative observations by instructors when grading design projects. It was apparent that students often divided responsibility for parts of the project among different group members, as indicated by distinctly different grammar, structure, style, or general quality of different portions of the same project. Therefore, group formation would be irrelevant as a predictor of performance as the group element of the assignment frequently became, in effect, an individual assignment. In future studies, researchers could collect data on which students complete each section of the project or develop a design project that requires students to collaborate on specific parts of the project. Additionally, while this research is assessing the effectiveness of the CATME group formation method, it only used two group formation variables: schedule compatibility and previous-course grade point. Many possible variables are offered in the CATME system. These two variables do not provide an effect, but it is possible that different CATME variable permutations do.

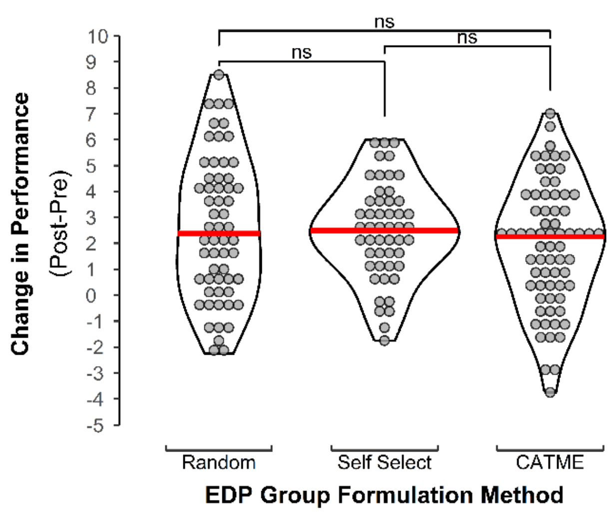

Figure 1. This graph illustrates the distribution of score changes (post-quiz minus pre-quiz) across the three group formation methods. Plots show the spread of data, while the red line indicates median values. Individual data points are represented to highlight trends and variability within each group. Statistical comparisons indicate no significant differences ("ns") in performance changes between groups.

Next, the paired data were visualized by plotting pre- and post-quiz scores for each group, showcasing trends and individual group variability. Boxplots were used to show score distribution changes over time within each group type (Figure 2). Lines connecting data points represent individual changes in performance, allowing for the observation of trends such as consistent improvements or outliers. Figure 2 also shows color-coded information reflecting participant performance on course exams. Using a paired analysis, a significant difference ($p << 0.01$) between pre- and post-quiz performance is shown for all group formation methods (Figure 2). Students improved on average, but no discernible trend could be explained by grouping method when analyzing the changes in individual quiz scores.

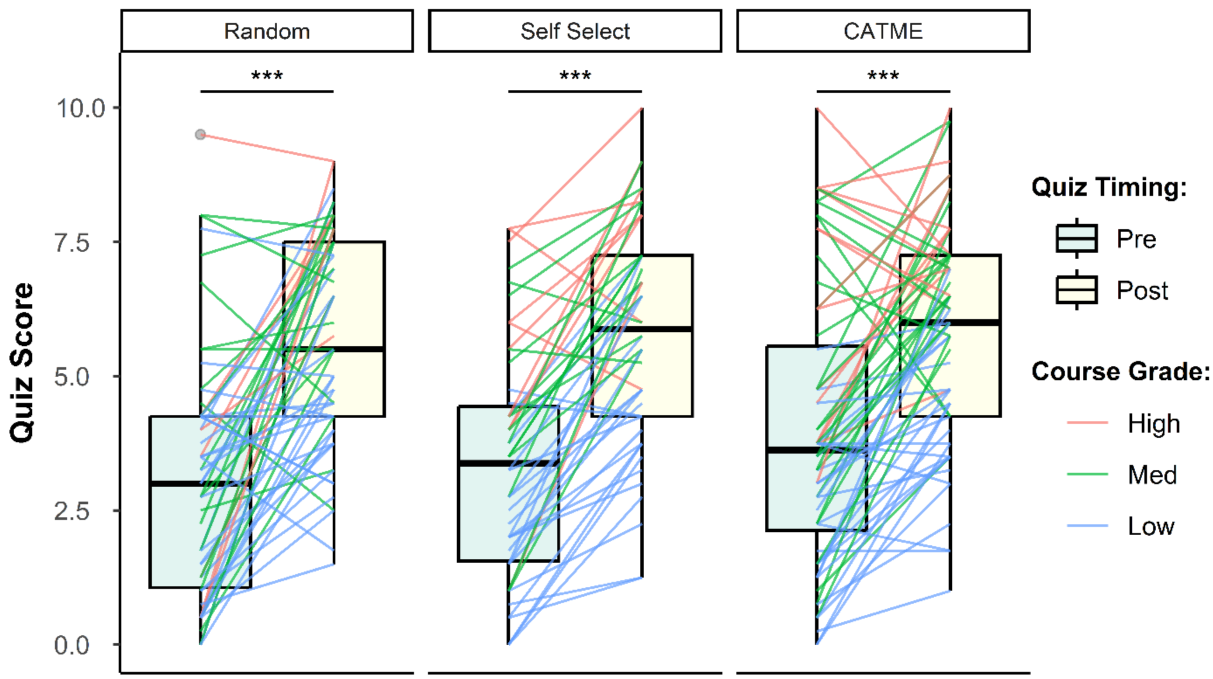

Figure 2. This graph depicts individual and group-level changes in quiz scores from pre- to post-assessment for each group formation method. Boxplots show the distribution of scores at each time point, while lines connecting individual data points illustrate within-subject changes. Color coding represents participants' performance categories, highlighting potential interactions between baseline performance and group dynamics. Indicators (***) signify that pre- and post-quiz scores were significantly different within each group type, as determined by paired Wilcoxon signed rank tests.

A linear model provided additional insight into the dataset by showing significant effects on score changes from some predictor variables and interactions. The model showed that students taught by two instructors (B and D) had significantly less improvement ($p = 0.025$ and 0.0015) as compared to the reference group. Further, the model showed a significantly positive effect on the scores of the CATME group taught by Instructor D ($p = 0.047$). Neither group formation nor exam performance category had a significant effect on quiz score changes within this dataset. The instructors in this study had a wide range of experience instructing. While they collaborated on teaching methodology, each ultimately taught independently. Therefore, modeling results seem to suggest that how information is taught to students matters and that variation in teaching has the potential to affect student performance compared to others in the same course. The idea that instructor methods can affect student performance is far from a novel idea [11], but these data emphasize the point. The results further suggest that synchronization among instructors may be prudent to help ensure equitable provision of information to students. However, the results from this model should be interpreted cautiously. The standard error for this model (2.14) was high considering the intercept predicted value of an average score change of 3.12 for the reference group. That error, taken with the multiple R-squared for the model (0.248) suggests that the model needs to be improved by the addition of pertinent predictor variables. This model can be used to explain this dataset in conjunction with numerical tests and graphical depictions previously discussed, but the model should not be extrapolated out and applied beyond this study. We encourage researchers investigating similar group performance dynamics in the future to design their work to collect myriad predictor variables and assess multiple performance measures as response variables. With a larger dataset, the broader engineering education community might be able to better determine key indicators of student success.

Given this flexibility, educators may prioritize group formation methods based on convenience, student preference, or heterogeneity. For instance, randomly-selected groups can streamline logistics when simplicity is important, whereas allowing students to self-select partners may enhance motivation and group cohesion in settings where peer relationships are valued. Alternatively, tools like CATME can be employed when the goal is to create groups of students with similar past academic performance or accommodate scheduling constraints. The results highlight the potential for focusing on other factors, such as group dynamics or instructional methods, to enhance the effectiveness of collaborative learning. However, this study's findings are constrained by its specific context, including the sample of undergraduate students and the controlled nature of the learning environment. Future research should explore additional variables, such as different group formation variables, to identify key factors influencing collaborative success. While this study focused on schedule compatibility and prior course performance, future studies could incorporate additional options in the CATME Team-Maker tool, such as academic discipline, specific skills like writing ability, or cognitive styles such as big-picture versus detail-oriented thinking. Depending on the attribute, group formation could be tailored to create more homogeneous or heterogeneous teams. Expanding the use of CATME variables would allow researchers to assess how targeted group composition influences collaborative learning outcomes, using a combination of performance metrics and peer feedback.

4. Conclusion

Despite the widespread recognition of collaborative learning as a critical andragogical approach, the specific influence of group formation strategies on student outcomes has remained an open question. Our results indicate that while all groups exhibited learning improvements, the group formation method had no statistically significant effect on individual performance. These findings contribute valuable insights into optimizing collaborative learning practices. Specifically, they suggest that educators may have flexibility in choosing group formation strategies without risking significant disparities in individual performance outcomes. Further insights into key predictors of student success during collaborative engineering groupwork could be optimized by increasing the collection of predictor data points during future studies. Overall, this research underscores the importance of evidence-based approaches to group formation while encouraging educators to prioritize strategies that foster effective learning environments. These insights provide a foundation for enhancing collaborative learning practices across diverse educational settings.

References

[1] J. Kreie, R. W. Headrick, and R. Steiner, "Using team learning to improve student retention," *College Teaching*, vol. 55, no. 2, pp. 51–56, Apr. 2007, doi: 10.3200/CTCH.55.2.51-56.

[2] "Welcome to CATME - Smarter Teamwork." https://info.catme.org/ (accessed Jan. 12, 2025).

[3] R. A. Layton, M. L. Loughry, M. W. Ohland, and G. D. Ricco, "Design and Validation of a Web-Based System for Assigning Members to Teams Using Instructor-Specified Criteria," *Advances in Engineering Education*, no. 2, pp. 1–28, 2010.

[4] M. Baideme, K. Newhart, C. Robbins, M. Butkus, and A. Pfluger, "Influence of Group Learning in Environmental Engineering: A Curriculum and Course-level Assessment," presented at the 2023 ASEE Annual Conference & Exposition, Jun. 2023, doi: 10.18260/1-2--43689.

[5] B. J. Millis and P. G. Cottell Jr, *Cooperative Learning for Higher Education Faculty. Series on Higher Education.* Oryx Press, P.O. Box 33889, Phoenix, AZ 85067-3889; phone: 800-279-6799; fax: 800-279-4663 ($39.95)., 1997.

[6] L. Vygotsky, *Mind in society: Development of higher psychological processes. L. S. Vygotsky*. Harvard University Press, 1978.

[7] M. C. Schippers, "Social loafing tendencies and team performance: the compensating effect of agreeableness and conscientiousness," *Academy of Management Learning & Education*, vol. 13, no. 1, pp. 62–81, Mar. 2014, doi: 10.5465/amle.2012.0191.

[8] R. R. Hake, "Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses," *Am. J. Phys.*, vol. 66, no. 1, p. 64, 1998, doi: 10.1119/1.18809.

[9] R. M. Felder and R. Brent, "Cooperative Learning," in *Active Learning: Models from the Analytical Sciences*, vol. 970, P. A. Mabrouk, Ed. Washington, DC: American Chemical Society, 2007, pp. 34–53.

[10] D. R. Bacon, K. A. Stewart, and W. S. Silver, "Lessons from the Best and Worst Student Team Experiences: How a Teacher can make the Difference," *Journal of Management Education*, vol. 23, no. 5, pp. 467–488, Oct. 1999, doi: 10.1177/105256299902300503.

[11] S. E. Carrell and J. E. West, "Does Professor Quality Matter? Evidence from Random Assignment of Students to Professors," *Journal of Political Economy*, vol. 118, no. 3, pp. 409–432, Jun. 2010, doi: 10.1086/653808.