# Gender-Based Performance in a Collaborative Learning Engineering Classroom

**Hanwei Wang, University of Wisconsin - Madison**

Hanwei Wang is a PhD candidate in Environmental Chemistry and Technology with a doctoral minor in Cartography and Geographic Information Systems at the University of Wisconsin–Madison. She received her Bachelor's degree in Environmental Engineering from Dalian University of Technology and then earned an MS in Environmental Health from Johns Hopkins School of Public Health. Building on her teaching experiences, Hanwei developed a teaching-as-research project investigating gender-based performance in a collaborative engineering classroom and received the Delta Certificate in Research, Teaching and Learning.

**Prof. Katherine D McMahon, University of Wisconsin - Madison**

# Gender-Based Performance in a Collaborative Learning Classroom


## Abstract

The underrepresentation of women in engineering continues to be a significant challenge. As of 2021, women comprised only 31% of environmental engineers in the U.S. While there has been an increase in the number of women pursuing engineering degrees and careers, gender disparities in performance and representation continue to persist. Addressing these disparities requires a deeper understanding of gender dynamics in undergraduate education, where foundational engineering skills are developed.

In this study, we collected data of 598 students enrolled in an introductory course, *Environmental Engineering*, over eight semesters (2015-2023), taught by the same female instructor. In this course, collaborative learning strategies were implemented to enhance peer instruction. Learning groups of 3-5 students were assigned, with the guiding principle of avoiding gender imbalance that could marginalize female students. These groups worked together weekly on problem sets, submitting a single solution for grading, with all group members receiving the same grade. Students were assessed individually using weekly quizzes and three exams.

This retrospective study performed quantitative analyses to investigate 1) gender-based performance difference across various assessments; 2) the influence of the gender composition of collaborative learning groups on individual student performance; 3) the impact of academic level on performance outcomes.

Our findings showed that the percentage of female students enrolled in the course increased from 25% in 2015 to 51% in 2023. Using grade point average (GPA) as a performance indicator, a two-sample t-test revealed no significant difference in overall performance between male and female students ($p = 0.28$) across all semesters aggregated. To assess the impact of group gender composition, students were categorized as females in single-gender groups, females in mixed-gender groups, males in single-gender groups, and males in mixed-gender groups. Analyses of variance (ANOVA) with Tukey's honest significantly difference (HSD) tests were performed using problem set scores as the direct reflection of collaborative learning outcomes. Our findings revealed that female students in mixed-gender groups performed significantly better than those in single-gender groups ($p = 6.8 \times 10^{-6}$), while no significant difference was found between males in single and mixed-gender groups ($p = 0.27$), or between female and male students in mixed gender group ($p = 0.94$).

Similarly, ANOVA analyses with Tukey's HSD tests of GPA by academic level indicated that prior academic experience influenced performance. Graduate students and seniors performed better than sophomores ($p = 0.038$ and $p = 0.0049$, respectively), though no significant difference was observed between sophomores and juniors ($p = 0.23$).

In conclusion, no significant gender difference was observed in the overall course performance. However, group dynamics played a crucial role in collaborative learning settings, and stage of education could influence student success. The finding that female students performed better in mixed-gender groups aligns with previous studies on both K-12 engineering education and

college economic decision-making, which advocates for the use of mixed-gender learning groups for improved outcomes. Our findings provide valuable insights for the design of inclusive and effective learning environments in engineering education, recognizing the importance of addressing factors beyond gender composition to foster equitable and supportive academic experiences.

## Introduction

The field of engineering has historically exhibited one of the largest gender gaps among science, technology, engineering, and mathematics (STEM) disciplines, with women significantly underrepresented in both academic and professional settings.[1-4] Despite concerted efforts to increase the proportion of women engineers, progress has been gradual. For example, in environmental engineering, only 23% of professionals in the United States were women in 2014, according to Census Bureau data, and this figure rose modestly to 31% by 2021 (Fig. 1).[5] While this increase is notable, the workforce remains predominantly male, underscoring the persistent challenges in achieving gender parity.

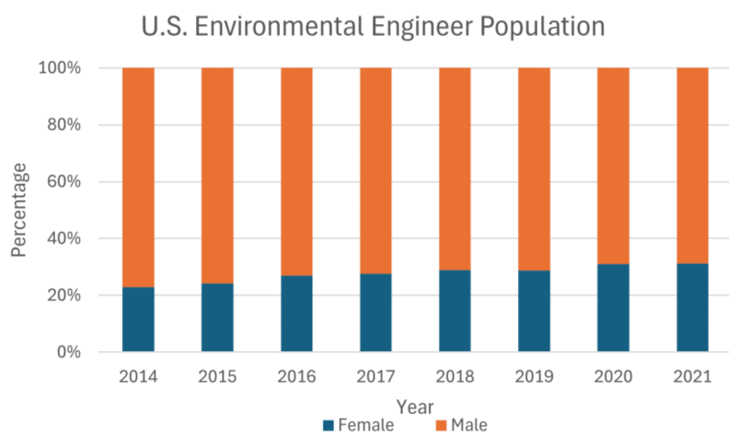


Figure 1. Gender Distribution of U.S. Environmental Engineers (2014–2021).

Understanding why women are underrepresented in engineering is complex, involving societal, institutional, and cultural factors.[4, 6, 7] However, one critical area of inquiry is the academic experience of female engineering students compared to their male counterparts. Undergraduate education serves as a pivotal stage for shaping the future composition of the engineering workforce, making it essential to explore how gender dynamics influence student performance and engagement in college engineering courses.[8]

Collaborative learning is widely regarded as a key pedagogical strategy in engineering education, reflecting the importance of teamwork as a professional competency.[9-11] However, the success of collaborative learning depends on factors such as group size, gender composition, and prior knowledge, which can significantly influence group dynamics and individual outcomes.[12-14] These variables prompt key questions regarding engineering education: 1) How does gender-based performance differ across various assessments? 2) How does the gender composition of collaborative learning groups affect individual student performance? 3) How does course performance vary across academic levels of students? Addressing these questions provides critical insights into the factors contributing to gender disparities in engineering education, informing strategies to create more equitable learning environments and, ultimately, a more diverse engineering workforce.

**Course Context**

*Environmental Engineering* is an introductory class required for undergraduate students in the Department of Civil and Environmental Engineering at the University of Wisconsin–Madison, and elective for students in other programs.. The semesters under investigation were offered by the same female instructor with over 20 years of teaching experience, ensuring consistency in instructional style and course content. The instructor has employed evidence-based teaching strategies, such as active learning and peer instruction, to enhance student engagement and comprehension, since 2012.

The course is structured as a blended learning experience that integrates both asynchronous and synchronous components (Fig. 2). For the asynchronous portion, students are required to watch short "pre-flight" instructional videos and complete associated quizzes prior to lectures. These preparations engage students with the content and support active participation during the synchronous portion, which includes 50-minute lectures held three times a week and weekly 115-min discussion sessions. These discussions focus on solving problem sets in small groups, constructed to encourage active learning. Most synchronous activities are conducted in person, with exceptions during 2020 and 2021 when the course was adapted to an online format due to the COVID-19 pandemic.

The active learning groups in this class were an integral component of the course design, fostering collaborative learning and engagement. These groups were assigned by the instructor at the beginning of each semester and remained unchanged throughout the semester to encourage consistency in group dynamics. Each group typically consisted of 3-5 students, with group composition adhering to a specific principle: avoiding situations where female students were minoritized; namely, groups were designed to ensure that if female students were present, their numbers were equal to or exceeded those of male students. This approach aimed to create a balanced and inclusive environment that supported equitable participation and minimized the potential for isolation.[15]

These groups collaborated weekly on problem sets during discussion sessions. To ensure collective accountability, each group was required to reach a consensus on the final solutions and submit one unified version for grading. All group members received the same grade for each problem set, emphasizing the importance of teamwork and shared responsibility. This structure promoted active engagement, encouraged diverse perspectives, and aligned with the course's broader objective of integrating collaborative learning into the engineering curriculum.

Weekly quizzes, administered either in class or as take-home, formed a significant component of the assessments, while major assessments in the course included two midterm exams and a final exam. In some semesters, a field project, essay, or news report was assigned to encourage students to relate class knowledge with real-world applications. Peer evaluations of group work and overall professional attitude further emphasized the importance of collaboration and professional conduct. This course design balances individual accountability with collaborative learning, ensuring that students develop both technical knowledge and essential teamwork skills necessary for success in environmental engineering.

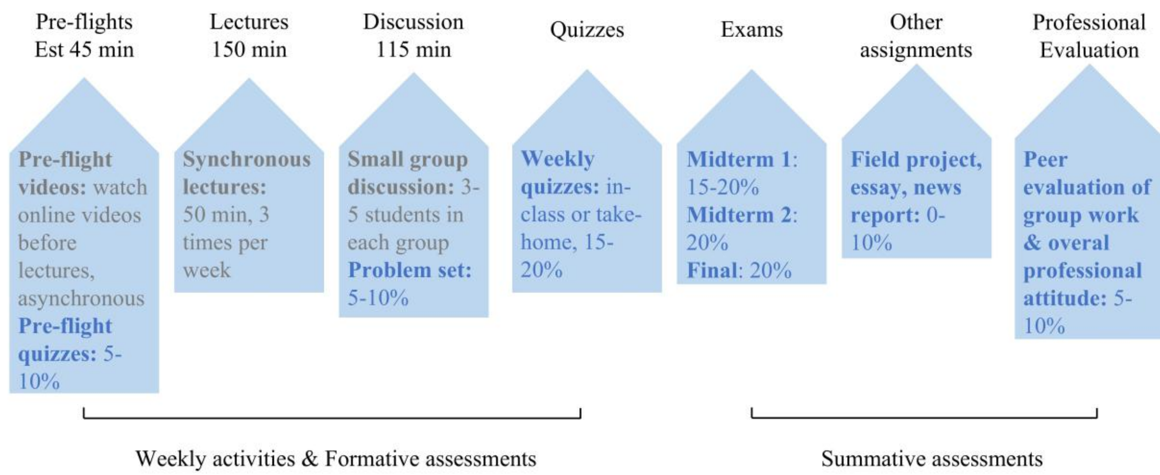


Figure 2. Course structure of *Environmental Engineering*. Activities are shown in gray, while assessments and their representative grading weights are indicated in blue.

**Methods**

This study employed a retrospective design to investigate gender-based performance differences in an undergraduate engineering class. The analysis spanned eight semesters between 2015 and 2023, encompassing a total of 598 students (Table 1). The course was designed based on evidence-based teaching practices, including active learning and collaborative group work. Quantitative data were collected from students' separate assessment grades and overall grade point average (GPA). The metrics of GPA conversion was shown in Table 2.

Table 1. Student enrollment per semester

| Year | Total Number of Student Enrolled |
|---|---|
| 2015 | 69 |
| 2016 | 66 |
| 2017 | 90 |
| 2018 | 93 |
| 2020 | 85 |
| 2021 | 78 |
| 2022 | 36 |
| 2023 | 81 |

For statistical analyses, two primary approaches were employed. First, a two-sample t-test was conducted to compare assessment performances between female and male students across all semesters, providing insights into overall gender-based performance differences. Second, analysis of variance (ANOVA) was used to evaluate the influence of group gender composition on student performance. In this analysis, the semesters in 2015 and 2023 were excluded due to changes in group composition during the semester. Group gender composition was categorized as follows: 1) female students in single-gender groups, 2) male students in single-gender groups, 3) female students in mixed-gender groups, and 4) male students in mixed-gender groups. To identify the specific differences between these gender compositions, Tukey's Honestly Significant Difference (HSD) test was applied as a post hoc analysis. Additionally, similar ANOVA and Tukey's HSD tests were implemented to examine whether students' stage of education influenced their performance outcomes, in which academic levels were categorized as sophomore, junior, senior, and graduate.

Table 2. GPA conversion scale for letter grades

| Letter Grade | GPA |
|---|---|
| A | 4.0 |
| AB | 3.5 |
| B | 3.0 |
| BC | 2.5 |
| C | 2.0 |
| D | 1.0 |
| F | 0.0 |

## Results and Discussion

Gender distribution and gender-based course performance

Throughout the eight semesters from 2015 to 2023, there was a noticeable and gradual increase in the proportion of female students, despite that male students consistently constituted the majority of the class (Fig. 3). The most pronounced growth in female representation occurred between 2018 and 2023, with 2023 marking the highest percentage of female students at 51%, doubling the 25% representation observed in 2015. This upward trend indicates a positive shift in gender diversity within the course, aligning with broader efforts to increase female participation in engineering disciplines.

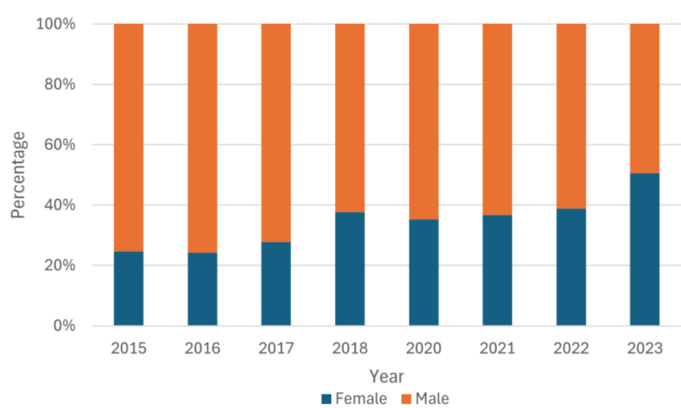


Figure 3. Gender distribution of students enrolled in *Environmental Engineering*.

Regarding gender-based course performance, the p-values from two-sample t-tests for each assessment and overall course GPA are summarized in Table 3. Foremost, there was no significant difference in GPA-based overall course performance between female and male students in seven out of eight semesters (Fig. 4). In 2015, female students achieved a significantly higher average GPA than their male counterparts ($p = 0.043$). Across all semesters, no significant differences were observed for midterm and final exams, or in-class quizzes.

In specific assessments, female students outperformed their male counterparts, including pre-flight quizzes in 2015 (female average = 95.2, male average = 90.7) and 2022 (female average = 90.3, male average = 84.5) with the overall p-value also showing significance (female average = 92.4, male average = 89.9), problem sets in 2015 (female average = 94.0, male average = 90.3) and 2021 (female average = 97.0, male average = 96.3) with total p-value across all semesters also being significant (female average = 96.4, male average = 95.5), as well as take-home quizzes in 2022 (female average = 90.3, male average = 84.5). Additionally, some differences were observed in the field project assignment in 2015 (female average = 94.2, male average = 90.2) and essay in 2018 (female average = 95.1, male average = 92.7).

Table 3. p-values from two-sample t-tests comparing course assessment performances between female and male students

| **Year** | **GPA** | **PF quizzes** | **IC quizzes** | **TH quizzes** | **Problem sets** | **MT #1** | **MT #2** | **Final** | **Others** |
|---|---|---|---|---|---|---|---|---|---|
| 2015 | 0.043 | 0.038 | 0.066 | / | 0.011 | 0.61 | 0.15 | 0.52 | 0.35 0.028 |
| 2016 | 0.46 | 0.086 | 0.87 | / | 0.17 | 0.65 | 0.18 | 0.36 | 0.13 |
| 2017 | 0.43 | 0.49 | 0.12 | / | 0.23 | 0.96 | 0.79 | 0.57 | 0.15 |
| 2018 | 0.27 | 0.15 | 0.064 | / | 0.77 | 0.55 | 0.97 | 0.36 | 0.0086 |
| 2020 | 0.56 | 0.41 | / | 0.81 | 0.54 | 0.36 | 0.63 | 0.29 | / |
| 2021 | 0.63 | 0.14 | / | 0.48 | 0.042 | 0.72 | 0.28 | 0.10 | / |
| 2022 | 0.66 | 0.048 | / | 0.021 | 0.053 | 0.34 | 0.097 | 0.75 | / |
| 2023 | 0.54 | 0.37 | 0.26 | / | 0.93 | 0.19 | 0.31 | 0.86 | / |
| Total | 0.28 | 0.0021 | 0.18 | 0.18 | 0.012 | 0.56 | 0.78 | 0.44 | / |

Note: “PF quizzes” refers to “pre-flight quizzes”; “IC quizzes” refers to “in-class quizzes”; “TH quizzes” refers to “take-home quizzes”; “MT” refers to “midterm exam”; “Final” refers to “final exam”; “others” refers to all other assignments, including essay, news report, or field project. Significant p-values ($p < 0.05$) are highlighted in red to indicate statistically significant differences.

The findings suggest that gender-based performance differences were more pronounced in formative assessments, such as preflight quizzes and problem sets than in summative assessments. Further analysis is required to evaluate the impact of time constraints on assessment performance based on gender. These findings provide a valuable context for interpreting subsequent analyses of group gender composition.

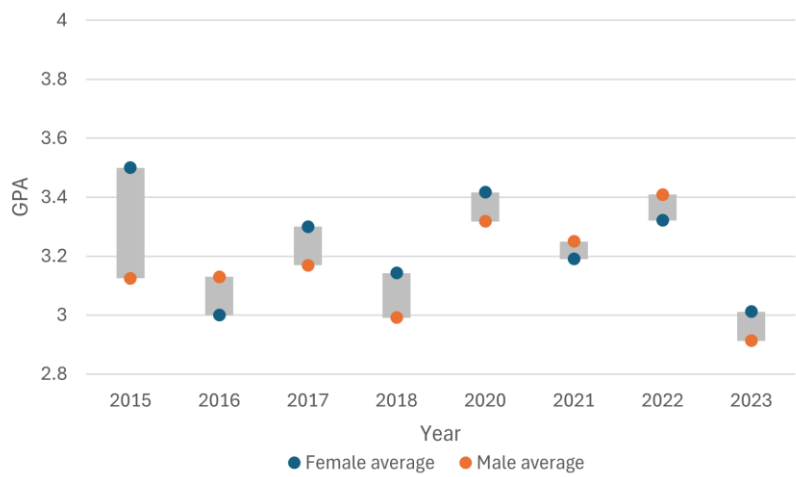


Figure 4. Average course GPA of female and male students per semester.

Influence of group gender compositions

Table 4 summarizes the p-values from ANOVA analyses of each assessment and overall course GPA among four group gender compositions: female students in single-gender groups, male students in single-gender groups, female students in mixed-gender groups, and male students in mixed-gender groups. Overall, no statistically significant differences were observed for GPA or assessments, including quizzes, midterm and final exams, across any individual semester or overall. A significant difference was observed in pre-flight quizzes across all semesters, though no single year showed significant differences. Notably, significant differences in performance were identified for problem sets, particularly in 2020 ($p = 0.027$), 2021 ($p = 0.0032$), and 2022 ($p = 0.0036$), with the total p-value across all semesters being highly significant ($p = 6.0 \times 10^{-6}$). These results suggest that group gender composition has a notable impact on performance in problem sets, which serves as the most straightforward indicator of small group collaboration.

Table 4. ANOVA p-values of course assessments among four group gender compositions

| **Year** | **GPA** | **PF quizzes** | **IC quizzes** | **TH quizzes** | **Problem sets** | **MT #1** | **MT #2** | **Final** |
|---|---|---|---|---|---|---|---|---|
| 2016 | 0.36 | 0.23 | 0.99 | / | 0.38 | 0.29 | 0.34 | 0.63 |
| 2017 | 0.61 | 0.64 | 0.30 | / | 0.47 | 0.86 | 0.65 | 0.48 |
| 2018 | 0.51 | 0.37 | 0.15 | / | 0.98 | 0.87 | 0.65 | 0.71 |
| 2020 | 0.85 | 0.51 | / | 0.54 | 0.027 | 0.45 | 0.80 | 0.78 |
| 2021 | 0.68 | 0.25 | / | 0.32 | 0.0032 | 0.76 | 0.60 | 0.30 |
| 2022 | 0.79 | 0.22 | / | 0.12 | 0.0036 | 0.50 | 0.31 | 0.64 |
| Total | 0.78 | 0.017 | 0.18 | 0.45 | $6.0\times10^{-6}$ | 0.83 | 0.28 | 0.15 |

Note: “PF quizzes” refers to “pre-flight quizzes”; “IC quizzes” refers to “in-class quizzes”; “TH quizzes” refers to “take-home quizzes”; “MT” refers to “midterm exam; “Final” refers to “final exam”. Significant p-values ($p < 0.05$) are highlighted in red to indicate statistically significant differences.

Tukey’s HSD test was applied post hoc to identify the specific differences between these gender compositions (Fig. 5). The results showed that female students in single-gender groups exhibited the lowest problem set scores, which were significantly lower than those female students in mixed-gender groups ($p = 6.8 \times 10^{-6}$), male students in mixed-gender groups ($p = 0.00015$), and male students in single-gender groups ($p = 0.0032$). This suggests that group gender composition strongly influences female students' performance on problem sets, with mixed-gender settings being more favorable. In contrast, male students showed no significant differences in problem set scores between single-gender and mixed-gender groups ($p = 0.27$). Notably, the differences between female and male students in mixed-gender groups were also insignificant ($p = 0.94$), indicating that mixed-gender groups promote equitable performance.

Given that no significant differences were observed in exams, quizzes, and GPA among the four gender composition groups, these findings suggest that while group gender composition influences performance on collaborative problem sets, it does not affect individual assessments. This distinction highlights the importance of group dynamics in collaborative learning and underscores the potential benefits of fostering mixed-gender group environments to support female students' performance in collaborative tasks. It is also worth noticing that the significant differences in problem set performance started in 2020, coinciding with the onset of the COVID-19 pandemic. Further research is needed to explore how the pandemic might relate with changes in group dynamics.

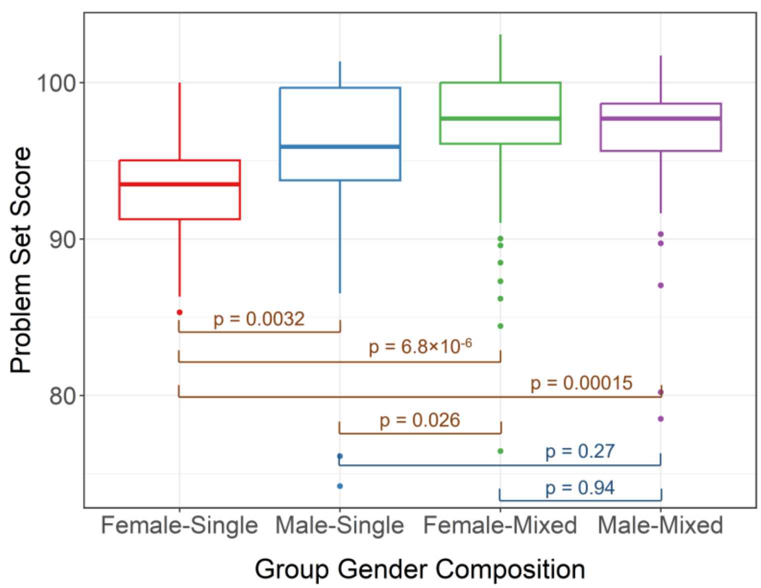


Figure 5. Problem set score distribution across different group gender compositions

Influence of academic level

The influence of prior knowledge on student performance was also investigaed using ANOVA and Tukey's HSD tests. The data reveal significant differences in course GPA between certain academic levels, reflecting the influence of academic progression on student performance (Fig. 6). Sophomore students exhibited lower GPAs compared to seniors ($p = 0.0049$) and graduate students ($p = 0.038$), suggesting that higher academic levels correspond with improved performance. This trend may indicate the cumulative benefits of academic experience, such as improved study habits, greater familiarity with engineering concepts, and increased engagement with course material. In contrast, no significant difference in GPA was observed between sophomores and juniors (p = 0.23), nor among juniors, seniors, and graduate students (p = 0.25, p = 0.24, and p = 0.82, respectively).

These results suggest that the most substantial improvements in GPA occur later in students' academic journeys, likely between the sophomore and senior years, when foundational skills are solidified, and students engage with advanced engineering coursework. These results underscore the importance of tailoring teaching strategies to meet the developmental needs of underclassmen taking a required course relatively early. Such populations may benefit from additional support to enhance their academic outcomes.

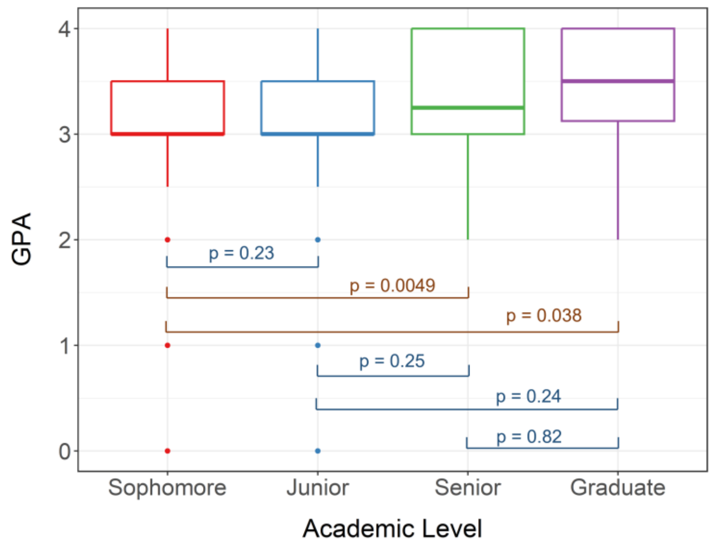


Figure 6. Course GPA distribution across academic levels

**Limitations and Implications**

Limitations

This study has several limitations that should be considered when interpreting the results. First, gender information was inferred by the instructor rather than self-reported by students, which may introduce inaccuracies. Additionally, gender was categorized using a binary system, excluding non-binary or gender-diverse identities, which could have enriched the analysis. Last but no least, the 2020 semester also presents a unique challenge as an outlier due to the lack of in-person interactions caused by the COVID-19 pandemic, potentially impacting collaborative learning and group dynamics in unprecedented and ongoing ways.

Implications

Despite the limitations, this study offers valuable insights into gender dynamics and academic performance in an engineering classroom. One key finding is that there is no overall gender gap in the performance of this course. Across all semesters, no significant differences were observed in time-constrained assessments, such as quizzes and exams, indicating parity in performance in individual, high-stakes testing scenarios. Interestingly, in some semesters, female students outperformed their male counterparts in time-unconstrained assessments, including pre-flight quizzes, and essay or field project assignments. This finding suggests that female students may excel in tasks that allow for reflective thinking and extended engagement with course materials.

The study also underscores the potential benefits of mixed-gender group settings for female students, aligning with findings from previous research. In a qualitative study on collaborative group dynamics of sixth- and seventh-grade students in an engineering afterschool program, females in mixed-gender groups were found to achieve better learning outcomes compared to those in single-gender groups.[16] Similarly, in a business simulation setting, the all-female teams were outperformed by any other gender combination.[17] These differences have been attributed to variations in work dynamics and interactional styles between females in mixed- and single-gender settings.[16, 17] This was further evidenced by the study conducted under anonymous gender composition, where the all-female teams exhibited the highest productivity. Moreover, a hidden positive effect of female presence on the productivity of other team members was highlighted under such conditions.[18] Additionlly, it was also observed that both female and male students in mixed-gender groups performed better than their peers in single-gender groups.[9] Overall, cooperative learning environments alone do not necessarily lead to improved performance;[19] rather, it is the positive and supportive teamwork experiences that are more likely to contribute to students' success.[20] Factors independent of gender, such as communication styles or leadership roles, should be considered to avoid reinforcing stereotypes.[13, 15]

Academic progression also plays a significant role in performance. Students with more than two years of undergraduate prerequisites tended to perform better, likely due to the foundational skills and knowledge acquired over time. These findings suggest that targeted support for underclassmen could help bridge the gap in performance and better prepare students for advanced coursework.

This study raises important questions for future research:

1. **Group Dynamics and Self-Efficacy**: How do group dynamics and individual experiences interact with self-efficacy and performance, and how might deliberate facilitation of group performance by the instructor impact the outcomes? Furthermore, how might the course experiences shape students' future career decisions in engineering?
2. **Prior Knowledge**: What specific prior courses or knowledge have the greatest impact on student performance in this course? Identifying these key factors could guide curriculum design and preparatory coursework.
3. **Pandemic Effects**: How did the COVID-19 pandemic alter students' study patterns and collaborative behaviors? Understanding these changes could inform strategies to support learning in hybrid or remote environments.

These questions provide a roadmap for further exploration to enhance gender equity and optimize learning outcomes in engineering education.

**Acknowledgements**

The authors would like to thank the startup fund from the Department of Civil and Environmental Engineering, College of Engineering, the Office of the Vice Chancellor for Research and Graduate Education (OVCRGE) at the University of Wisconsin–Madison, and the Wisconsin Alumni Research Foundation (WARF) for the support. Additional support was provided by the National Science Foundation (CBET-2132026). This project was conducted through the Delta Program in Research, Teaching, and Learning at UW-Madison. We thank the Associate Directors of Delta, Dr. Jessica Maher and Dr. Devin Wixon, for valuable feedback on study design, as well as Prof. Haoran Wei for supporting Hanwei.

## References

[1] J. S. Long and M. F. Fox, "Scientific Careers - Universalism and Particularism," *Annu. Rev. Sociol.*, vol. 21, pp. 45-71, 1995, doi: DOI 10.1146/annurev.so.21.080195.000401.

[2] G. Sonnert and G. J. Holton, *Who Succeeds in Science? The Gender Dimension*. New Brunswick, N.J: Rutgers University Press, 1995.

[3] G. Sonnert and G. J. Holton, *Gender differences in science careers : the project access study*. New Brunswick, N.J: Rutgers University Press, 1995.

[4] T. E. S. Charlesworth and M. R. Banaji, "Gender in Science, Technology, Engineering, and Mathematics: Issues, Causes, Solutions," *J. Neurosci.*, vol. 39, no. 37, pp. 7228-7243, Sep 11 2019, doi: 10.1523/Jneurosci.0475-18.2019.

[5] U.S. Census Bureau. *American Community Survey (ACS) Public Use Microdata Sample (PUMS) 5-Year Estimate*. [Online]. Available: https://www.census.gov/programs-surveys/acs/microdata.html

[6] M. B. Ross, B. M. Glennon, R. Murciano-Goroff, E. G. Berkes, B. A. Weinberg, and J. I. Lane, "Women are credited less in science than men," *Nature*, vol. 608, no. 7921, pp. 135-+, Aug 4 2022, doi: 10.1038/s41586-022-04966-w.

[7] Y. Yang, T. Y. Tian, T. K. Woodruff, B. F. Jones, and B. Uzzi, "Gender-diverse teams produce more novel and higher-impact scientific ideas," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 119, no. 36, Sep 6 2022, doi: 10.1073/pnas.2200841119.

[8] G. Sonnert and M. F. Fox, "Women, Men, and Academic Performance in Science and Engineering: The Gender Difference in Undergraduate Grade Point Averages," *J. High. Educ.*, vol. 83, no. 1, pp. 73-+, Jan-Feb 2012, doi: 10.1080/00221546.2012.11777235.

[9] D. Gnesdilow, A. Siebert-Evenstone, J. Rutledge, S. Jung, and S. Puntambekar, "Group Work in the Science Classroom: How Gender Composition May Affect Individual Performance," in *International Conference on Computer Supported Collaborative Learning*, Madison, WI, 2013, vol. 2: International Society of the Learning Sciences, doi: 10.13140/2.1.1718.5285.

[10] A. A. Niler, R. Asencio, and L. A. DeChurch, "Solidarity in STEM: How Gender Composition Affects Women's Experience in Work Teams," *Sex Roles*, vol. 82, no. 3-4, pp. 142-154, Feb 2020, doi: 10.1007/s11199-019-01046-8.

[11] J. M. Buckley, S. Grajeda, A. E. Trauth, and D. P. Roberts, "A Framework for Quantifying Student Self-Confidence and Task Choice in Engineering Design-related Activities," in *ASEE Annual Conference*, Tampa, FL, 2019: American Society for Engineering Education, doi: 10.18260/1-2--31957.

[12] Q. N. Feng, H. Luo, W. H. Li, T. J. Chen, and N. N. Song, "Effects of gender diversity on college students' collaborative learning: From individual gender to gender pairing," *Heliyon*, vol. 9, no. 6, Jun 2023, doi: 10.1016/j.heliyon.2023.e16237.

[13] P. L. Curseu, M. M. H. Chappin, and R. J. G. Jansen, "Gender diversity and motivation in collaborative learning groups: the mediating role of group discussion quality," *Soc. Psychol. Educ.*, vol. 21, no. 2, pp. 289-302, Apr 2018, doi: 10.1007/s11218-017-9419-5.

[14] I. A. G. Wilkinson and I. Y. Y. Fung, "Small-group composition and peer effects," *Int. J. Educ. Res.*, vol. 37, no. 5, pp. 425-447, 2002/01/01/ 2002, doi: 10.1016/S0883-0355(03)00014-4.

[15] S. Ingram and A. Parker, "The influence of gender on collaborative projects in an engineering classroom," *IEEE T. Prof. Commun.*, vol. 45, no. 1, pp. 7-20, Mar 2002, doi: 10.1109/47.988359.

[16] J. Schnittka and C. G. Schnittka, ""Can I Drop It This Time?" Gender and Collaborative Group Dynamics in an Engineering Design-Based Afterschool Program," *J. Pre-Coll. Eng. Educ. Res.*, vol. 6, no. 2, pp. 1-24, 2016.

[17] J. Apesteguia, G. Azmat, and N. Iriberri, "The Impact of Gender Composition on Team Performance and Decision Making: Evidence from the Field," *Manage. Sci.*, vol. 58, no. 1, pp. 78-93, Jan 2012, doi: 10.1287/mnsc.1110.1348.

[18] H. G. Song, M. Restivo, A. van de Rijt, L. Scarlatos, D. Tonjes, and A. Orlove, "The hidden gender effect in online collaboration: An experimental study of team performance under anonymity," *Comput. Hum. Behav.*, vol. 50, pp. 274-282, Sep 2015, doi: 10.1016/j.chb.2015.04.013.

[19] S. Takeda and F. Homberg, "The effects of gender on group work process and achievement: an analysis through self- and peer-assessment," *Brit. Educ. Res. J.*, vol. 40, no. 2, pp. 373-396, Apr 2014, doi: 10.1002/berj.3088.

[20] K. Miller, G. Kestin, and O. Miller, "How gender composition and group formation impact the effectiveness of group work in two-stage collaborative exams," *Phys. Rev. Phys. Educ. R.*, vol. 18, no. 2, Nov 28 2022, doi: 10.1103/PhysRevPhysEducRes.18.020137.