Pre- and Post-AI Adoption: Performance in Introductory Online EngineeringCourses

Dr. Ghazal Barari, Embry-Riddle Aeronautical University

Dr. Ghazal Barari is an Associate Professor of Engineering at Embry-Riddle Aeronautical University Worldwide. With a Ph.D. in Mechanical Engineering from the University of Central Florida, her research pivoted from combustion and biofuel's reaction mechanism to engineering education and student success pathways. She studies AI applications in curriculum design, simulation-based learning, and students' engagement in online environments. She leads the Engineering Fundamentals program's ABET accreditation, serves as an editor for International Journal of Aviation, Aeronautics, and Aerospace, and has been recognized for her innovative teaching and service.

Julianna Gesun, Embry–Riddle Aeronautical University, Daytona Beach

Julianna Gesun, Ph.D., is currently a postdoc researcher at Embry-Riddle. Her research broadly focuses on understanding and supporting the process by which engineering programs facilitate the environments for students to develop optimal functioning in undergraduate engineering programs. Her research interests intersect the fields of positive psychology, engineering education, and human development to understand the intrapersonal, cognitive, social, behavioral, contextual, cultural, and outcome factors that influence thriving in engineering.

Prior to joining Embry-Riddle, she was a National Science Foundation/American Society for Engineering Education engineering postdoctoral fellow at the University of New Hampshire. She received her Ph.D. in Engineering Education at Purdue University, where she was an NSF Graduate Research Fellow and the winner of Purdue's 2021 Three Minute Thesis competition for her work in developing research and courses on engineering thriving. She also received dual bachelor's degrees in Industrial Engineering and Human Development and Family Studies at the University of Illinois at Urbana-Champaign. Her prior work experiences include product management, consulting, tutoring, marketing, and information technology.

Pre- and Post-AI Adoption: Performance in Introductory Online Engineering Courses

Introduction

This brief research paper examines whether the widespread availability of generative artificial intelligence (AI) corresponds with observable shifts in student performance in an online engineering gateway course. Generative AI has rapidly become embedded in students' academic workflows, particularly in online learning environments. Following the widespread release of large language models, students have increasingly used AI tools to plan, draft, revise, and refine coursework, often without direct instructional guidance. These tools are especially accessible in online engineering courses and take-at-home exams, where assessments are frequently un-proctored, flexible, and emphasize advisor evaluations [1], [2]. As a result, generative AI represents a structural change in the learning environment rather than an instructional innovation. Thus, understanding the intersections between this shift and student performance is an increasing research need in engineering education.

Despite growing attention to generative AI, empirical evidence linking independent student AI use to performance outcomes remain limited. Much of the existing discourse emphasizes academic integrity risks, ethical ambiguity, content accuracy, and overreliance [3]-[5], while empirical studies often focus on structured or instructor-sanctioned AI use [6]-[11]. At the same time, campus- and course-level surveys suggest that most students self-teach AI workflows and selectively apply them across assignments, resulting in highly variable and individualized usage patterns [12]. In engineering education, experiential and technology-enhanced approaches have been shown to enhance engagement and workforce relevance [13]-[16]. Advances in continual and representation learning demonstrate how AI systems can incrementally acquire knowledge while mitigating catastrophic forgetting, supporting adaptive learning processes [17]-[20].

Un-proctored assessments in online engineering courses provide a particularly important and understudied context for examining AI-related performance patterns in engineering education [21]. At our university, assignments such as discussions, take-home work, and multi-week projects are commonly used to assess higher-order thinking, iteration, and collaboration. Prior research suggests that scaffolded and collaborative assessments such as engineering capstone courses yield stable performance distributions over time, and better engagement [2], [22]-[25]. However, few studies have directly examined whether student performance on such assessments changed following the widespread availability of generative AI tools [26]-[31]. Therefore, comparing pre-AI and post-AI performance on comparable un-proctored assessments offers a useful approach for assessing whether AI availability corresponds with meaningful shifts in academic achievement.

These findings indicate that generative AI is already influencing student behavior, yet its relationship to measurable academic outcomes remains underexplored. This gap highlights the need for performance-focused analyses grounded in real coursework rather than hypothetical or experimental settings. In this research brief, we address these gaps by examining student performance in online engineering courses before and after the widespread emergence of generative AI. Specifically, the study investigates two research questions: (1) How did performance on un-proctored activities change from a pre-AI period (2022–2023) to a post-AI period (2024–2025)? and (2) How do performance outcomes, including percentage scores and mastery rates (≥80%), compare across cohorts defined by AI availability? This study does not directly measure individual AI usage; instead, AI availability serves as a contextual cohort distinction. Outcomes are further disaggregated by student demographics and prior achievement to examine subgroup-specific patterns. By grounding the analysis in authentic course data and student self-reports, this study reports empirical evidence on the alignments between independent generative AI use and academic performance in online engineering education. Building on this context, the present study examines whether the widespread availability of generative AI corresponds with observable changes in student performance in an authentic online engineering course setting.

Using an introductory engineering course as a case study, student outcomes on a multi-week, un-proctored team project are compared across course offerings delivered before and after the emergence of generative AI. By leveraging naturally occurring course data and focusing on distributional patterns, benchmark attainment, and subgroup-level differences, this work provides an empirical, performance-based perspective on AI's alignment with learning outcomes in online engineering education. The following sections describe the study design and analytical approach, present comparative results across cohorts, and discuss implications for assessment and instructional practice.

Methodology

This study compares student performance in ENGR 101 before and after the widespread emergence of generative AI tools (*pre-AI* vs. *post-AI* course offerings). The focal assessment is a multi-week team (group) project completed over four weeks. Student teams progressed through staged deliverables, submitting weekly components (milestones) and a final project submission at the end of the four-week period. The analysis reported here uses each student's final project score (the summative score recorded for the team project) as the outcome measure.

A quasi-experimental, comparative cohort design was used. Two naturally occurring cohorts were defined based on term timing relative to the emergence of generative AI at the end of 2023: Pre-AI cohort (2022–23) and Post-AI cohort (2024–25).While the pre-AI cohort (2022–2023) followed the immediate COVID-19 pandemic disruption period, ENGR 101 was delivered in a , fully online format in both cohorts using the same project prompt, milestone structure, grading rubric, and instructional team. No substantial policy or structural assessment changes occurred between terms. Thus, while broader shifts in online learning habits may have continued to evolve post-pandemic, the focal assessment conditions remained consistent across cohorts.

Final project scores and student demographic attributes were extracted from course reporting exports for the pre-AI and post-AI terms. The project was open-note and open-internet in both cohorts, with no proctoring mechanisms in place. Records were cleaned by retaining valid numeric scores for the final project outcome, standardizing demographic labels (e.g., gender categories; age groups as Gen X, Millennial, Gen Z; ethnicity categories), and ensuring one final score per student record (the final score used for all comparisons). Analyses focused on describing and comparing performance distributions across cohorts using four complementary views:

1. **Score-range distribution (performance bands):** Final scores were grouped into score ranges (bins) to compare how students were distributed across performance levels pre-AI vs. post-AI. Percent-of-students (or percent-of-records, depending on the report export) in each band was plotted to highlight shifts in the distribution rather than only changes in averages.
2. **Boxplot comparison (distributional change):** Boxplots were used to compare the median, interquartile range (IQR), and outliers of final scores between pre-AI and post-AI cohorts. This visual emphasizes changes in spread and the presence of unusually low/high scores.
3. **Threshold analysis (≥80% benchmark):** A criterion-referenced comparison was performed by calculating the percentage of students scoring ≥80% in each cohort. This supports interpretation in terms of attainment of a common performance benchmark.
4. **Disaggregation by student characteristics (equity-focused comparisons):** The ≥80% metric was further disaggregated and compared across Gender (Male, Female), Age group (Gen X, Millennial, Gen Z), and Ethnicity (White, Hispanic/Latino, Black/African American, Asian, and Other where remaining categories were combined to maintain interpretability and adequate group sizes).

These subgroup comparisons were used to examine whether the pre/post patterns were consistent across groups or concentrated in specific populations. Additionally, some variables (e.g., enrollment composition, instructor experience, evolving student familiarity with online collaboration tools) may also contribute to observed patterns. Future studies can expand to include data on these additional variables.

Results and Discussion

This section presents a comparative analysis of ENGR 101 group-project performance across course offerings before and after the widespread emergence of generative AI. Using students' final project scores, results are summarized through the following graphs:

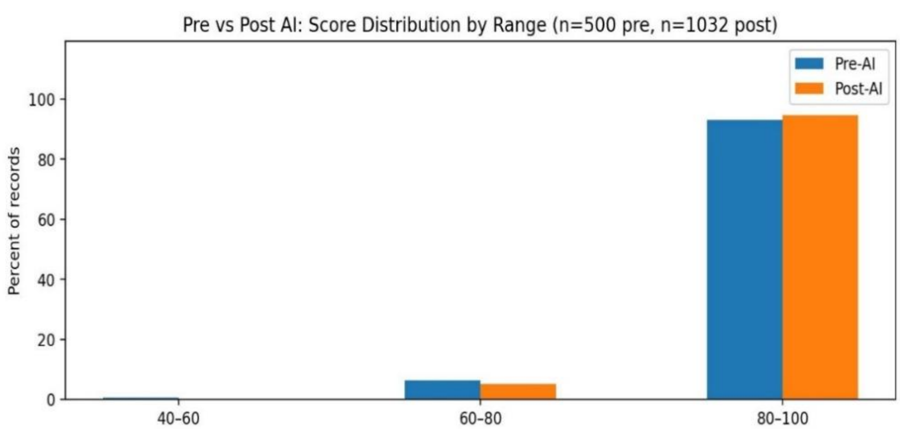

Figure 1: Distribution of ENGR 101 team project scores before vs. after AI emergence across score ranges (40–60, 60–80, 80–100), reported as the percentage of records in each bin (with sample sizes noted on the figure).

Figure 1 summarizes the proportion of scores in the upper performance bands (40–60, 60–80, 80–100). In both cohorts, the vast majority of scores fall in the 80–100 band. The 60–80 band constitutes a small share in both cohorts and appears slightly smaller post-AI, while 40–60 is minimal in both cohorts. This figure visually reinforces that the overall distribution is heavily concentrated at the high end in both periods, with only minor changes pre-to-post AI.

Figure 2 provides a distribution-level comparison of final scores. The medians for pre-AI and post-AI are both high and appear very similar, indicating that typical performance did not substantially change. The IQRs are also comparable, suggesting similar score spread among the middle 50% of students across cohorts. Both cohorts include outliers, including low-score outliers, indicating that a small subset of students performed substantially below the typical range in both periods. Overall, the boxplot supports the conclusion that post-AI performance is distributionally similar to pre-AI, with no major shift in central tendency.

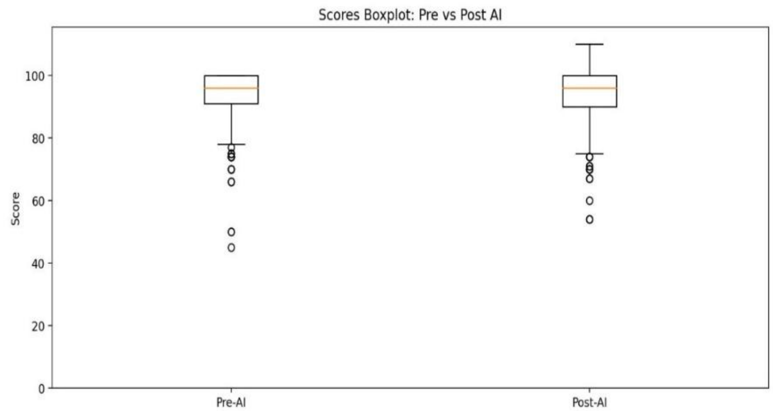

Figure 2: Boxplot comparison of ENGR 101 team project score distributions before vs. after AI emergence, showing median, interquartile range, and outliers.

Figure 3 compares the percent ≥80% for Male and Female students. For Male students, the post-AI cohort shows a clear increase relative to pre-AI. For Female students, the ≥80% rate remains high in both cohorts but shows a slight decrease post-AI relative to pre-AI. Importantly, both genders remain at high attainment overall; the key observation is that the direction of change differs by gender (increase for males, small decrease for females), which is worth noting when interpreting overall cohort shifts. One possible explanation for the divergent pattern between male and female students may relate to differences in technology adoption patterns or differential confidence in leveraging AI tools within collaborative settings. Alternatively, team role dynamics may influence how AI-supported contributions are distributed. These causal explanations warrant further investigation.

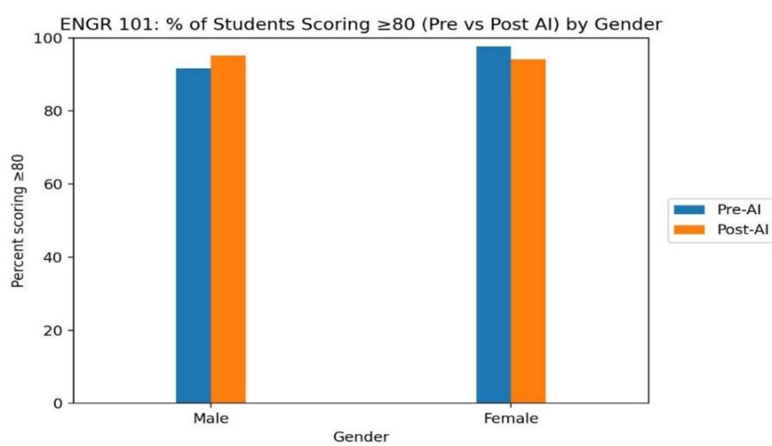

Figure 3: Percentage of ENGR 101 students achieving scores ≥ 80% in the project pre- and post AI, disaggregated by gender

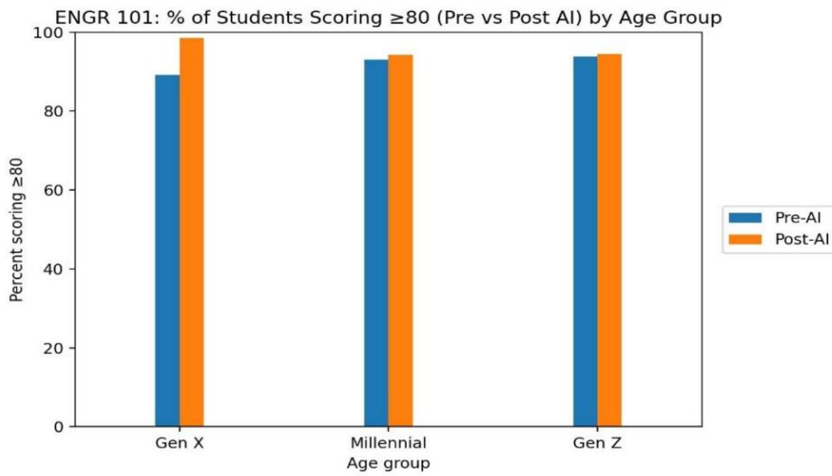

Figure 4: Percentage of ENGR 101 students achieving scores ≥ 80% in the project pre- and post AI, disaggregated by *age group*

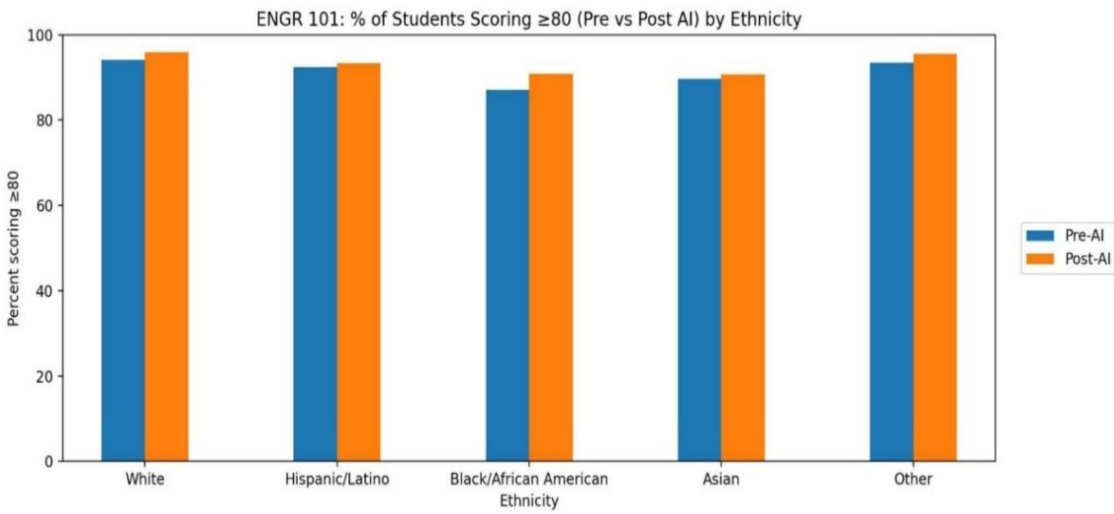

Figure 5: Percentage of ENGR 101 students achieving scores ≥ 80% in the team project before vs. after AI emergence, disaggregated by *ethnicity* (White, Hispanic/Latino, Black/African American, Asian, and Other).

Figure 4 shows the ≥80% benchmark by age group. All three age groups exhibit high attainment in both cohorts. The most pronounced change is for Gen X, where the post-AI cohort reaches an especially high (near-ceiling) ≥80% rate compared to pre-AI. Millennials and Gen Z show small increases post-AI (or essentially stable levels), indicating that benchmark attainment is broadly consistent across age groups, with no evidence of a post-AI drop in performance.

Figure 5 compares the percentage of students scoring ≥80% in the pre-AI and post-AI cohorts across the five ethnicity groupings (White, Hispanic/Latino, Black/African American, Asian, and Other). Overall, the pattern is consistently high attainment in both cohorts, with post-AI generally matching or slightly exceeding pre-AI for most categories. The largest visible gain occurs for Black/African American students, where the post-AI ≥80% rate is noticeably higher than pre-AI. White and "Other" groups also show small increases post-AI. Hispanic/Latino and Asian show modest upward shifts or near-stability. Taken together, Figure 5 suggests that the "≥80% attainment" benchmark remained strong across ethnicities and did not decline after AI became widely available.

The slight changes observed across cohorts suggests that certain assessment features may be more resilient in AI-rich environments. These include multi-week scaffolding, iterative milestone feedback, team accountability structures, and evaluation criteria emphasizing integration rather than isolated answers. Assessments structured around synthesis, coordination, and documentation may be less susceptible to single-instance AI substitution compared to isolated take-home tasks.

One interpretation of the observed slight changes is that students optimize toward grading criteria rather than deeply engaging with learning objectives. If AI tools assist in meeting rubric expectations without altering final scores, performance metrics alone may be insufficient to capture shifts in cognitive processing or skill development. It is worth to mention that performance equivalence does not necessarily imply learning equivalence. AI-assisted workflows may alter how tasks are completed (e.g., delegation of drafting, revision, or brainstorming), even when final scores remain stable. Future research should incorporate process-tracing methods, rubric-level subscore analysis, or artifact comparison to examine potential shifts in cognitive engagement.

Conclusion

This work examines whether student performance in ENGR 101 changed after the widespread emergence of generative AI by comparing cohort outcomes from course offerings delivered before and after AI became broadly accessible. Performance is evaluated using students' final scores on a four-week, scaffolded team project with weekly milestones culminating in a final submission. The analysis focuses on overall shifts in score distributions and benchmark attainment (≥80%), and then explores whether observed pre/post patterns differ across key student subgroups, including gender, age group and ethnicity. Overall, the comparison of pre-AI and post-AI ENGR 101 cohorts indicates the availability of generative AI did not correspond with a decline in final group-project performance. Score distributions remained strongly concentrated in the 80–100 range, and benchmark attainment (≥80%) was stable to modestly higher in the post-AI cohort, with subgroup patterns suggesting small gains for some populations and minor shifts for others.

It is important to note that individual AI usage was not directly measured, the observed stability reflects cohort-level comparisons rather than isolated effects attributable solely to AI adoption. Performance equivalence does not necessarily imply equivalence in learning processes. AI-assisted workflows may alter how tasks are completed, even when final grades remain similar, suggesting the need for future research that incorporates process-level measures, student surveys, artifact analysis, or rubric-level sub score comparisons.

Future work should extend beyond a single team-based project to examine multiple assessment types, including individual assignments, exams, laboratories, written reflections, and programming or design artifacts, to determine whether cohort-level performance stability persists across tasks with varying cognitive demands and collaboration structures. Incorporating both proctored and unproctored formats may help distinguish shifts in learning and skill development from differences in assessment conditions, particularly in environments where AI-enabled support is more accessible. Expanding the analysis across additional courses and instructors would strengthen external validity, while pairing performance data with process-oriented measures, such as student surveys, artifact comparisons, or rubric-level subscore analysis, would provide insight into how AI-supported workflows shape engagement and learning in evolving technological contexts.

References

[1] Lande, M. (2023, June). Roles for Take-Home Exams from the Perspective of Engineering Students and Instructors. In *2023 ASEE Annual Conference & Exposition*.

[2] Znidi, F., Uddin, M. N., & Morsy, M. (2024, April). Evolving engineering education: online vs. in-person capstone projects compared (EEE-OIPC). In *Frontiers in Education* (Vol. 9, p. 1403781). Frontiers Media SA.

[3] Barari, G. Basu, D. Lin Y., and Ortega-Moody, J. (2026). Evaluating Large Language Models in Engineering Education: A Multi-Course Assessment Across Bloom's Taxonomy, in *IEEE Access*, doi: 10.1109/ACCESS.2026.3686339.

[4] Bittle, K., & El-Gayar, O. (2025). Generative AI and academic integrity in higher education: A systematic review and research agenda. *Information*, *16*(4), 296.

[5] Dempere, J., Modugu, K., Hesham, A., & Ramasamy, L. K. (2023, September). The impact of ChatGPT on higher education. In *Frontiers in Education* (Vol. 8, p. 1206936). Frontiers Media SA.

[6] Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big?🦜. In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623).

[7] Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, *103*, 102274.

[8] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators?. *International journal of educational technology in higher education*, *16*(1), 1-27.

[9] Currie, G. M. (2023, September). Academic integrity and artificial intelligence: is ChatGPT hype, hero or heresy?. In *Seminars in Nuclear medicine* (Vol. 53, No. 5, pp. 719-730). WB Saunders.

[10] Khatri, B. B., & Karki, P. D. (2023). Artificial intelligence (AI) in higher education: Growing academic integrity and ethical concerns. *Nepalese Journal of Development and Rural Studies*, *20*(01), 1-7.

[11] Uzun, L. (2023). ChatGPT and academic integrity concerns: Detecting artificial intelligence generated content. *Language Education and Technology*, *3*(1).

[12] Ovi, J. A., Fierro, G., & Smith, C. (2025). Assessing Student Adoption of Generative Artificial Intelligence across Engineering Education from 2023 to 2024. arXiv preprint arXiv:2503.04696.

[13] Janke, C., Lin, Y., & Barari, G. (2025). BOARD# 15: From drones to airplanes, Lessons learned from uncrewed aviation for teaching innovative and sustainable electric aviation propulsion. In *Proceedings of the 2025 ASEE Annual Conference & Exposition*. American Society for Engineering Education.

[14] Ward, T., Jenab, K., Ortega-Moody, J., Barari, G., & Molina Acosta, L. D. C. (2025). Virtual classrooms, real impact: A framework for introducing virtual reality to K–12 STEM learning based on best practices. *Applied Sciences*, *15*(21), 11356. https://doi.org/10.3390/app152111356

[15] Siebold, K., & Barari, G. (2026, February). Enhancing Undergraduate Engineering Education through HLA-Based Simulation Exploration Experience (SEE). In *2026 CIEC*.

[16] Barari, Ghazal, and Karl Siebold. "Curriculum Alignment through the Industry Pulse Platform." In *2026 CIEC*. 2026.

[17] Barari, N., & Barari, G. (2025). Continual learning models in aviation systems. *The Collegiate Aviation Review International*, *43*(1).

[18] Barari, N., & Kim, E. (2021, November). Linking sparse coding dictionaries for representation learning. In 2021 International Conference on Rebooting Computing (ICRC) (pp. 84-87). IEEE.

[19] Barari, N., Lian, X., & MacLellan, C. J. (2024). Incremental concept formation over visual images without catastrophic forgetting. *arXiv preprint arXiv:2402.16933*.

[20] Barari, N., Kim, E., & MacLellan, C. (2025). Explaining Robustness to Catastrophic Forgetting Through Incremental Concept Formation. *arXiv preprint arXiv:2510.23756*.

[21] Nguyen, H. T. T. (2023). Unproctored assignment-based online assessment in higher education: Stakeholder evaluation of issues. *Issues in Educational Research*, *33*(1), 207-226.

[22] Barari, G., & Sanders, B. (2025, June). Application of FLASH to Forecast Student Engagement in Online Engineering Courses. In *2025 ASEE Annual Conference & Exposition*.

[23] Barari, G., & Sanders, B. (2024, February). The effectiveness of virtual environments for increasing engagement in engineering technology courses. In *2024 CIEC*.

[24] Barari, G., Janke, C., & Lin, Y. (2026). Integrating T-shaped competency models in interdisciplinary aviation and engineering education. Collegiate Aviation Review International, 44(1), 136–159. Retrieved from https://ojs.library.okstate.edu/osu/index.php/CARI/article/view/10584/9344
[25] Mishur, J., Niemi, S. R., Underhill, J. L., & Traum, M. J. (2023, October). Impact of Online Versus In-Person Group Work on Behavioral Peer Evaluations in Engineering Capstone Team Projects. In *ASME International Mechanical Engineering Congress and Exposition* (Vol. 87653, p. V008T09A027). American Society of Mechanical Engineers.
[26] Bower, M., Torrington, J., Lai, J. W., Petocz, P., & Alfano, M. (2024). How should we change teaching and assessment in response to increasingly powerful generative Artificial Intelligence? Outcomes of the ChatGPT teacher survey. *Education and Information Technologies*, *29*(12), 15403-15439.
[27] Khlaif, Z. N., Ayyoub, A., Hamamra, B., Bensalem, E., Mitwally, M. A., Ayyoub, A., ... & Shadid, F. (2024). University teachers' views on the adoption and integration of generative AI tools for student assessment in higher education. *Education Sciences*, *14*(10), 1090.
[28] Gundu, T. (2024). Strategies for e-Assessments in the Era of Generative Artificial Intelligence. *Electronic Journal of e-Learning*, *22*(7), 40-50.
[29] Smolansky, A., Cram, A., Raduescu, C., Zeivots, S., Huber, E., & Kizilcec, R. F. (2023, July). Educator and student perspectives on the impact of generative AI on assessments in higher education. In *Proceedings of the tenth ACM conference on Learning@ Scale* (pp. 378-382).
[30] Mao, J., Chen, B., & Liu, J. C. (2024). Generative artificial intelligence in education and its implications for assessment. *TechTrends*, *68*(1), 58-66.
[31] Khlaif, Z. N., Alkouk, W. A., Salama, N., & Abu Eideh, B. (2025). Redesigning assessments for AI-enhanced learning: A framework for educators in the generative AI era. *Education Sciences*, *15*(2), 174.