Impact of the FossilSketch Application on Students' Use of Generative Artificial Intelligence for Conceptual Assessment

Dr. Saira Anwar, Texas A&M University

Saira Anwar is an Assistant Professor at the Department of Multidisciplinary Engineering, Texas A and M University, College Station. She received her Ph.D. in Engineering Education from the School of Engineering Education, Purdue University, USA. The Department of Energy, National Science Foundation, and industry sponsors fund her research. Her research potential and the implication of her work are recognized through national and international awards, including the 2023 NSTA/NARST Research Worth Reading award for her publication in the Journal of Research in Science Teaching, 2023 New Faculty Fellow award by IEEE ASEE Frontiers in Education Conference, 2022 Apprentice Faculty Grant award by the ERM Division, ASEE, and 2020 outstanding researcher award by the School of Engineering Education, Purdue University. Dr. Anwar has over 20 years of teaching experience at various national and international universities, including the Texas A and M University - USA, University of Florida - USA, and Forman Christian College University - Pakistan. She also received outstanding teacher awards in 2013 and 2006. Also, she received the "President of Pakistan Merit and Talent Scholarship" for her undergraduate studies.

Dr. Daniel Sungje Bahng, Texas A&M University

Daniel Bahng received his PhD in Interdisciplinary Engineering at Texas A&M University's Department of Multidisciplinary Engineering. He also earned his degrees in Bachelor of Science and Master of Science in Computer Engineering from Texas A&M University's Department of Computer Science and Engineering. His research interests are in large language models, identity theory, and engineering education.

Anna Stepanova, Texas A&M University

Dr. Anna Stepanova is a researcher at the Sketch Recognition Lab at Texas A&M University. She holds a Ph.D., Master's and Bachelor's in geology. Anna's research interests are in geosciences, micropaleontology and education.

Christina Belanger, Texas A&M University
Dr. Tracy Anne Hammond, Texas A&M University

Dr. Tracy Anne Hammond is a Professor of Computer Science and Engineering at Texas A&M University and Director of the Sketch Recognition Lab. She earned her Ph.D. in Electrical Engineering and Computer Science from the MIT Computer Science and Artificial Intelligence Laboratory at Massachusetts Institute of Technology, along with additional degrees from Columbia University in computer science, mathematics, applied physics, and anthropology.

Her research focuses on artificial intelligence, sketch recognition, intelligent user interfaces, haptics, and engineering education, with applications in education, health, accessibility, and related domains. She contributes to research, teaching, and professional service in AI and engineering education, with an emphasis on responsible and inclusive design.

Katherine Sarah Cherry, Texas A&M University
Syeda Fizza Ali, Texas A&M University

Syeda Fizza Ali is a PhD candidate in Interdisciplinary Engineering at Texas A&M University. Her research explores how academic stress relates to computational thinking among undergraduate engineering students, with a focus on creating learning environments that support student well-being. She also works as a graduate research assistant, focusing on educational technology tools and improving student engagement.

Impact of the FossilSketch Application on Students' Use of Generative Artificial Intelligence for Conceptual Assessment

Abstract

Prior research suggests an increasing reliance of undergraduate students on Generative Artificial Intelligence (GenAI) tools. As part of a larger project to develop and evaluate the effectiveness of the FossilSketch, a web-based learning tool, this paper examines how GenAI use changed before and after using the app. For this study, we hypothesized that effective learning environments, such as the FossilSketch, can impact (i.e., lower) the use of GenAI. For this purpose, we introduced the FossilSketch, designed to teach micropaleontology concepts to diverse students in three courses. Although all three courses were undergraduate general science courses, they varied in teaching modalities (in-person or online), nature of lab sessions (no labs, or required labs), level of coverage of micropaleontology concepts before FossilSketch (none, partial through guest lectures, or frequently through lectures), and the way the FossilSketch was introduced (extra-credit assignment or required assignment). Considering these variations, each course offers a unique perspective on the use of FossilSketch.

In all three courses, students completed a pre- and post-conceptual assessment comprising 4 open-ended constructed-response questions. For the pre-assessment, the researchers identified commonalities between student responses and GenAI responses. The researchers also identified writing characteristics that helped classify students' responses as either GenAI-generated or student-generated. To explore the change from before and after the use of the FossilSketch, this paper addresses two research questions: 1) How does the use of GenAI vary across courses before and after the use of the FossilSketch application? 2) How does the use of GenAI vary between questions before and after the use of the FossilSketch application?

In this cross-sectional study, we collected data from 111 undergraduate students ($n1$=40, $n2$=36, $n3$=35) at a large R1 university. To prepare the data, we classified each student's response as either student-generated (0) or GenAI-generated (1) for both pre-assessment and post-assessment. The data were analyzed using descriptive statistics and chi-square tests of students' use of GenAI in pre- and post-assessments across three courses. Also, chi-square tests were used to examine students' use of GenAI across questions in the pre- and post-assessments. Results indicated that the number of students who used GenAI was significantly lower in the post-assessment than in the pre-assessment. Also, we found that the variation at the question level was similar across courses. In line with our hypothesis, this study's results indicate that authentic context and activities can influence students' decisions to exercise their own judgment and become self-reliant. The study explains the need for additional research to examine how integrating active learning across different types of courses can mitigate factors that could push students to become overly reliant on GenAI tools.

Introduction

After the public introduction of Generative Artificial Intelligence (GenAI) in 2022, undergraduate students began using these tools in their coursework [1], [2]. Over the years, it has been reported that more students are relying on GenAI tools to complete assignments [3], [4]. However, it is noteworthy that excessive use of these tools indicates students' overreliance on GenAI. This over-reliance could also produce a range of unintended and negative effects on students' learning outcomes [5], [6], [7]. For example, research has found that when students depend too heavily on AI tools for assistance in work or as automated problem solvers, they may experience diminished critical thinking and problem-solving skills [8]. Students with GenAI are less likely to engage with the underlying content and analytical processes that foster independent cognitive development [9]. This type of dependence can erode analytical abilities and encourage uncritical acceptance of information, potentially leading to poorer decision-making and reduced motivation to engage with challenging tasks independently [9].

Research indicates that students' reliance on GenAI tools is influenced by course factors, such as the types of assessments assigned [10], [11], [12] or the learning environment they are exposed to [13]. For example, environments emphasizing problem-based engagement can shape how students adopt and depend on AI, with implications for independence and cognitive engagement. However, this research is sparse, and more research is needed to examine students' reliance on GenAI across varying learning environments, especially in courses that use active learning approaches.

As part of a larger project, we are developing a learning tool, i.e., the FossilSketch, to teach micropaleontology concepts across various science courses. The motivation for this research also emerged when we conducted a pre-conceptual assessment before introducing our designed active learning tool in three science courses. We noticed that, for open-ended constructed-response questions in the assessment, students' answers were similar to those generated by GenAI tools. The observation presented a unique opportunity to examine how introducing the FossilSketch application will influence students' decisions to use GenAI tools for their solutions across three courses. Given that the FossilSketch provides ample opportunities to learn and practice micropaleontology concepts, we hypothesized that students' reliance on GenAI would decrease after the use of FossilSketch. For this purpose, this paper specifically addresses two research questions:
1) How does the use of GenAI vary across courses before and after the use of the FossilSketch application?
2) How does the use of GenAI vary between questions before and after the use of the FossilSketch application?

Literature Review

According to Bonwell and Eison [14], Active learning encompasses instructional approaches that teach students learning material through engaging activities. Prior literature extensively shows that implementing active learning in various types of courses, including those taught through traditional passive lectures, significantly improves students' performance and engagement with the material [15], [16], [17]. It engages students in problem-solving and inquiry-based exercises, which have been shown to help students' conceptual understanding in science and geoscience courses [15], [18]. Technology interventions further support active learning by enabling interactive and collaborative learning environments that extend beyond traditional classroom activities [19], [20]. In science and geoscience courses, these technology-supported active learning strategies allow students to work with authentic data and visualizations, reinforcing inquiry-based learning and conceptual development [18].

Besides the existence and use of such approaches, with the emergence of GenAI tools, students began using them without them being specifically integrated as active learning tools or interventions in the course [21]. The literature has provided several reasons for students' reliance on them for coursework. For example, many students find that GenAI tools' immediate feedback on their responses to course assessment questions is very useful [22] [23]. Because GenAI tools can instantly assess the correctness or quality of students' assessment work, they serve as an effective tutoring system. Also, students used GenAI tools for improving the quality of their work in course assessments [24] [25]. Among other reasons, students found them useful to brainstorm ideas for how to begin working on the assessments [22] [26]. It is noteworthy that less confident students may rely more on GenAI tools than more confident students [21].

Prior literature also reported the impact of students' overreliance on GenAI tools [24]. Examples of these impacts include reduced cognitive engagement and problem-solving skills [5] [6], lower critical thinking [7], [27], and lower knowledge retention [28]. Given these documented risks, it is critical to examine how instructional design and pedagogical approaches shape students' patterns of GenAI use [9]. Specifically, examining variations in GenAI use across courses before and after the implementation of the active learning tool can reveal how course design affects students' reliance on GenAI.

Although prior research on the use of the FossilSketch observed its effectiveness for students' learning experience in micropaleontology, in this study, we continued exploring the effectiveness of the FossilSketch as a learning tool by observing how implementing an active learning environment, such as the FossilSketch, impacted GenAI usage in different courses, each with unique factors and learning environments.

FossilSketch Design

The FossilSketch [29], [30] is a web-based tool that includes interactive activities to help students learn micropaleontological concepts. The app provides a flexible environment in which students can learn these concepts at their own pace [31]. The app provides three types of activities: educational videos, interactive games, and exercises [32], [33]. After completing each activity, students receive automated feedback, allowing them to assess their performance independently without the instructor's intervention.

Educational videos introduce key concepts and guide users through the steps of microfossil identification. Interactive Games include matching exercises in which students further practice sorting and distinguishing between different microfossil features. Exercises include more complex scenarios, such as assemblage analysis, where users identify and classify microfossils from sample images and interpret the results to reconstruct paleoenvironments. In the current version, the app comprises 13 modules covering various micropaleontological concepts that can be taught to students in geoscience, biology, and environmental science.

Research Design & Methods

This study uses a quantitative approach, specifically a pre-post cross-sectional design, to examine the impact of the FossilSketch on students' utilization of GenAI tools to complete activities.

Site and Participants

The FossilSketch was used at Texas A&M University, a large R1 southwestern public institute in three undergraduate general science courses: GEOL 207, ARCS 104, and GEOL 208. Of these three courses, 111 undergraduate students participated in this study.

GEOL 207 (Dinosaur World) introduces the students to topics related to dinosaurs, but did not cover marine microfossils. Some topics of the course include dinosaur paleobiology and paleoecology; terrestrial paleoclimate and paleoenvironments of the Mesozoic; dinosaur ancestors; appearance and radiation of dinosaurs; and the appearance and ancestry of birds. GEOL 207 was a fully online course that offered online lectures and laboratory assignments. Participants in GEOL 207 were offered to complete the FossilSketch's activities as extra credit independently. 40 students from GEOL 207 participated in this study.

ARSC 104 (Contemporary Issues in Science - Cosmos, Earth and Humanity) focuses on contemporary issues in the science of our universe or cosmos, Earth, and humanity. ARSC 104 was an in-person course that offered lectures but no laboratory assignments. While the course's focus was not micropaleontology, ARSC 104 still partially covered topics related to microfossils and climate. A micropaleontologist gave guest lectures on Earth history. It offered ARSC 104's students the opportunity to complete the FossilSketch's activities as extra credit. 36 students from this course participated in the study.

GEOL 208 (Life on a Dynamic Planet) introduces students to the critical events in the Earth's 4.6 billion-year history that shaped life. GEOL 208 was an in-person course that offered both lectures by a micropaleontologist and laboratory sessions led by a graduate student teaching assistant. The course frequently focused on micropaleontology. Participants in GEOL 208 were required to complete the FossilSketch's activities. 35 students in GEOL 208 consented to participate in the study.

Measures and Data Collection

To collect data on students' use of GenAI to answer micropaleontology-related questions, we developed a conceptual assessment that included four open-ended constructed-response questions. During the grading of these questions, it was observed that students' responses are not properly constructed and use words mostly chosen by GenAI tools. Table 1 provides the details on the open-ended questions.

Table I. Open-ended constructed-response questions

Questions
How can you use foraminifera to estimate oxygen levels?
How can you use ostracoda to estimate water salinity?
What are the steps to identify a foraminifer to genus level?
What are the steps to identify an ostracod to genus level?

For all three courses, we used the four open-ended constructed response questions from the conceptual assessment. The data were collected in a pre-post design, with data collected before and after introducing the FossilSketch.

To determine whether a student used GenAI to answer any items, micropaleontologist experts evaluated participants' responses. The analysis focused on writing patterns [34], coherence, and content quality to differentiate student-authored responses from those generated by GenAI tools.

Analysis

To address the first research question and examine how GenAI use varied across courses, we conducted two complementary analyses. First, we prepared the data by classifying students as GenAI users if they used GenAI for at least one question in either the pre- or post-assessment. Students were coded as 1 if they used GenAI and 0 if they did not. An additional category, labeled unsure, was used when expert reviewers could not determine whether a response was generated by a student or a GenAI tool. Responses were coded as missing when students did not provide an answer. A chi-square test of independence was then conducted to examine differences in GenAI use across courses.

Second, to examine variations in GenAI use before and after the implementation of the FossilSketch across courses, we independently coded pre- and post-intervention GenAI use as 0 (no use), 1 (use), 2 (unsure), or missing. Comparative chi-square analyses were conducted separately for each course to evaluate changes in GenAI use from pre to post.

To address the second research question and examine GenAI use across questions, we conducted question-level analyses by categorizing GenAI use for each question in both the pre- and post-assessments across all courses. Comparative chi-square analyses were then performed to assess differences in GenAI use across questions and Courses.

For all analyses, effect sizes were calculated using Cramér's V to assess the magnitude of observed associations. All statistical analyses were conducted using IBM SPSS Statistics (Version 31.0.1.0) [35].

Results

Variation across the courses

We first conducted a chi-square analysis across courses, classifying the use of GenAI as students using it for any question, regardless of pre- and post-variation.

Table II: Count of students with GenAI use by course (observed and expected)

GenAI Use	GEOL 207	ARSC 104	GEOL 208
Used GenAI	14 (10.99)	8 (6.71)	6 (10.28)
Not Used GenAi	17 (20.01)	11 (12.29)	23 (18.72
Missing*	9	17	6
Total*	40	36	35

* Not included in analysis; *Note.* Expected counts are shown in parentheses. Missing and Total were included for reporting only and excluded from the chi-square analysis

A chi-square test of independence was conducted to examine differences in GenAI use across three courses, excluding missing responses. The association between the courses and GenAI use was not statistically significant, $\chi^2(2, N = 79) = 5.24$, $p = .073$. The effect size was small to moderate (Cramér's $V = .26$), indicating some variation in GenAI use across courses. Post-hoc analyses using Pearson residuals did not reveal any statistically significant cell-level deviations, suggesting that a single course did not drive observed differences.

Second, we conducted chi-square tests of independence separately for each course to examine changes in GenAI use following the introduction of the FossilSketch, excluding missing responses. The results are presented in Tables 2 and 3.

Table III: Observed counts of pre- and post-GenAI use by Course

	GEOL 207		**ARSC 104**		**GEOL 208**	
GenAI Use	**Pre**	**Post**	**Pre**	**Post**	**Pre**	**Post**
Used GenAI	23	7	8	4	9	2
Not Used	0	17	2	10	9	22
Missing*	17	16	26	22	17	11
Total*	40	40	36	36	35	35

* Not included in analysis

Table IV: Results of Chi-square analysis pre- and post-use of the FossilSketch across courses

GenAI Use	**GEOL 207**	**ARSC 104**	**GEOL 208**
χ^2	26.13	5.49	8.91
P	<.001**	0.019*	.003**
Cramér's *V*	0.75	0.48	0.46

**p<0.05, **p<0.01*

Results indicated a significant reduction in GenAI use in GEOL 207 with χ^2 (1, N = 47) = 26.13, $p < .001$, with a large effect (Cramér's V = .75). ARSC 104 also showed a significant decrease in GenAI use from pre to post, with χ^2(1, N = 24) = 5.49, $p = .019$, with a moderate effect (V = .48). Similarly, GEOL 208 demonstrated a significant reduction in GenAI use with χ^2(1, N = 42) = 8.91, $p = .003$, with a moderate effect size (V = .46). Together, these findings suggest that the use of FossilSketch could be a factor in consistent decreases in GenAI use across courses, though the magnitude of change varied by course, which could be due to the individual variations within the courses.

Variation across the questions within courses

To answer research question 2 and explore variation at the question level at pre- and post-assessments within courses, we conducted a series of chi-square tests of independence comparing pre- and post-intervention AI use for each course across each question. The results are presented in Table 4.

Table V: Results of Chi-square analyses pre- and post-use of the FossilSketch for each question across courses

Course	**Question**	χ^2	**p**	**Cramér's V**
GEOL 207	Q1	15.0	<.001**	0.61
	Q2	18.0	<.001**	0.67
	Q3	17.4	<.001**	0.66

	Q4	17.4	<.001**	0.66
ARSC 104	Q1	4.44	.035*	0.44
	Q2	6.08	.014*	0.51
	Q3	2.57	.109	0.34
	Q4	3.53	.060	0.39
GEOL 208	Q1	12.6	<.001**	0.55
	Q2	16.4	<.001**	0.63
	Q3	5.67	.017*	0.39
	Q4	8.26	.004**	0.47

$p<0.05$, $^{**}p<0.01$

A series of chi-square tests examined changes in AI use from pre- to post-intervention for four questions across three courses. In GEOL 207, significant reductions in AI use were observed for all questions (Q1: $\chi^2(1, N=45)=15.0$, $p<.001$, V=0.61; Q2: $\chi^2=18.0$, $p<.001$, V=0.67; Q3: $\chi^2=17.4$, $p<.001$, V=0.66; Q4: $\chi^2=17.4$, $p<.001$, V=0.66), with large effect sizes indicating strong practical impact. ARSC 104 showed significant decreases in AI use for Q1 ($\chi^2=4.44$, p=.035, V=0.44) and Q2 ($\chi^2=6.08$, p=.014, V=0.51), with moderate effects, while reductions for Q3 (p=.109) and Q4 (p=.060) were not statistically significant but demonstrated small-to-moderate effect sizes suggesting meaningful trends. GEOL 208 demonstrated significant AI use reductions across all questions (Q1: $\chi^2=12.6$, $p<.001$, V=0.55; Q2: $\chi^2=16.4$, $p<.001$, V=0.63; Q3: $\chi^2=5.67$, p=.017, V=0.39; Q4: $\chi^2=8.26$, p=.004, V=0.47), with effect sizes ranging from small to moderate to large, indicating the intervention effectively decreased AI use. Overall, question-based variations differed across courses. Further magnitude of change varied by both the course and the question.

Discussion

This study investigated how undergraduate students' use of GenAI varied across three courses and four open-ended constructed response assessment questions before and after the implementation of an active learning tool, the FossilSketch. The results indicate a consistent and statistically significant reduction in GenAI use following exposure to the app across all three courses, supporting our hypothesis that effective learning environments can influence students' reliance on GenAI tools.

At the course level, although baseline differences in GenAI use were not statistically significant, all courses demonstrated significant pre–post reductions in GenAI use, with moderate to large effect sizes. These findings align with prior research indicating that instructional design and learning environments shape how students adopt and depend on GenAI tools [13], [25]. In particular, the observed reductions are consistent with the active learning literature, which shows that environments emphasizing engagement, inquiry, and practice promote deeper cognitive processing and reduce surface-level strategies [14], [18], [19], [36]. The variability in effect sizes across courses may be attributed to differences in course modality, prior exposure to micropaleontology concepts, and the integration of the FossilSketch, suggesting that contextual factors play an important role in shaping GenAI use [37], [38].

At the question level, GenAI use decreased for most questions across courses, although the magnitude of change varied by question and course. This variation reflects prior findings that students' GenAI use is task-dependent, with students more likely to rely on AI tools for complex, open-ended, or procedural tasks [10], [25], [26]. In this study, all the questions required students to construct responses and to demonstrate disciplinary identification skills. A reduction in the use of GenAI could be due to the use of FossilSketch, where students gained the practice and feedback needed for self-reliance. This interpretation is consistent with research indicating that students' confidence and perceived usefulness influence their reliance on GenAI tools [9], [22].

Importantly, these findings extend prior literature documenting the risks associated with students' overreliance on GenAI, including reduced cognitive engagement, critical thinking, and knowledge retention [7], [8], [27], [32]. Rather than framing GenAI use solely as an academic integrity concern, this study provides evidence that pedagogical interventions, specifically active learning supported by educational technology (e.g., the FossilSketch), can mitigate unproductive reliance on GenAI. By offering authentic practice opportunities and feedback, the FossilSketch appears to encourage students to engage more directly with content and exercise independent judgment, consistent with calls for instructional approaches that shape responsible and cognitively meaningful GenAI use [23].

Limitations

This study has several limitations that should be considered when interpreting the results. First, the sample sizes within each course were relatively small, which may have limited statistical power and the stability of effect size estimates. Although statistically significant results were observed, small sample sizes can reduce the reliability and generalizability of the findings. Second, the determination of GenAI use was based on expert judgment rather than student self-report or automated detection methods. While experts rely on writing characteristics and response patterns, this approach introduces potential subjectivity and misclassification, particularly when student writing and GenAI-generated text overlap. Third, the analyses were conducted using aggregated pre- and post-assessment data rather than paired student-level data. As a result, individual-level changes in GenAI use could not be directly examined, limiting causal interpretations regarding how specific students altered their behavior following the intervention. Fourth, the study did not include process data, such as interaction logs from the FossilSketch or direct measures of how students used GenAI during learning. The absence of such data limits insight into when, why, and how students choose to rely on GenAI tools. Finally, the study was conducted at a single R1 institution within geoscience-related courses, which may constrain the generalizability of the findings to other disciplines, institutions, or student populations.

Future Work

Future research should address these limitations by employing larger, multi-institutional samples to improve statistical power and generalizability. Collecting paired pre–post data would enable longitudinal analyses of individual students' GenAI use patterns and provide stronger evidence of causal relationships between instructional interventions and AI reliance. In addition, future studies should incorporate mixed-methods approaches, including student surveys, interviews, and classroom observations [39], [40] , to understand students' motivations, confidence levels, and decision-making processes related to GenAI use. Integrating process-level data, such as app usage logs and time-on-task measures, would further clarify how engagement with active learning tools mediates reliance on GenAI. Finally, future work should explore how GenAI can be intentionally and transparently integrated into active learning environments, rather than implicitly avoided, to distinguish productive from unproductive uses and support students' development of ethical and metacognitive AI practices.

Conclusion and Implications

This study provides novel empirical evidence that an active learning technology, i.e., the FossilSketch application, can significantly reduce undergraduate students' reliance on GenAI tools across courses and constructed response assessment questions. Unlike prior research that primarily documents the prevalence, perceptions, or risks of GenAI use [1], [8], [22], this study demonstrates how instructional design and active learning approaches can help shape students' decisions to rely on GenAI.

The novelty of this work lies in its focus on GenAI use as a pedagogical outcome influenced by learning environment design, rather than as a fixed or inevitable student behavior. By examining GenAI use at both the course and question levels, this study offers a nuanced understanding of how active learning and authentic practice can promote student independence and cognitive engagement. As GenAI tools become increasingly embedded in higher education, these findings underscore the importance of designing learning environments that encourage students to exercise their own judgment and engage deeply with disciplinary content.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Award# 2337105. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

[1] J. Freeman, "Student generative AI survey 2025." Higher Education Policy Institute, Feb. 2025.

[2] S. F. Ali, A. S. Rege, S. A. Ponniah, H. Ehsan, and S. Anwar, "BOARD# 215: Role of Generative AI in Enhancing Pre-College Engineering Design Thinking (Work in Progress)," presented at the 2025 ASEE Annual Conference & Exposition, 2025.

[3] S. Grassini, "Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings," *Educ. Sci.*, vol. 13, no. 7, p. 692, 2023.

[4] C. K. Y. Chan and W. Hu, "Students' voices on generative AI: Perceptions, benefits, and challenges in higher education," *Int. J. Educ. Technol. High. Educ.*, vol. 20, no. 1, p. 43, 2023.

[5] E. Kasneci *et al.*, "ChatGPT for good? On opportunities and challenges of large language models for education," *Learn. Individ. Differ.*, vol. 103, p. 102274, Apr. 2023, doi: 10.1016/j.lindif.2023.102274.

[6] S. Koltovskaia, "Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study," *Assess. Writ.*, vol. 44, p. 100450, Apr. 2020, doi: 10.1016/j.asw.2020.100450.

[7] B. Z. Larson, C. Moser, A. Caza, K. Muehlfeld, and L. A. Colombo, "Critical thinking in the age of generative AI," *Acad. Manag. Learn. Educ.*, vol. 23, no. 3, pp. 373–378, Sep. 2024, doi: 10.5465/amle.2024.0338.

[8] C. Zhai, S. Wibowo, and L. D. Li, "The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review," *Smart Learn. Environ.*, vol. 11, no. 1, p. 28, 2024.
[9] Z. Smutny and F. Sudzina, "What Affects Work Performance When Using AI Chatbots? Investigating Mediations and Factors Affecting Performance Expectancy and Intentions to Use ChatGPT," *Int. J. Human–Computer Interact.*, pp. 1–25, 2025.
[10] K. E. Singer-Freeman, K. Verbeke, and B. Barre, "Generative AI Use Among University Students Depends on Academic Level and Task," *High. Learn. Res. Commun.*, vol. 15, no. 2, p. 8, 2025.
[11] C. Hou, G. Zhu, V. Sudarshan, F. S. Lim, and Y. S. Ong, "Measuring undergraduate students' reliance on Generative AI during problem-solving: Scale development and validation," *Comput. Educ.*, p. 105329, 2025.
[12] U. Farooq and S. Anwar, "ChatGPT Performance on Standardized Testing Exam--A Proposed Strategy for Learners," *ArXiv Prepr. ArXiv230914519*, 2023.
[13] S. Otto, S. Ejsing-Duun, and E. Lindsay, "Disruptive tensions and emerging practices: an exploratory inquiry into student perspectives on generative Artificial Intelligence in a problem-based learning environment," *Educ. Inf. Technol.*, pp. 1–30, 2025.
[14] C. C. Bonwell and J. A. Eison, *Active learning: creating excitement in the classroom*. in ASHE-ERIC higher education report, no. 1, 1991. Washington, DC: School of Education and Human Development, George Washington University, 1991.
[15] S. Freeman *et al.*, "Active learning increases student performance in science, engineering, and mathematics," *Proc. Natl. Acad. Sci.*, vol. 111, no. 23, pp. 8410–8415, Jun. 2014, doi: 10.1073/pnas.1319030111.
[16] S. Anwar, "Role of different instructional strategies on engineering students' academic performance and motivational constructs," 2020.
[17] S. Anwar and M. Menekse, "Unique Contributions of Individual Reflections and Teamwork on Engineering Students' Academic Performance and Achievement Goals," *Int. J. Eng. Educ.*, vol. 36, no. 3, Art. no. 3, 2020.
[18] C. Huguet *et al.*, "Improving the motivation of students in a large introductory geoscience course through active learning," *J. Geosci. Educ.*, vol. 68, no. 1, pp. 20–32, 2020.
[19] S. Freeman *et al.*, "Active learning increases student performance in science, engineering, and mathematics," *Proc. Natl. Acad. Sci.*, vol. 111, no. 23, pp. 8410–8415, 2014.
[20] S. Anwar, "Impact of Educational Technology-Based Learning Environment on Students' Achievement Goals, Motivational Constructs, and Engagement," presented at the Proceedings of the 2019 ACM Conference on International Computing Education Research, 2019, pp. 321–322.
[21] H. Johnston, R. F. Wells, E. M. Shanks, T. Boey, and B. N. Parsons, "Student perspectives on the use of generative artificial intelligence technologies in higher education," *Int. J. Educ. Integr.*, vol. 20, no. 1, p. 2, 2024.
[22] C. K. Y. Chan and W. Zhou, "Deconstructing student perceptions of generative AI (GenAI) through an Expectancy Value Theory (EVT)-based instrument," May 02, 2023, *arXiv*: arXiv:2305.01186. doi: 10.48550/arXiv.2305.01186.
[23] Y. K. Dwivedi *et al.*, "Opinion paper: 'So what if ChatGPT wrote it?' Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy," *Int. J. Inf. Manag.*, vol. 71, p. 102642, Aug. 2023, doi: 10.1016/j.ijinfomgt.2023.102642.

[24] J. H. Gruenhagen, P. M. Sinclair, J.-A. Carroll, P. R. A. Baker, A. Wilson, and D. Demant, "The rapid rise of generative AI and its implications for academic integrity: Students' perceptions and use of chatbots for assistance with assessments," *Comput. Educ. Artif. Intell.*, vol. 7, p. 100273, Dec. 2024, doi: 10.1016/j.caeai.2024.100273.

[25] H. Johnston, R. F. Wells, E. M. Shanks, T. Boey, and B. N. Parsons, "Student perspectives on the use of generative artificial intelligence technologies in higher education," *Int. J. Educ. Integr.*, vol. 20, no. 1, p. 2, Feb. 2024, doi: 10.1007/s40979-024-00149-4.

[26] E. Kasneci *et al.*, "ChatGPT for good? On opportunities and challenges of large language models for education," *Learn. Individ. Differ.*, vol. 103, p. 102274, Apr. 2023, doi: 10.1016/j.lindif.2023.102274.

[27] D. S. Bahng *et al.*, "Students' Critical Thinking: Impact of FossilSketch, an Active Learning Educational Application," presented at the 2025 IEEE Frontiers in Education Conference (FIE), IEEE, 2025, pp. 1–7.

[28] A. Barcaui, "ChatGPT as a cognitive crutch: Evidence from a randomized controlled trial on knowledge retention," *Soc. Sci. Humanit. Open*, vol. 12, p. 102287, 2025, doi: 10.1016/j.ssaho.2025.102287.

[29] C. Belanger *et al.*, "Fossilsketch- An Interactive Software that Improves Student Ability to Learn Microfossil Identification Independently," presented at the Geological Society of America Abstracts, 2022, p. 380355.

[30] A. Stepanova, S. Anwar, C. Belanger, and T. Hammond, "Understanding the Impact of the FossilSketch Application using Students' Perceptions," presented at the 2025 IEEE Frontiers in Education Conference (FIE), IEEE, 2025, pp. 1–8.

[31] A. Stepanova *et al.*, "Using the interactive software FossilSketch to teach micropaleontology to undergraduate students," *J. Geosci. Educ.*, vol. 73, no. 2, pp. 133–153, 2025.

[32] A. Stepanova, "Improving undergraduate student critical thinking and ability to solve environmental problems with fossil records through FossilSketch application," *NSF Award Number 2337105 Dir. STEM Educ.*, vol. 23, no. 2337105, p. 37105, 2024.

[33] A. Stepanova, S. Anwar, C. Belanger, T. A. Hammond, and C. A. Stanley, "Board 416: Undergraduate student experiences with FossilSketch software to learn basics of micropaleontology," presented at the 2023 ASEE Annual Conference & Exposition, 2023.

[34] S. J. Bang and S. Anwar, "Utilizing Natural Language Processing for Assisting in Writing English Sentences," presented at the 2024 ASEE Annual Conference & Exposition, 2024.

[35] *IBM SPSS Statistics for Windows*. (2025).

[36] S. Anwar *et al.*, "Role of Engagement in Predicting Sixth-grade Students' Performance in an Integrated STEM Life Sciences Unit," in *ASEE Virtual Annual Conference*, Jun. 2020.

[37] S. Anwar, A. A. Butt, and M. Menekse, "Exploring relationships between academic engagement, application engagement, and academic performance in a first-year engineering course," presented at the 2022 IEEE Frontiers in Education Conference (FIE), IEEE, 2022, pp. 1–5.

[38] S. Anwar, A. A. Butt, and M. Menekse, "Utilizing an NLP-supported mobile reflection application to explore academic engagement, application engagement, and performance in engineering and physics courses," *Int. J. STEM Educ.*, vol. 12, no. 1, p. 41, 2025.

[39] S. Anwar and M. Menekse, "A systematic review of observation protocols used in postsecondary STEM classrooms," *Rev. Educ.*, vol. 9, no. 1, pp. 81–120, 2021.

[40] U. Farooq, O. Al-Ani, S. F. Ali, R. Viruru, and S. Anwar, "A Multi-Modal Investigation of Instructional Practices in an Undergraduate Engineering Classroom," presented at the 2025 IEEE Frontiers in Education Conference (FIE), IEEE, 2025, pp. 1–9.