2026 ASEE ANNUAL CONFERENCE June 21–24, 2026 | Charlotte Convention Center, NC
Host Sponsor UNC Charlotte
Paper ID #52352

Incoming Engineering Students' GenAI Literacy: Differences Across Experience Levels

Udit Kumar Das, University of Louisville
Dr. Angela Thompson, University of Louisville

Dr. Angela Thompson is an Associate Professor in the Department of Engineering Fundamentals at the University of Louisville. Dr. Thompson received her PhD in Mechanical Engineering from the University of Louisville. Her research interests are in biomechanics and engineering education, particularly related to first-year students.

Ms. Campbell R Bego P.E., University of Louisville

Campbell Rightmyer Bego, PhD, PE, studies learning and retention in undergraduate engineering programs in the Department of Engineering Fundamentals at the University of Louisville's Speed School of Engineering. She obtained a BS from Columbia University in Mechanical Engineering, a PE license in Mechanical Engineering from the state of New York, and an MS and PhD in Cognitive Science from the University of Louisville. Her current interests are generative AI and information literacy for engineering, individualized, first-year persistence interventions, and the effectiveness of evidence-based practices in the engineering classroom.

Incoming Engineering Students' GenAI Literacy: Differences Across Experience Levels

Abstract

This Complete Research Paper evaluates incoming first-year engineering students' Generative AI (GenAI) literacy by comparing objective knowledge and subjective self-perceptions across prior experience levels. The widespread availability of GenAI tools has led to rapid adoption among students. However, the complexity of these tools requires a level of literacy to understand their capabilities, limitations, and ethical considerations to avoid misuse or over-reliance. Using a quantitative survey design, we assessed 212 incoming engineering students using a mixed-format instrument: an objective knowledge test and a subjective self-efficacy scale. We compared literacy outcomes between students with low versus high self-reported prior GenAI experience. It was anticipated that prior experience would predict higher literacy across both measures. Results revealed a critical misalignment: objective and subjective scores were uncorrelated. While the high-experience group reported significantly higher confidence, they demonstrated significantly lower objective knowledge than the low-experience group. We conclude frequent usage may build confidence without building the necessary technical understanding. Therefore, engineering education must provide AI literacy instruction in introductory courses on the limitations and ethics of GenAI to ensure students possess the critical skills required for responsible use.

Introduction

Generative Artificial Intelligence (GenAI) has been widely adopted by students for academic work. Recent surveys suggest use of GenAI tools like ChatGPT, Gemini, and Claude for schoolwork has grown rapidly. In a three-semester survey of first-year college students, ChatGPT use rose from 15.6% to 80.9% from Spring 2023 to Spring 2024 [1]. Responses also showed a shift from trying ChatGPT for non-academic use toward using it for homework assignments [1]. More recent surveys in the United Kingdom in December 2024 and in the United States in July 2025 reported 92% and 85% of undergraduate students using AI for coursework, respectively [2], [3]. Students most frequently use GenAI tools to explain concepts, summarize papers, and brainstorm ideas. [2], [3]. However, given that GenAI tools are no longer novel and students are regularly using them for school-related tasks, there is a need to ensure students understand the capabilities, limitations, and ethical risks of GenAI tools to support responsible and ethical use.

GenAI tools generate content from probabilistic outcomes and use machine learning algorithms to predict the most likely next word or phrase based on their training. This technology has been demonstrated as a powerful tool for learning and efficiency. For example, in our prior research analyzing the impact of integrating GenAI use into a first-year engineering course, we found a clear benefit on students' programming outcomes when using GenAI in the classroom, indicating that students were able to verify and correct AI-generated code outputs [4]. However, users must

also be aware of the limitations of GenAI tools. These systems can provide weak reasoning in a confident tone and produce outputs that look correct but are incorrect or misleading. These features create risks for learning when students accept outputs without verification or treat GenAI as an authoritative source. Students often maintain high levels of trust in ChatGPT output even when it is unlikely to respond accurately [5]. Moreover, GenAI outputs can raise copyright and plagiarism concerns and can also increase bias and unfairness, which may negatively affect teaching and learning outcomes [6]. Additionally, if students use GenAI tools inappropriately or unethically, it can circumvent their learning [6], [7]. Prior work in education has highlighted this combination of opportunity and risk as a central challenge for teaching and learning [6].

Early research on GenAI technologies indicated that students were more optimistic about the tools' capabilities and less worried about the limitations than their professors [8]. Students' enthusiasm raises important questions about whether students have the right knowledge to use the tools responsibly and effectively. Students use GenAI technologies extensively, but they also express ethical concerns around information accuracy, over-reliance, and academic integrity [7]. These questions highlight the need for AI literacy instruction in colleges. As GenAI becomes more popular among students, they must learn to use these tools while maintaining intellectual rigor and academic integrity. Educators must address how students use these tools, and their understanding of GenAI's capabilities, limitations, and ethical considerations [9]. Additionally, educators need to understand what students actually know and can do when engaging with AI tools for classwork, and how to assess students' AI literacy.

Background

Understanding students' actual literacy levels is important for guiding students to use GenAI tools ethically and responsibly. GenAI literacy involves both competence and judgment: knowing how to use the tool and critically evaluate AI outputs [10]. Long and Magerko [11] define AI literacy as a set of competencies that help people critically evaluate AI technologies, communicate and collaborate effectively with AI, and use AI as a tool in everyday contexts. Ng et al. [12] proposed that AI literacy spans four key dimensions: (1) know and understand, (2) use and apply, (3) evaluate and create, and (4) ethical issues.

A major challenge is how GenAI literacy is assessed. GenAI literacy is often measured using self-reports alone [13], [14]. For example, Wang et al. [15] developed and validated a 12-item survey measuring subject perceptions of AI literacy around 4 constructs: awareness, use, evaluation, and ethics. Subjective measures can be biased and misaligned with actual competence [13], [14]. Kelly et al. [16] conducted a survey with university students and found that confidence in using GenAI increased with experience, yet a segment of students who had never used a GenAI tool still felt confident in their ability to use GenAI. The Dunning–Kruger effect explains this misalignment, where individuals with lower skill tend to overestimate their abilities in self-assessments [17]. This means students who feel very tech-savvy might assume they understand AI well, when in fact their objective knowledge is limited [18], [19].

Recent work has pushed toward objective performance based assessments of GenAI literacy [10], [13], [14], [20]. Objective AI literacy assessments often utilize knowledge-based multiple-choice questions to minimize bias and provide a standardized measurement of competence. Chiu et al. [20] developed and validated a 25-item multiple-choice AI literacy test for middle school students. Ding et al. [14] developed and validated an objective AI literacy assessment with 25 multiple-choice items for pre- and in-service teachers in the United States. Their results showed that even educators demonstrated a moderate level ($M = .65$, $SD = 0.11$) of AI literacy on the test. Similarly, Jin et al. [10] developed the Generative AI Literacy Assessment Test (GLAT), a 20-item multiple-choice instrument, to objectively measure college students' GenAI literacy. They reported that GLAT scores were significant predictors of students' actual performance on GenAI-supported tasks outperforming self-reported measures. Researchers have also developed rubric-based assessment methods to evaluate GenAI information literacy and applied this evaluation in a first-year engineering programming assignment [21].

Given the prevalence of AI use by college students and the recency of AI literacy assessments, there is a need to measure AI literacy in different student populations. It is important to assess what our students know first, entering college, to inform interventions and instruction in AI literacy. To the authors' knowledge, only two studies have examined both objective and subjective measures in the same population. Jin et al. [10] compared their objective GLAT scale with a subjective measure from a different instrument, and Zhang et al. [22], a recent preprint, specifically investigated the misalignment between self-reported and objective-based AI literacy measures. The two types of assessments are seldom used together and compared. This gap in the literature points to the need for more studies that evaluate self-perceptions and actual AI knowledge to better understand whether confidence reflects competence.

Research Questions

This study evaluates incoming first-year engineering students' GenAI literacy using both objective and subjective assessments, and examines how these measures differ across self-reported GenAI experience levels. We pose the following research questions:

- **RQ1:** How well do incoming first-year engineering students perform on objective and subjective GenAI literacy assessments?
- **RQ2:** How are students' objective GenAI literacy scores related to their subjective GenAI literacy perceptions?
- **RQ3:** Do students with higher versus lower self-reported GenAI experience differ in objective and subjective GenAI literacy scores?

Methods

This study was approved by the Institutional Review Board at the University of Louisville.

Participants

The participants were first-year engineering students (N = 263) enrolled in ENGR 110, an introductory engineering course, in Fall 2025 at the University of Louisville. The survey was administered online during the first week of class. After data collection, responses were screened for completeness and quality. Some participants' responses were excluded from analysis due to the following reasons:

- Incomplete survey submissions ($n = 30$)
- No variability response on Likert items: Surveys where a respondent gave the same rating for all subjective (Likert-scale) questions ($n = 7$)
- No variability in true/false answers: Participants marked "False" for every true/false knowledge item ($n = 3$)
- No variability after first true/false question: Cases where a respondent selected only one "False" (and "True" for all others) out of seven true/false items (n = 8), or conversely only one "True" (and "False" for all others) ($n = 3$)

The final sample was $N = 212$ students.

Study Procedures and Materials

Students were invited to participate in the survey in their first week of classes. We used a mixed-format survey instrument consisting of three parts: an objective knowledge test, a subjective self-assessment and a single item measuring prior experience with generative AI tools. Twenty objective questions were selected and adapted: 11 from the Generative AI Literacy Assessment Test [10] and 9 from Ding et al. [14]. Objective questions asked students about what GenAI is, what it does, how it works, and how it should or should not be used. Subjective questions ask students about their skills identifying and using GenAI tools. Not all questions from the originally published scales were included. Items deemed too technical for a general incoming first-year cohort were omitted. We selected items that were most appropriate and relevant for our undergraduate engineering student population. For example, we excluded highly technical items such as "*In the context of Generative AI, what is 'zero-shot learning*?" from the GLAT [10], and "*Machine learning is a kind of statistical inference(T/F)*" from Ding et al. [14] was not included in our survey because we did not think most students would understand this term. The objective section contained true/false and multiple-choice questions to measure students' generative AI literacy across three constructs: Know & Understand, Use & Evaluate, and Ethics. The subjective section had Likert-scale items (1 = "Strongly Disagree," 7 = "Strongly Agree") adapted from Wang et al. [15] AI Literacy Scale framework. These self-report items were grouped into four constructs: Awareness, Usage, Evaluation, and Ethics. .

The items from Ding et al. [14] were excluded from further analysis because they showed very low internal consistency (Cronbach's α = 0.18, with several items correlating negatively with the total scale).

The reliability for the three objective measure subscales were low (Know & Understand: Cronbach's α = .27, McDonald's ω = .33; Use & Evaluate: α = .29, ω = .35; Ethics: α = .39, ω = .44). Removing individual items did not meaningfully improve reliability within any subscale. Because of this weak internal consistency, we decided not to interpret the objective items at the subscale level. Instead, we treated the 11 items as a single objective knowledge test. The combined objective scale showed moderate internal consistency (α = .53, ω = .55), and no item substantially increased reliability if removed. All 11 objective AI literacy items were retained.

For the subjective self-assessment, we evaluated reliability for four subscales (Awareness, Usage, Evaluation, and Ethics). Reliability was moderate for Awareness (Cronbach's α = .57, McDonald's ω = .60) and Ethics (α = .49, ω = .57), and acceptable for Usage (α = .71, ω = .74) and Evaluation (α = .76, ω = .76). For Awareness, dropping the reverse-scaled item improved α = .66, and for Ethics, dropping the reverse-scaled item improved Cronbach's α = .66. Reverse-keyed items reduced subscale reliability and showed evidence of misunderstanding in response screening; therefore, these items were excluded. Usage also included a reverse-keyed item, but because the subscale reliability was already acceptable (α = .71), we retained the item. The overall scale showed good internal consistency (α = .77; ω = .79). 10 items were retained for further analysis.

The objective and subjective items retained in the final instrument are summarized in Table 1 and the full assessment is provided in the Appendix.

Data Analysis

The survey data was deidentified prior to analysis by removing direct personal identifiers. For each student, we computed the mean of the 11 objective questions. We normalized the Likert-scale subjective responses to a 0–1 range and the subjective mean was computed using the 10 questions retained. Subjective subscale means were also calculated for four categories: Awareness, Usage, Evaluation, and Ethics. Self-reported prior GenAI experience was coded as an ordinal variable, and for group comparisons, experience was dichotomized into low and high groups as shown in Table 2.

We analyzed descriptive statistics and distribution to summarize objective performance and subjective perceptions (RQ1). To test associations (RQ2), we computed Pearson correlations between Objective Mean and Subjective Mean (including each subscale). To answer RQ3, we tested differences in objective and subjective scores between low and high experience groups using two separate ANOVAs. To further examine subjective assessments by the 4 constructs, we conducted a repeated-measures ANOVA with subjective subscale (Awareness, Usage, Evaluation, Ethics) as the within-subjects factor and experience level (Low, High) as the between-subjects factor. Greenhouse-Geisser corrected results were reported when sphericity was violated. Bonferroni-adjusted pairwise comparisons and estimated marginal means were used for follow-up tests. Statistical significance was set at $\alpha = .05$.

Table 1. GenAI literacy scales included in the survey.

Section	Scale	Response	# Items	Example
Objective Knowledge Test	Know & Understand	Multiple-choice (scored 0 or 1, 0 = incorrect & 1 = correct)	4	"Which of the following best describes 'Generative AI'?"
	Use & Evaluate	Multiple-choice (scored 0 or 1, 0 = incorrect & 1 = correct)	4	"As a student using Generative AI to gather information for an assignment, how should you approach the information it provides?"
	Ethics	Multiple-choice (scored 0 or 1, 0 = incorrect & 1 = correct)	3	"When a Generative AI system is used for screening job applications, what issue might arise concerning the quality and fairness of hiring decisions?"
Subjective Self-Assessment	Awareness	1 = Strongly Disagree to 7 = Strongly Agree	2	"I can distinguish between smart devices and non-smart devices."
	Usage	1 = Strongly Disagree to 7 = Strongly Agree	3	"I can skillfully use Generative AI tools to help me with my daily work."
	Evaluation	1 = Strongly Disagree to 7 = Strongly Agree	3	"I can evaluate the capabilities and limitations of a Generative AI tool after using it for a while."
	Ethics	1 = Strongly Disagree to 7 = Strongly Agree	2	"I always comply with ethical principles when using Generative AI tools."

Results

RQ1: Students' objective GenAI literacy performance averaged $M = .58$, $SD = .19$. The median was .64. The distribution was unimodal (Figure 1). Based on the Shapiro–Wilk test, Objective Mean deviated from normality ($W = .97$, $p < .001$).

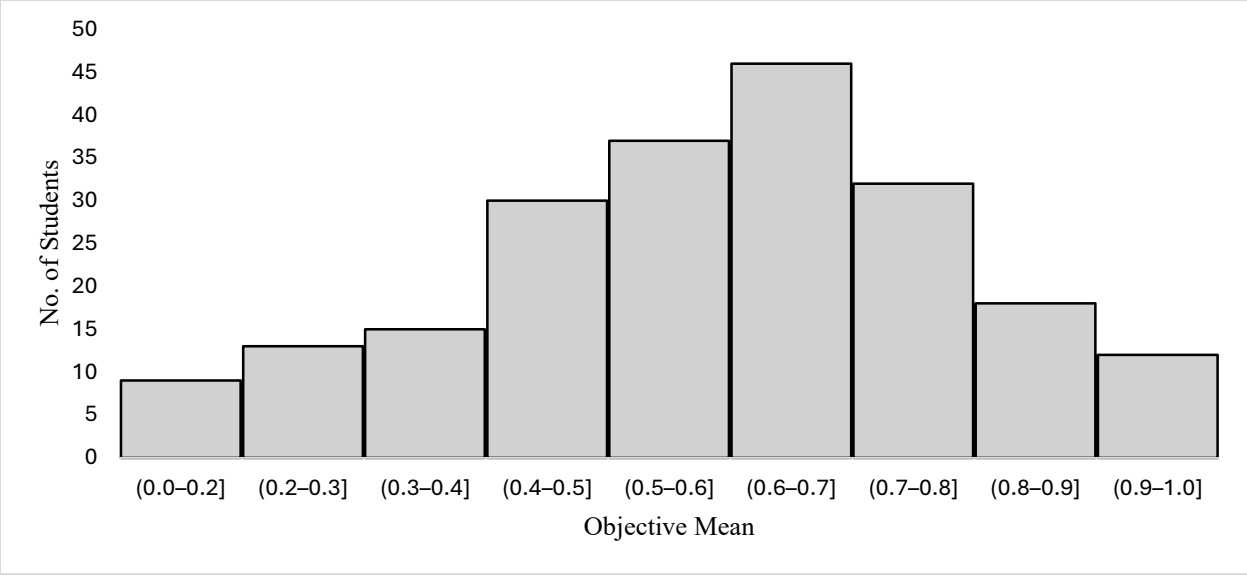

Figure 1. Objective AI literacy Mean Distribution

The subjective scale overall averaged M = .64, SD = .14. The median was .65, indicating that students perceived themselves as moderately to highly competent in GenAI literacy. The Shapiro–Wilk test did not indicate a significant deviation from normality for subjective mean (W = .99, p = .056), and the distribution was unimodal (Figure 2). Mean scores for the subjective subscales were Awareness (M = .69, SD = .19), Usage (M = .61, SD = .20), Evaluation (M = .60, SD = .20), and Ethics (M = .68, SD = .23).

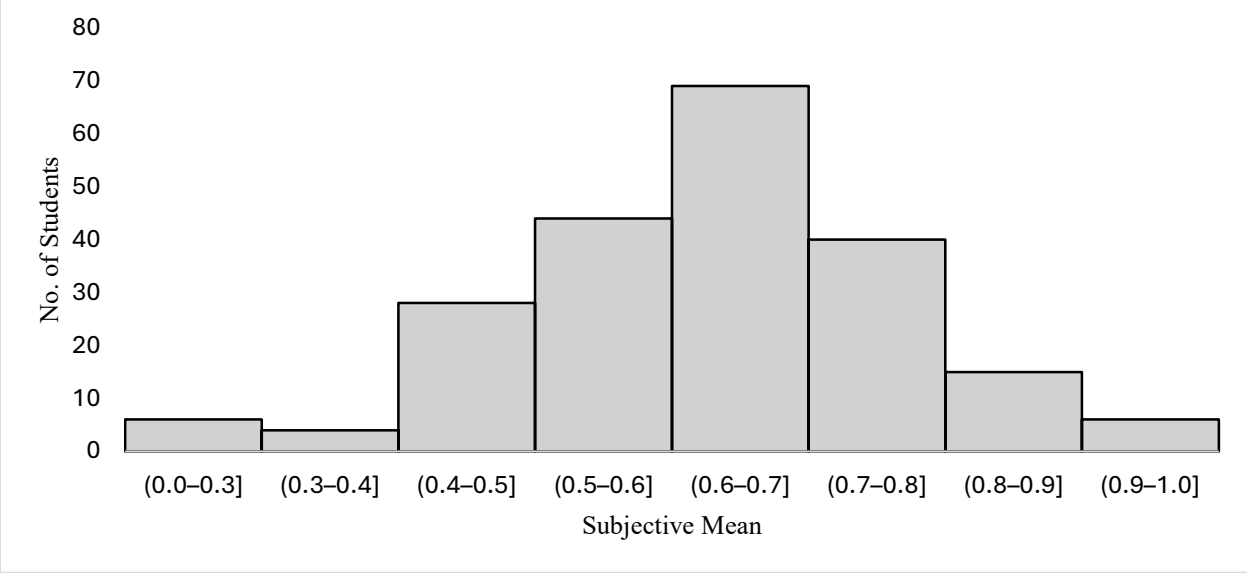

Figure 2. Subjective Mean Distribution

Prior GenAI experience (0–6 scale) had a mean of M = 2.44, SD = 1.28, indicating that most students reported low-to-moderate prior experience (Table 2).

Table 2. Experience Group Distribution

Experience Group	Experience Response	Value	*n* (%)	Group Total (*N*, %)
Low	None	0	13 (6.1%)	111 (52.4%)
	Circumstantial	1	37 (17.5%)	
	Minimal use	2	61 (28.8%)	
High	Moderate	3	59 (27.8%)	101 (47.6%)
	Regular	4	33 (15.6%)	
	Explorer	5	5 (2.4%)	
	Frequent	6	4 (1.9%)	
Total				212 (100%)

RQ2: Pearson correlations revealed that students' objective GenAI literacy was not significantly correlated with their subjective GenAI literacy ($r = .07$, $p = .288$). At the subscale level, objective GenAI literacy had a significant positive correlation with subjective Awareness ($r = .20$, $p = .004$), but objective GenAI literacy was not significantly related to the other subjective GenAI literacy subscales (Usage, $r = -.05$, $p = .479$; Evaluation, $r = .01$, $p = .865$; Ethics, $r = .11$, $p = .101$).

RQ3: Objective means significantly differed by experience level, $F(1, 210) = 4.49$, $p = .035$, $\eta^2p = .021$, $\omega^2 = .016$. The low experience group scored higher ($M = .60$) than the high experience group ($M = .55$). Levene's test was not significant ($p = .220$). Shapiro–Wilk indicated deviation from normality ($p = .036$), but a Kruskal-Wallis non-parametric test also revealed a statistically significant difference in objective means between low and high experience groups ($p = 0.038$). Subjective mean also significantly differed by experience Level, $F(1, 210) = 18.3$, $p < .001$, $\eta^2p = .080$, $\omega^2 = .075$. The high experience group reported higher subjective GenAI literacy than the low experience group (Low: $M = .60$, High: $M = .68$). Assumption checks were acceptable (Shapiro–Wilk $p = .128$; Levene's $p = .160$).

For the subjective assessment, there was a significant main effect of subscale, $F(2.56, 537.07) = 14.8$, $p < .001$, $\eta^2p = .066$, indicating that subjective ratings differed across Awareness, Usage, Evaluation, and Ethics (Figure 3). There was also a significant main effect of experience, $F(1, 210) = 11.4$, $p < .001$, $\eta^2p = .051$, indicating differences in subscale ratings between the Low and High groups. The Subscale × Experience Level interaction was significant, $F(2.56, 537.07) = 23.2$, $p < .001$, $\eta^2p = .100$, meaning subscales differed across experience groups.

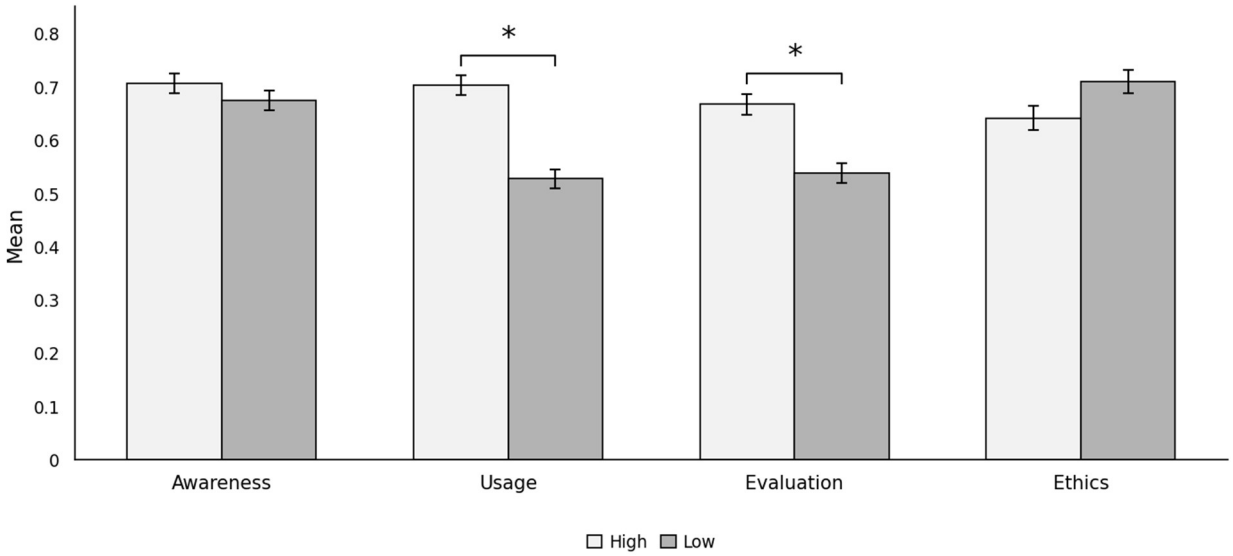

Figure 3. Estimated marginal means of subjective subscale scores (awareness, usage, evaluation, ethics) and experience level (high, low). Error bars indicate ±SE.

Post-hoc tests for the subjective assessment subscales showed that between groups, Usage mean was lower for the low experience group (M = .53, SE = .017) than the high experience group (M = .70, SE = .018; $p < .001$), and Evaluation mean was lower for the low experience group (M = .54, SE = .018) than the high experience group (M = .67, SE = .019; $p < .001$). Awareness means did not differ between groups (Low: M = .68, SE = .018; High: M = .70, SE = .019; p = 1.000), and Ethics means did not differ between groups (Low: M = .72, SE = .021; High: M = .64, SE = .022; p = .464)

Additionally, within the low experience group, Awareness (M = .68, SE = .018) was higher than Usage (M = .53, SE = .017; $p < .001$) and Evaluation (M = .54, SE = .018; $p < .001$). Ethics (M = .72, SE = .021) was also higher than Usage (M = .53, SE = .017; p < .001) and Evaluation (M = .54, SE = .018; $p < .001$). Usage and Evaluation did not differ (M = .53 vs .54; p = 1.000), and Awareness and Ethics did not differ (M = .68 vs .72; p = 1.000). Within the high experience group, none of the subscale differences were significant after correction ($p \geq .50$).

Discussion

This study evaluated incoming first-year engineering students' GenAI literacy and experience using an objective assessment, a subjective self-assessment, and a prior experience survey question, all adapted from recent literature [10], [14], [15]. The distribution of student performance on the objective GenAI literacy assessment was unimodal with a mean of 0.58 (on a 0-1 scale) with a standard deviation of 0.19, indicating that students enter engineering school with a wide range of GenAI literacy.

Subjective GenAI literacy results revealed that students generally perceived themselves as moderately competent. The average subjective score was M = 0.64, which is higher than the

objective average. Students rated their "Awareness" and "Ethics" particularly high. This suggests incoming engineering students (in Fall 2025) felt confident in their general understanding of AI concepts and ethical obligations, even if their technical understanding is less developed.

We hypothesized that prior experience using GenAI would relate to higher literacy. However, results revealed the opposite for the objective assessment. Students with low prior experience scored significantly higher on the objective test (M =. 60) compared to students with high prior experience (M = .55). Conversely, the high experience group reported significantly higher subjective confidence (M = .68) than the low experience group (M = .60). When looking at the breakdown by subscale, High experience students specifically rated their ability to "Use" and "Evaluate" GenAI tools much higher than low experience group, despite lower objective knowledge.

These results show a gap between self-perception and demonstrated knowledge. We found no significant correlation between objective and subjective GenAI literacy scores (r = 0.07). This misalignment may be explained by the Dunning-Kruger effect, where individuals with lower competence overestimate their abilities and experts underestimate theirs [17]. This finding also aligns with Kelly et al. [16], who found that sometimes student confidence/self-perception was greater than actual understanding. In our population, high confidence did not translate to high competence.

This misalignment with confidence and competence has important implications for engineering education. If students have increased perceptions of their understanding, they may make poor decisions about when and how to use these new technologies appropriately. However, if students have decreased perceptions of their understanding, they may avoid using a tool that could help their learning. This highlights the need for instruction and interventions in GenAI literacy, particularly in the first college year, to bring students' objective knowledge into alignment with their subjective perceptions.

Limitations

The objective knowledge test showed moderate reliability (α = .53). This was lower than GLAT's reported reliability (α = .80) for the full 20-item GLAT [10] and a subsequent study applying the full 20-item GLAT to medical and nursing students reported α = .76 [23]. Our lower reliability likely stems from three factors. First, we used a shortened version (11 items) rather than the full Jin et al. [10] and Ding et al. [14] scales, and reliability inherently decreases with fewer items. Second, we combined items from Jin et al. [10] and Ding et al. [14], creating a mixed instrument that may not function cohesively. Third, the original studies surveyed broader populations, whereas our sample consisted entirely of incoming first-year students.

The reliability of our overall subjective items was acceptable (α = .77). Wang et al. [15] reported reliability, α = .83, which is slightly higher than our data. Prior implementation of the Wang et al.

[15] subjective scale has shown higher reliability, α = .85 [24]. However, Uğraş et al. [24] studied 440 pre-service teachers, a much larger and different sample than our 212 engineering students.

This study was conducted at a single university with one cohort of first-year engineering students. The findings may not be generalizable to students at other institutions or in different disciplines. Additionally, measuring complex skills like "Evaluation" and "Ethics" through multiple-choice questions is inherently difficult. While objective tests reduce self-report bias, they may not fully capture a student's ability to apply GenAI in open-ended, real-world scenarios. Prior experience was self-reported and may not accurately represent the frequency or ways that students engaged with GenAI tools.

Future work may revert to the full GLAT assessment or explore additional objective knowledge questions to see whether construct reliability is improved in this population. Objective assessments of AI literacy are critical to evaluate the effectiveness of AI literacy interventions and instruction.

Conclusion

This study examined incoming engineering students' GenAI literacy on both objective and subjective assessments. Results revealed a misalignment in student performance on objective and subjective assessments. Students with higher self-reported experience using GenAI performed higher on the subjective self-assessments of their literacy but lower on the objective knowledge-based literacy assessments than students with less experience. In other words, students with more experience may have higher confidence using AI tools, but lower actual competence. This has important implications for educators; given the high prevalence of GenAI use by college students on schoolwork, and potential for misuse, there is a need for AI literacy instruction in introductory college courses to ensure students are using GenAI in ways that support their learning.

References

[1] M. Frydenberg, K. Mentzer, and A. Patterson, "The Rapid Rise of Generative AI Adoption among First-Year College Students," *Inf. Syst. Educ. J.*, vol. 24, no. 1, pp. 4–18, 2025, doi: 10.62273/HVNN2048.

[2] J. Freeman, "Student Generative AI Survey 2025." Policy Note 61. Higher Education Policy Institute (HEPI), February 26, 2025. https://www.hepi.ac.uk/reports/student-generative-ai-survey-2025/.

[3] C. Flaherty, "How AI Is Changing—Not 'Killing'—College," Inside Higher Ed. Accessed: Jan. 20, 2026. [Online]. Available: https://www.insidehighered.com/news/students/academics/2025/08/29/survey-college-students-views-ai

[4] U. K. Das et al., "The Impact of a GenAI Integration on First-Year Engineering Students' Programming Outcomes," in 2025 IEEE Frontiers in Education Conference (FIE), Nov. 2025, pp. 1–6. doi: 10.1109/FIE63693.2025.11328708.

[5] C. R. Bego *et al.*, "Working towards GenAI literacy: Assessing first-year engineering students' attitudes towards, trust in, and ethical opinions of ChatGPT," presented at the 2024 ASEE Annual Conference and Exposition, American Society for Engineering Education, Jun. 2024. doi: 10.18260/1-2--48555.

[6] 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.
[7] W. Yan, T. Nakajima, and R. Sawada, “Beyond tool use: Tracking the evolution of generative AI literacy among university students through a process-oriented investigation,” *Comput. Educ. Artif. Intell.*, vol. 9, p. 100465, Dec. 2025, doi: 10.1016/j.caeai.2025.100465.
[8] C. K. Y. Chan and K. K. W. Lee, “The AI generation gap: Are Gen Z students more interested in adopting generative AI such as ChatGPT in teaching and learning than their Gen X and millennial generation teachers?,” *Smart Learn. Environ.*, vol. 10, no. 1, p. 60, Nov. 2023, doi: 10.1186/s40561-023-00269-3.
[9] N. McDonald, A. Johri, A. Ali, and A. H. Collier, “Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines,” *Comput. Hum. Behav. Artif. Hum.*, vol. 3, p. 100121, Mar. 2025, doi: 10.1016/j.chbah.2025.100121.
[10] Y. Jin, R. Martinez-Maldonado, D. Gašević, and L. Yan, “GLAT: The generative AI literacy assessment test,” *Comput. Educ. Artif. Intell.*, vol. 9, p. 100436, Dec. 2025, doi: 10.1016/j.caeai.2025.100436.
[11] D. Long and B. Magerko, “What is AI Literacy? Competencies and Design Considerations,” in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, Honolulu HI USA: ACM, Apr. 2020, pp. 1–16. doi: 10.1145/3313831.3376727.
[12] D. T. K. Ng, J. K. L. Leung, S. K. W. Chu, and M. S. Qiao, “Conceptualizing AI literacy: An exploratory review,” *Comput. Educ. Artif. Intell.*, vol. 2, p. 100041, 2021, doi: 10.1016/j.caeai.2021.100041.
[13] A. Markus, A. Carolus, and C. Wienrich, “Objective measurement of AI literacy: Development and validation of the AI competency objective scale (AICOS),” *Comput. Educ. Artif. Intell.*, vol. 9, p. 100485, Dec. 2025, doi: 10.1016/j.caeai.2025.100485.
[14] L. Ding, S. Kim, and R. A. Allday, “Development of an AI literacy assessment for non-technical individuals: What do teachers know?,” *Contemp. Educ. Technol.*, vol. 16, no. 3, p. ep512, Jul. 2024, doi: 10.30935/cedtech/14619.
[15] B. Wang, P.-L. P. Rau, and T. Yuan, “Measuring user competence in using artificial intelligence: validity and reliability of artificial intelligence literacy scale,” *Behav. Inf. Technol.*, vol. 42, no. 9, pp. 1324–1337, Jul. 2023, doi: 10.1080/0144929X.2022.2072768.
[16] A. Kelly, M. Sullivan, and K. Strampel, “Generative artificial intelligence: University student awareness, experience, and confidence in use across disciplines,” *J. Univ. Teach. Learn. Pract.*, vol. 20, no. 6, Aug. 2023, doi: 10.53761/1.20.6.12.
[17] K. Mahmood and University of the Punjab, “Do People Overestimate Their Information Literacy Skills? A Systematic Review of Empirical Evidence on the Dunning-Kruger Effect,” *Comminfolit*, vol. 10, no. 2, p. 199, 2016, doi: 10.15760/comminfolit.2016.10.2.24.
[18] M. C. Laupichler, A. Aster, and T. Raupach, “Delphi study for the development and preliminary validation of an item set for the assessment of non-experts’ AI literacy,” *Comput. Educ. Artif. Intell.*, vol. 4, p. 100126, Jan. 2023, doi: 10.1016/j.caeai.2023.100126.
[19] T. Lintner, “A systematic review of AI literacy scales,” *Npj Sci. Learn.*, vol. 9, no. 1, p. 50, Aug. 2024, doi: 10.1038/s41539-024-00264-4.

[20] T. K. F. Chiu *et al.*, “Developing and validating measures for AI literacy tests: From self-reported to objective measures,” *Comput. Educ. Artif. Intell.*, vol. 7, p. 100282, Dec. 2024, doi: 10.1016/j.caeai.2024.100282.
[21] M. S. Yates *et al.*, “WIP - Developing a Rubric for GenAI Information Literacy in Programming,” in *2025 IEEE Frontiers in Education Conference (FIE)*, Nashville, TN, USA: IEEE, Nov. 2025, pp. 1–5. doi: 10.1109/FIE63693.2025.11328245.
[22] S. Zhang *et al.*, “How to Assess AI Literacy: Misalignment Between Self-Reported and Objective-Based Measures,” Jan. 03, 2026, *arXiv*: arXiv:2601.06101. doi: 10.48550/arXiv.2601.06101.
[23] Y. Jin, K. Yang, R. Martínez-Maldonado, D. Gavsevi’c, and L. Yan, “The Agency Gap: How Generative AI Literacy Shapes Independent Writing after AI Support,” Jul. 2025. Accessed: Jan. 20, 2026. [Online]. Available: https://www.semanticscholar.org/paper/The-Agency-Gap%3A-How-Generative-AI-Literacy-Shapes-Jin-Yang/63fdf1257360f007cd9f7873143bc0eac272fbbc
[24] H. Uğraş, M. Doğan, and M. Uğraş, “Adaptation of Artificial Intelligence Literacy Scale into Turkish: A Sample of Pre-Service Teachers,” *E-Kafkas J. Educ. Res.*, vol. 11, Art. no. 4, doi: 10.30900/kafkasegt.1429630.

Appendix

AI Literacy Survey

Instructions:

In this survey, the term "generative AI" refers to the technology that powers tools like ChatGPT, Claude.AI, Gemini, and many others.

1. How much have you used generative AI tools prior to this course? Select the description that best matches your experience:
    a. **None** – I've actively avoided generative AI.
    b. **Circumstantial** – I have seen generative AI results on google searches, etc, but have not logged onto a generative AI tool website.
    c. **Minimal** – I have played around with a generative AI tool a little bit, but not much.
    d. **Moderate** – I have used it intentionally more than a few times.
    e. **Regular** – I use a generative AI tool or tools regularly for school or personal uses.
    f. **Frequent** – I use generative AI tools every day.
    g. **Explorer** – I continue to explore new generative AI tools to see what they can do, and use them regularly whenever possible.

Objective Questions (#2-12) *adapted from*
Jin, Yueqiao, Roberto Martinez-Maldonado, Dragan Gašević, and Lixiang Yan. 2025a. "GLAT: The Generative AI Literacy Assessment Test." *Computers and Education: Artificial Intelligence* 9(December):100436. https://doi.org/10.1016/j.caeai.2025.100436.
Answers Bolded

2. Which of the following best describes "Generative AI"?
    a. AI that creates new content like text, images, or music by learning from existing data.
    b. An AI system designed to enhance the speed and accuracy of data retrieval in search engines.
    c. A form of artificial intelligence that focuses on translating languages in real-time.
    d. AI technology used primarily for managing and organizing large databases.
3. Which of the following tasks can Generative AI perform with a high degree of accuracy?
    a. Predicting stock market trends
    b. Making ethical decisions in complex scenarios
    c. Diagnosing rare diseases
    d. Generating human-like text based on prompts
4. Which of the following statements best describes how generative AI works?
    a. It generates text by analyzing and summarizing large volumes of web content.
    b. It generates text by predicting the next word based on the context of previous words.
    c. It generates text by translating input text into multiple languages simultaneously.

d. It generates text by using pre-defined templates and filling in the blanks.

5. It is unlikely for generative AI to provide an accurate summary of the latest financial market trends in real-time. Is this statement true or false?
    a. True, because the generative AI tool's **data may be outdated due to its knowledge cutoff.**
    b. True, because the generative AI tool is not good at handling numbers and structured data.
    c. False, because the generative AI tool frequently updates its knowledge base.
    d. False, because the generative AI tool is capable of synthesizing the latest market data automatically.
6. When using generative AI to help write a good marketing pitch, which of these is NOT very useful?
    a. Giving the generative AI tool details about the customers
    b. Asking the generative AI tool to mention what makes the product special an it's benefits
    c. Telling the generative AI tool to use convincing words
    d. Providing the generative AI tool with a list of competitors' products
7. What does the term "token" refer to in the context of generative AI?
    a. A token is a unit of text, such as a word or a sub-word, that the model processes individually.
    b. A token is a unique identifier assigned to each user interacting with the language model.
    c. A token is a security measure used to authenticate API requests to the language model.
    d. A token is a reward given to users for contributing valuable data to train the language model.
8. As a student using generative AI to gather information for an assignment, how should you approach the information it provides?
    a. The generative AI tool's answers are always more trustworthy than any information you will find on the internet, so you can use them without further verification.
    b. The generative AI tool's answers are generally more trustworthy than internet sources, but you should still verify the information with other reliable sources.
    c. The generative AI tool's **answers are not necessarily more trustworthy than internet sources, and you should cross-check the information with other credible references.**
    d. The generative AI tool's answers are less trustworthy than internet sources because it relies on outdated information.
9. A generative AI tool has provided a summary of a research paper. The summary states that increased screen time is directly correlated with decreased attention spans in children aged 8-12." What is your next step?

a. Accept the summary as accurate because generative AI tools are generally reliable.
b. Ask the generative AI tool to provide more details about the study's methodology and results.
c. Cross-check the summary with the original research paper.
d. Use another generative AI tool to generate a summary for comparison and evaluate the consistency between both summaries

10. When a generative AI system is used for screening job applications, what issue might arise concerning the quality and fairness of hiring decisions?
    a. The generative AI system might overlook applicants' unique achievements and extracurricular activities.
    b. The generative AI system could misinterpret minor formatting differences in resumes.
    c. The generative AI system might not effectively handle applications submitted in various languages.
    d. The generative AI system could reinforce existing biases found in historical hiring data.
11. What are the potential copyright implications for a journalist using an AI-generated image in a commercial article?
    a. The journalist needs to check the licensing policy of the AI tool they used.
    b. The AI-generated image is automatically free to use without any restrictions.
    c. The journalist must pay a standard licensing fee to use the AI-generated image.
    d. The image cannot be used in any commercial context because it is AI Generated.
12. In a healthcare startup, an accurate AI model recommends treatments, but doctors don't trust it because they can't understand how the model arrived at its conclusions. What core issue does this scenario illustrate?
    a. The generative AI model uses obsolete training data.
    b. The training dataset lacks sufficient diversity.
    c. The treatment guidelines input are incorrect.
    d. The generative AI model behaves as a black box.

Subjective Questions (#1-10 below) *adapted from*

Wang, Bingcheng, Pei-Luen Patrick Rau, and Tianyi Yuan. 2023. "Measuring User Competence in Using Artificial Intelligence: Validity and Reliability of Artificial Intelligence Literacy Scale." *Behaviour & Information Technology* 42 (9): 1324–37. https://doi.org/10.1080/0144929X.2022.2072768.

Instructions:

The following questions ask about *your* experience and feelings about generative AI tools.

All responses use a 7-point Likert scale from Strongly Agree to Strongly Disagree

1. I can distinguish between smart devices and non-smart devices.
2. I can identify the generative AI technology employed in the applications and products I use.
3. I can skillfully use generative AI tools to help me with my daily work.
4. It is usually hard for me to learn to use a new generative AI tool.
5. I can use generative AI tools to improve my work efficiency.
6. I can evaluate the capabilities and limitations of a generative AI tool after using it for a while.
7. I can choose a proper solution from various solutions provided by a generative AI tool.
8. I can choose the most appropriate generative AI tool from a variety of tools for a particular task.
9. I always comply with ethical principles when using generative AI tools.
10. I am always alert to the abuse of generative AI tools.