Journal of Data Science 0 (0), 1–17
??? 2026

DOI: 10.6339/26-JDS1233
Education in Data Science

Perceptions and Utilization of GenAI Tools Among Data Science Students and Faculty

Abeer M. Hasan[1,*] and Sayed A. Mostafa[1]

[1]*Department of Mathematics & Statistics, North Carolina A&T State University, Greensboro, NC, USA*

Abstract

This study investigates perceptions and use of generative artificial intelligence (GenAI) tools among students and faculty in statistics and data science at a historically Black college or university. Survey data from 119 valid student responses and 14 faculty responses were used to examine familiarity, usage patterns, perceived benefits, awareness of limitations, and instructional support needs. Students reported substantial use of GenAI, with ChatGPT as the dominant tool, primarily for coding assistance and writing support. Although student perceptions of AI in data science workflows and careers were generally positive, confidence in interpreting AI-generated outputs was limited, and concerns about accuracy, reliability, and over-reliance were common. Faculty also viewed GenAI favorably, but self-rated proficiency and the frequency of classroom integration remained limited. Comparisons across student subgroups suggested that familiarity with GenAI and awareness of its limitations varied more by academic level than by gender. These findings highlight a gap between AI adoption and AI literacy and underscore the need for structured training, validation practices, and clearer institutional guidance for responsible AI integration in data science education.

Keywords *AI literacy; data science (DS) education; generative AI (GenAI); limitations; perceptions*

1 Introduction & Background

Generative AI refers to a class of models that can create new data based on patterns and structures learned from existing data. These models rely on deep learning techniques and neural networks to generate content across domains such as text, images, and code. Since OpenAI released ChatGPT at the end of 2022, a competitive landscape of powerful tools, including Claude, Gemini, and Copilot, has emerged as a disruptive force across industries (Michel-Villarreal et al., 2023).

The field of data science (DS) is central to this transformation. GenAI tools are reshaping professional workflows by automating complex tasks in data wrangling, visualization, and analysis, thereby accelerating insight generation and fostering actionable intelligence. This technological shift has profound implications for the DS workforce and, consequently, for higher education. Early evidence indicates that AI is already altering the job market, depressing early-career employment in AI-exposed occupations (Brynjolfsson et al., 2025). This reality creates an urgent imperative for universities to adapt curricula to equip students with the skills to navigate and leverage this new AI-augmented ecosystem.

*Corresponding author. Email: amhasan1@ncat.edu.

Received February 18, 2026; Accepted May 28, 2026

A growing body of literature suggests that integrating GenAI can significantly enhance the learning experience in DS. AI tools can automate repetitive tasks such as writing computer code, allowing students to focus on conceptual understanding and critical thinking (Holmes et al., 2023). Studies suggest that GenAI-driven educational tools can improve academic performance (Jauhiainen and Guerra, 2023). Beyond simple automation, AI is being positioned as a pedagogical partner. For example, AI-generated adaptive feedback has been shown to improve students' statistical skills (Bauer et al., 2025).

Despite clear opportunities, the literature extensively documents formidable barriers to integrating GenAI. Perceptions among educators vary widely, ranging from enthusiasm for innovation to skepticism about the ethical implications and integrity of course assessments (Prilop et al., 2025). A primary concern is the blurring of the line around academic misconduct, as the ease of content generation complicates traditional definitions of plagiarism and necessitates rethinking assessment strategies (Duah and McGivern, 2024).

Students' reliance on AI to complete assignments complicates the evaluation of their true understanding and effort. Xie et al. (2023) discuss the potential consequences of AI-assisted cheating, including risks to educational quality, fairness, and the credibility of academic institutions. While (Wang et al., 2026) focus on the limitations of AI detection tools, the inadequacies of traditional teaching methods in this context, and the potential for the responsible integration of AI into educational practices. Significant ethical concerns further complicate the adoption. GenAI models can perpetuate and amplify biases present in their training data, raising questions about equity in educational content (Gupta and Nyamapfene, 2025). The security and privacy of student data also remain key institutional issues (Ko and Chan, 2025).

These multifaceted challenges underscore that the successful integration of GenAI tools into DS is not merely a technical task but a complex socio-technical problem. It requires a coordinated effort to address faculty and student needs, establish clear ethical frameworks, and redesign curricula to balance the benefits of AI with the development of core competencies.

Our study provides empirical data on the student and faculty perspectives in DS education. Understanding the perspectives of DS students and faculty is crucial for navigating this transition; however, research on their specific use and needs remains relatively limited (Chan and Hu, 2023). This paper aims to fill this gap by analyzing survey data from DS students and faculty regarding the integration of GenAI tools into their curricula.

Contribution and Study Relevance

Existing research has largely emphasized broad perceptions of AI in education or the technical capacities of AI systems to support analysis, leaving a gap in understanding how AI tools are perceived and integrated within data-intensive STEM disciplines. A recent American Statistical Association (ASA) member survey provided an initial snapshot of how statisticians in academia and industry are engaging with generative AI, highlighting emerging use cases alongside persistent concerns about reliability, bias, ethics, and governance (Glickman and Yan, 2025). However, the descriptive nature of this effort underscores the need for more systematic, data-driven studies that rigorously examine patterns of use, perceptions, awareness of limitations, and educational implications across stakeholder groups, especially DS students and educators.

This study offers a discipline-specific perspective by focusing on students and faculty at a Historically Black College or University (HBCU), where questions of fairness, trustworthiness, and responsible AI adoption are closely linked to broader concerns of equity and representation in STEM fields. This context provides an important perspective on responsible and equitable

AI integration in technical education.

In addition to the discipline-specific questions, we developed parallel survey instruments for students and faculty. By integrating input from both students and educators, we present a comprehensive understanding of the experience and needs of both groups. The findings aim to inform the responsible and equitable integration of GenAI in DS education.

Our study aims to address the following research questions:

RQ1 How familiar are students with GenAI tools for DS?

RQ2 What are the students' perceptions of the role of GenAI tools in DS education and workflow?

RQ3 To what extent are students aware of the limitations and challenges of GenAI technology?

RQ4 How do gender or academic class levels influence students' overall familiarity, perceptions of GenAI usage, and awareness of limitations in DS?

RQ5 What are the faculty perceptions, competence, and concerns regarding integrating GenAI tools into DS education and workflow?

RQ6 What are the institutional support and training needs of the faculty to effectively integrate GenAI literacy and tools into their DS courses?

The remainder of the article is organized as follows. Section 2 discusses the study methodology. Section 3 discusses findings for Questions 6–7, 10–14, 18–19, 21, and 24 of the student survey. Section 4 presents findings from Questions 7–20, and 22–23 of the faculty survey. Additional descriptive results, including detailed response distributions for individual survey items, are reported in Section S1 of the Supplementary Materials. Section 5 presents study limitations, conclusions, and recommendations. Sections S2 and S3 of the Supplemental Materials include the complete survey instrument for students and faculty, respectively.

2 Methodology

This study was designed to evaluate the perceptions, adoption patterns, and principal concerns regarding AI tools among students and faculty in statistics and DS. To obtain a comprehensive view, we developed and administered two versions of the survey: one for students currently enrolled in a statistics or DS course and one for faculty who teach these courses. Participants were recruited via email invitations and flyers posted in the departments responsible for these courses. Data collection occurred during the academic year 2024-2025.

The survey instruments were adapted and expanded from the work of Chan and Hu (2023) to specifically address data science curricula, workflows, and tools. We expanded the survey instrument to include 24 questions for students and 23 questions for faculty. Using a combination of multiple-choice and Likert scale questions, the surveys evaluated familiarity with AI tools, common use cases, perceived benefits and risks, awareness of institutional policies, and training needs for faculty and students. The responses were analyzed using descriptive statistics and thematic categorization.

We constructed three composite scores reflecting student familiarity with AI tools, perceptions of AI integration in data science, and awareness of AI limitations. Confirmatory factor analysis was used to evaluate the internal structure. Welch's t-test was used for 2-group comparisons (Males vs. females), and ANOVA with Tukey post hoc was used for 3-group comparisons (lower-division, upper-division, and graduate students).

The survey was administered through Qualtrics. Data analysis and visualizations were performed using R version 4.4.2 (R Core Team, 2024). The complete student and faculty survey questions are provided in Sections S2 and S3 of the supplementary materials. To ensure reproducibility, we provided complete anonymized data from both surveys and R code with question-by-question data processing, summarization, and visualizations in our project's GitHub repository (https://github.com/Nothgisrandom/AI4DS).

3 Results from the Student Survey

We defined a valid response as one in which the participant spent at least 120 seconds and answered at least 25% of the questions. Filtering out invalid responses yielded a sample of 119 students. The respondents could skip any question they did not feel comfortable answering. The number of valid (nonempty) responses varied across questions, and some questions allowed multiple selections. When applicable, we used the percentage of valid responses as the denominator to calculate relative frequencies and indicated the number of valid responses in each relevant figure.

The gender breakdown of the student respondents was: 71 female (61.2%), 40 male (34.5%), and 5 in other reported categories. The class distribution of the student participants was 28 (25.2%) lower-division students (freshmen or sophomores), 57 (51.4%) upper-division students (juniors or seniors), and 26 (23.4%) graduate students. Most of the declared majors of student participants were DS, computer science/IT, or statistics. Freshmen were underrepresented in our survey because of the prerequisite sequence for our DS courses.

3.1 Familiarity with GenAI Tools for DS

To explore students' familiarity with AI tools for DS (**RQ1**), we used five survey questions. The majority of students reported at least some engagement with AI tools in their coursework: occasional use was the most common response (36.2%), followed by frequent (20.7%) and very frequent use (15.5%). In contrast, 15.5% had never used AI tools, and 12.1% used them rarely. Students primarily engaged with AI for coding assistance (69.4%) and proofreading or writing support (52.3%), with more specialized applications such as mathematical proofs and literature reviews reported less frequently. These results indicate that students use AI tools predominantly for implementation support and refinement rather than for content generation.

Among specific tools, *ChatGPT* was the most widely used (80%), followed by *Grammarly* (54%). Familiarity with domain-specific platforms such as *Google Colab* and *AI-driven analytics tools* was notably lower, suggesting that students' current exposure is concentrated in general-purpose AI assistants rather than specialized data analytics tools.

We also asked students to rate how often they use AI tools for specific DS tasks, including data wrangling and visualization, predictive modeling, exploratory data analysis, and the presentation of findings. Figure 1 summarizes the responses. Across all five task domains, the proportion of frequent users remained around 20%, while a substantial share reported rare or no use. These patterns may reflect limited training on DS-specific tools, restrictions imposed by faculty, and potential underreporting due to students' reluctance to disclose AI use. Detailed tool-by-tool familiarity ratings and task-level usage distributions are provided in Section S1.1 of the Supplementary Materials.

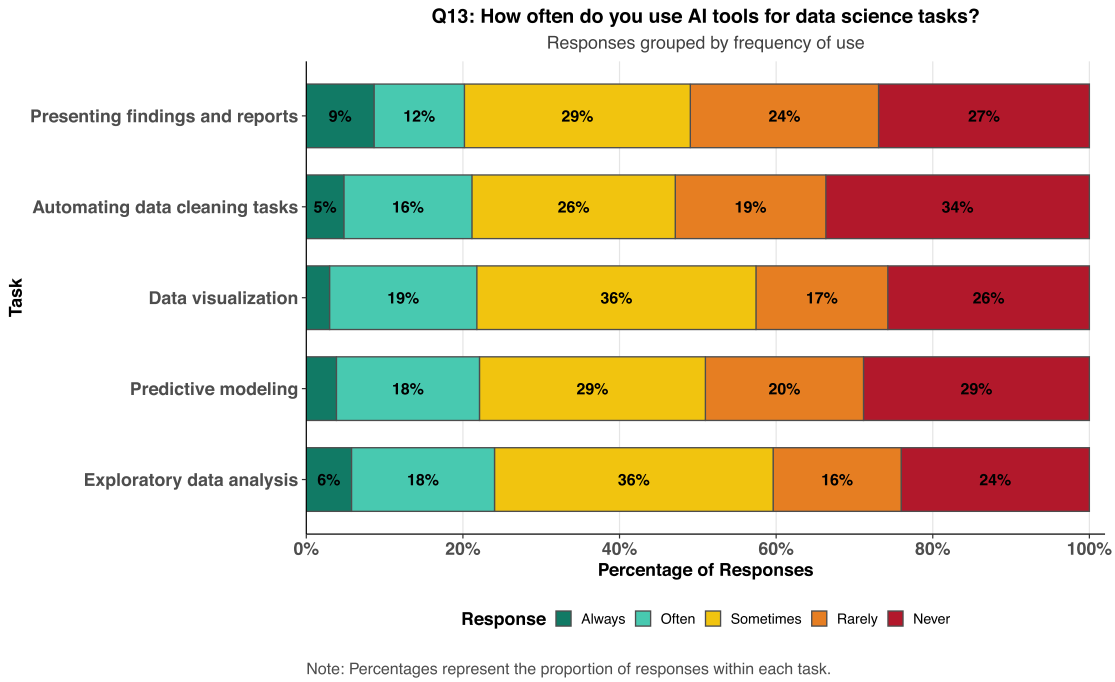

Figure 1: Students' self-reported usage frequency of AI tools in specific DS tasks.

3.2 Perceptions of GenAI Tools in DS

We used a set of Likert-scale questions to assess students' perceptions of AI tools in DS (**RQ2**). Survey responses indicate that most students viewed AI tools favorably; they seemed optimistic about AI's potential to automate processes, improve accuracy, and generate new insights. Nearly 70% of student respondents recognize the potential of AI tools to enhance accuracy, provide new insights, and automate data processing (Q12).

Figure 2 presents students' perceptions of AI's potential to enhance the DS workflow (Q21). Across all five workflow dimensions, agreement substantially outweighed neutrality and disagreement. The strongest endorsements appeared for supporting faster decision-making and automating repetitive tasks, while perceptions regarding predictive modeling accuracy showed relatively more neutrality. This pattern suggests that students associate AI primarily with efficiency gains while recognizing that modeling performance depends on statistical rigor and data quality.

Students also recognized the importance of AI for DS careers, when asked "Do you think students should learn how to use AI tools for their data science careers?", (Q16): 90.2% recommended learning AI tools for DS careers and considered AI an essential (32.4%) or a nice-to-have skill (57.8%), and 72.4% expressed willingness to incorporate AI tools into future DS projects (Q17). When asked whether AI tools should complement or replace traditional methods (Q24), a majority (54.9%) favored a balance between AI tools and traditional methods, while 31.4% favored using AI tools to complement traditional methods. Less than 8% supported replacement. These findings suggest a broad consensus favoring integrative rather than substitutive adoption of AI in DS. Complete response distributions for individual perception items are provided in Section S1.2 of the Supplementary Materials.

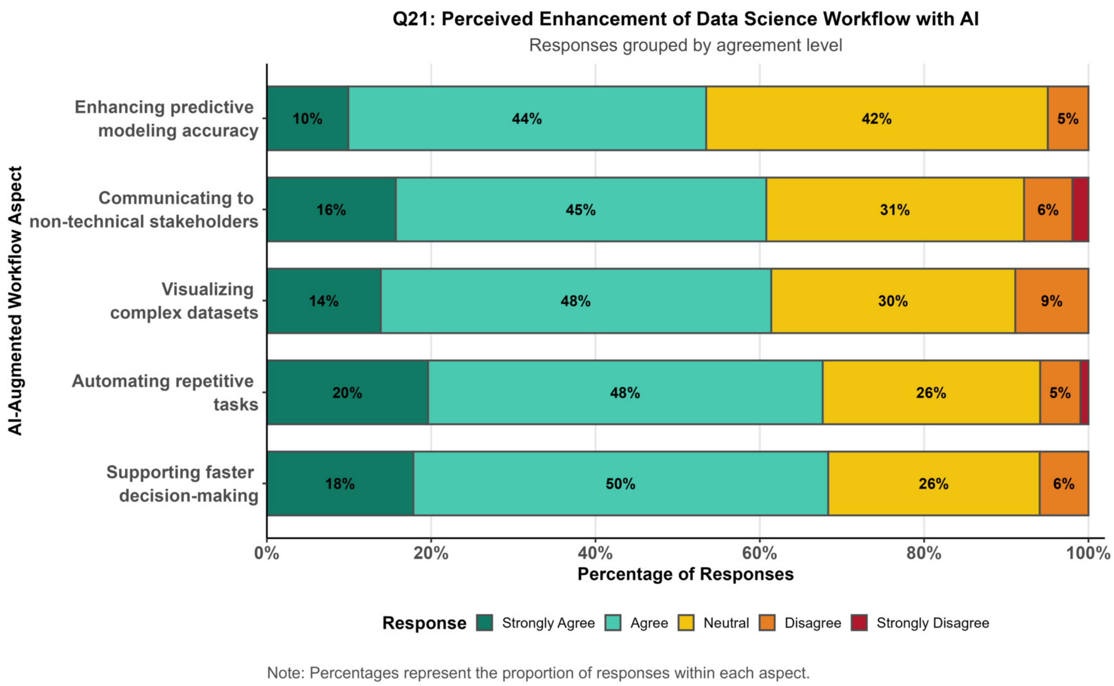

Figure 2: Students' perspective on AI's potential to enhance the DS workflow.

3.3 Awareness of GenAI Limitations

To explore students' awareness of AI limitations in DS tasks (**RQ3**), we used three sets of Likert-scale questions (Q18). More than 70% of students recognized the limitations of AI tools, particularly the risks of inaccurate or biased outputs and the need for expert validation. When asked whether they had encountered issues such as hallucinations, bias, or inaccuracy (Q20), nearly 70% reported experiencing such issues occasionally or frequently. Despite these encounters, students maintained an overall positive attitude toward AI's potential. Figure 3 summarizes students' awareness of specific AI limitations in DS workflows.

Students' confidence in explaining AI-generated outputs was notably limited (Q14): only 40.6% reported confidence, while 59.4% were either not confident or neutral. This gap between usage and interpretive confidence underscores the need for instruction in output validation and effective prompting strategies. Among students' primary concerns regarding AI in DS (Q19), accuracy and reliability were cited most frequently (23.6%), followed by privacy and security, skill degradation, and ethical concerns or bias. In the educational context (Q23), over-reliance on AI emerged as the dominant concern (52.4%), followed by hindering skill development (29.3%). These priorities indicate that epistemic risk-'can I trust the output?'-appears to outweigh occupational risk in students' current assessment of AI tools. Additional details on students' ranking of their concerns and challenges with AI integration into DS are provided in Section S1.3 of the Supplementary Materials.

3.4 Gender and Class Level Differences

In this section, we examine whether gender or academic class level influences students' familiarity with GenAI tools, perceptions of their role, and awareness of their limitations in DS (**RQ4**).

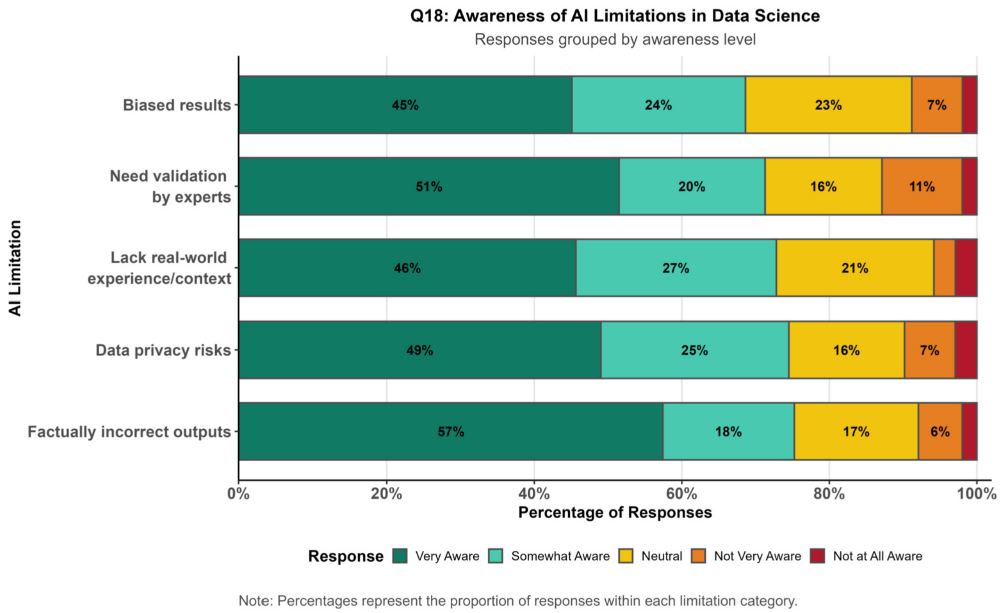

Figure 3: Students' awareness of the listed AI limitations in DS workflows.

To facilitate these comparisons, we constructed three composite scores from the survey items, validated each score using confirmatory factor analysis, and tested for group differences using Welch's t-test and one-way ANOVA.

3.4.1 Construction of Composite Scores

We constructed three composite scores by aggregating responses from related sets of Likert-scale survey items, as detailed in Table 1. For each item, responses were coded on a numeric scale from 1 to 5, with higher values indicating greater familiarity, more favorable perceptions, or greater awareness of limitations. For example, in Question 10, "Not at all Familiar" was coded as 1 and "Very Familiar" as 5; in Question 18, "Not at All Aware" was coded as 1 and "Very Aware" as 5. Each composite score was computed as the arithmetic mean of the coded item responses within its respective construct, yielding a score that ranges from 1 to 5. The three constructs are:

- **Familiarity with GenAI tools for DS**, derived from the five items in Question 10 (self-reported familiarity with specific AI tools);
- **Perceptions of GenAI use in DS**, derived from four items in Question 12 and five items in Question 21 (perceptions of AI's role in DS workflows);
- **Awareness of GenAI limitations**, derived from the five items in Question 18 (recognition of risks such as bias, hallucination, and privacy).

3.4.2 Confirmatory Factor Analysis

Before using these composite scores in group comparisons, we conducted confirmatory factor analysis (CFA) to evaluate whether the items within each construct adequately reflect a single underlying latent factor. CFA assesses model fit by comparing the covariance structure implied by the hypothesized factor model to the observed covariance structure in the data. We report the following fit indices, which are standard in applied survey research:

Table 1: Survey items used to construct composite measures.

Familiarity with GenAI tools for DS

Q10: How familiar are you with the following AI technologies in DS?

- ChatGPT
- Gemini
- Google Colab
- AI-powered data visualization tools (e.g., Tableau, Power BI with AI)
- AI-driven analytics platforms

Awareness of Limitations Score

Q18: To what extent are you aware of the following limitations of using AI for DS tasks?

- AI tools can produce biased results due to biased training data.
- AI tools may generate factually incorrect outputs (hallucinations).
- AI models lack real-world experience and context understanding.
- AI-generated results often need to be validated by experts.
- Using AI tools could pose data privacy risks.

Perceptions of AI Tools in DS (Q12 and Q21)

Q12: To what extent do you agree with the following statements regarding AI tools in DS?

- AI tools can enhance the accuracy of DS models.
- AI tools can automate tedious data processing tasks.
- AI tools can provide new insights or perspectives I would not have discovered otherwise.
- AI tools lack emotional intelligence, which may limit their utility in complex tasks.

Q21: Do you think AI tools can significantly enhance the following aspects of your DS workflow?

- Automating repetitive data processing tasks.
- Enhancing predictive modeling accuracy.
- Visualizing complex datasets more effectively.
- Supporting faster decision-making.
- Helping communicate findings to non-technical stakeholders.

- The **chi-square statistic** (χ^2) tests the null hypothesis that the model fits the data exactly. A non-significant p-value ($p > 0.05$) suggests an acceptable fit, although this test is sensitive to sample size.
- The **Comparative Fit Index (CFI)** and **Tucker-Lewis Index (TLI)** are incremental fit indices that compare the specified model to a baseline (independence) model. Values of 0.95 or above indicate a good fit (Hu and Bentler, 1999).
- The **Root Mean Square Error of Approximation (RMSEA)** quantifies the discrepancy between the model and the data per degree of freedom. Values below 0.06 indicate a close fit, while values between 0.06 and 0.08 reflect an acceptable fit. However, RMSEA is known to be upwardly biased in models with small degrees of freedom (Kenny et al., 2015), which is relevant for the five-item constructs examined here.
- **Cronbach's alpha** measures internal consistency, with values above 0.70 generally considered acceptable for survey research. We report both raw and standardized alpha values; close agreement between them indicates stable reliability across items.

For the *familiarity* construct (Q10), the model demonstrated adequate fit, $\chi^2(5) = 9.78$, $p = 0.082$, with a CFI of 0.96, TLI of 0.92, and RMSEA of 0.096. The slightly elevated RMSEA is consistent with the known upward bias of this index in models with few degrees of freedom (Kenny et al., 2015). The *perceptions* construct (Q12, Q21) also showed acceptable fit, $\chi^2(5) =$

Table 2: Model fit statistics from confirmatory factor analysis.

Composite	χ^2 Stat (p-value)	CFI	TLI	RMSEA	Raw Alpha	Standard Alpha
Familiarity	9.780 (0.082)	0.962	0.924	0.096	0.758	0.754
Perceptions	8.389 (0.136)	0.970	0.940	0.083	0.765	0.763
Awareness of Limitations	3.750 (0.586)	1.000	1.012	0.000	0.859	0.859

8.39, $p = 0.136$, with strong incremental indices (CFI = 0.97, TLI = 0.94) and RMSEA of 0.083. The *awareness of limitations* construct (Q18) exhibited excellent fit across all indices: $\chi^2(5) = 3.75$, $p = 0.586$, CFI = 1.00, TLI = 1.01, and RMSEA ≈ 0. Cronbach's alpha coefficients ranged from 0.76 to 0.86 across the three constructs, with standardized values closely aligned, indicating acceptable internal consistency. These results provide support for the internal consistency and unidimensional structure of the composite measures. Table 2 summarizes the CFA results for the Familiarity, Perceptions, and Limitations constructs.

3.4.3 Hypotheses and Statistical Tests

For each of the validated composite scores, we tested for statistically significant differences across gender and academic class level. For example, for the *Familiarity* score, we formulated the hypotheses in equations 1 and 2. Similar hypotheses were used for the *Perceptions* and *Limitations* scores.

$$\begin{cases} H_{01}: \text{There is no difference in } \textit{familiarity scores} \text{ between male and female students.} \\ H_{A1}: \text{There is a difference in } \textit{familiarity scores} \text{ between male and female students.} \end{cases} \tag{1}$$

$$\begin{cases} H_{02}: \text{There is no difference in } \textit{familiarity scores} \text{ among lower-division, upper-division,} \\ \qquad \text{and graduate students.} \\ H_{A2}: \text{At least one group differs in } \textit{familiarity score} \text{ among lower-division, upper-division,} \\ \qquad \text{and graduate students.} \end{cases} \tag{2}$$

All tests were evaluated at the overall 0.05 significance level. For two-group comparisons (1), we used Welch's two-sample t-test, which does not assume equal variances across groups and is robust for unequal sample sizes. Since we are comparing three composite scores for each subgroup, we adjusted our significance level to control the family-wise type I error rate using a Bonferroni correction, setting the adjusted significance level to $0.05/3 = 0.0167$. For the three-group comparison (2), we used one-way analysis of variance (ANOVA), followed by Tukey's post hoc pairwise comparisons where the omnibus test was significant.

3.4.4 Results

Table 3 presents the composite score means, standard deviations, and p-values for each group comparison.

Table 3: Composite score comparisons by gender and class levels.

Group	Familiarity			Perceptions			Awareness of limitations		
	n	Mean (SD)	P-value	n	Mean (SD)	P-value	n	Mean (SD)	P-value
Female	69	2.695 (0.921)		70	3.767 (0.597)		65	3.920 (0.939)	
Male	37	2.808 (0.963)	0.560	37	3.924 (0.592)	0.197	36	4.450 (0.572)	**0.001***
Lower-Division	27	2.265 (0.683)		27	3.811 (0.639)		24	3.658 (0.966)	
Upper-Division	54	2.755 (0.914)		54	3.822 (0.544)		52	4.196 (0.818)	
Graduate	26	3.162 (1.001)	**0.002***	26	3.831 (0.678)	0.993	25	4.360 (0.709)	**0.009***

Note. Pairwise comparisons were conducted using Welch's t-test, and three-group comparisons using ANOVA. To control the family-wise type I error rate using a Bonferroni correction, set the adjusted significance level to $0.05/3 = 0.0167$.

Gender Across the three composite scores, males had a higher arithmetic mean than females. Welch's t-tests revealed no statistically significant differences in *Familiarity* ($p = 0.560$) or *Perceptions* ($p = 0.197$) between female and male respondents. However, a statistically significant difference was observed for *Awareness of AI Limitations* ($p = 0.001$), with male respondents reporting higher mean awareness scores than female respondents.

Lower-Division, Upper-Division, and Graduate Across all three constructs, mean scores increased with academic standing. ANOVA tests indicated significant differences among the three class levels for *Familiarity* ($p = 0.002$) and *Awareness of AI Limitations* ($p = 0.009$), but not for *Perceptions* ($p = 0.993$).

Post hoc comparisons revealed that graduate students scored significantly higher than lower-division students on both *Familiarity* ($p_{adj} = 0.001$) and *Awareness* ($p_{adj} = 0.011$). The difference in *Awareness* between lower-division and upper-division students had ($p_{adj} = 0.027$). Note that the post-hoc pairwise comparisons were already adjusted for multiplicity using Tukey's HSD's method.

Overall, these comparisons suggest that differences in familiarity with GenAI tools and awareness of their limitations are more strongly related to academic experience and progression than with gender. The relative stability of perception scores across all subgroups indicates broadly consistent perceptions of AI in DS education, despite variation in exposure and self-reported knowledge.

4 Results from the Faculty Survey

We administered the faculty survey to faculty members who teach statistics or DS courses at our university. The number of valid responses was $n = 14$, with a gender breakdown of 9 females (64.3%) and 5 males (35.7%). We achieved a high response rate with nearly all statistics and data science faculty at our college completing the survey. The primary training of the faculty respondents was statistics or DS. They range in teaching experience from 1-3 years (38.5%) to over 10 years (23.1%). This section highlights the most interesting faculty findings; additional details are provided in Section S1.4-S1.6. The complete survey instrument, including item-level summaries and visualizations, is reported in Section S3 of the Supplementary Materials, with accompanying interpretive context.

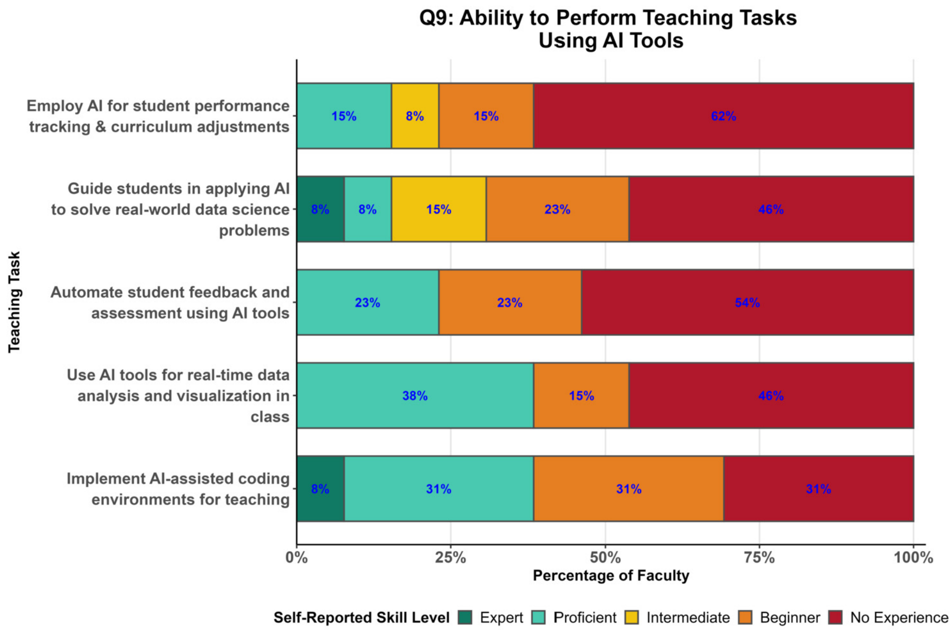

Figure 4: Faculty rating of their ability to perform teaching tasks using AI tools.

4.1 Faculty Integration of GenAI Tools in Teaching DS

To explore faculty familiarity with AI tools in teaching DS (**RQ5**), we asked whether faculty had incorporated AI tools into their teaching (Q7), and which tools they use (Q8). A majority (64.3%) reported incorporating AI tools at least sometimes, while 21.4% had never done so. All faculty respondents reported using AI learning assistants such as ChatGPT, followed by AI-powered coding platforms (64.3%) and data analysis or visualization tools (42.9%). Use of automated grading and project management tools was minimal, consistent with student responses indicating that ChatGPT was the most widely adopted GenAI tool.

When asked to rate their ability to perform teaching tasks using AI (Q9), over 60% of faculty rated their skill level as "Beginner" or "No Experience" (Figure 4). This is a notable contrast to student self-assessments and may reflect either faster student adoption or differences in self-assessment across groups. The finding underscores the importance of professional development and training support as faculty integrate new technologies into DS education. Across four specific teaching activities (Q10), fewer than 23% of faculty reported frequent AI use, and over 46% reported never or rarely using AI for each activity (Figure 5). The most common use cases were engaging students through AI-assisted interactive labs and generating teaching materials. Detailed distributions for faculty self-assessment and teaching activity usage are provided in Section S1.4 of the Supplementary Materials.

4.2 Faculty Perceptions of Using AI Tools in DS

We explored faculty perceptions of GenAI tools in teaching DS (**RQ5**). Overall, faculty expressed a generally positive view of AI adoption, coupled with nuanced concerns about its potential impact on student learning (Q11). Strong majorities agreed that AI tools can improve equity and inclusion in DS education (77%), automate repetitive teaching tasks (nearly 70%), and enhance instruction in technically complex topics such as machine learning and big data (over 75%). A

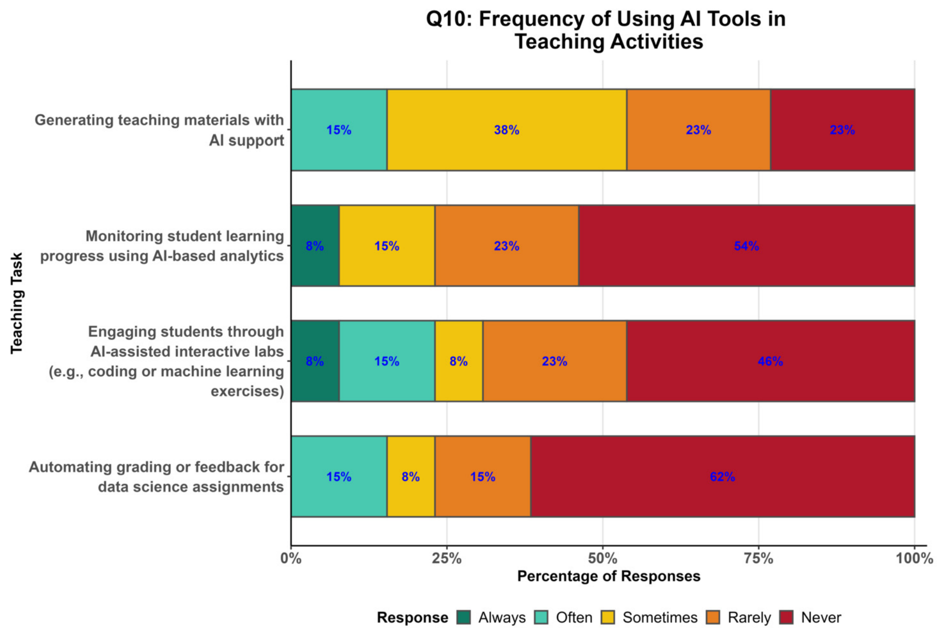

Figure 5: Faculty self-reported frequency of using AI tools in teaching activities.

majority (77%) also agreed that AI tools will be vital for the future of DS education. At the same time, faculty expressed notable concern that over-reliance on AI may reduce students' critical thinking skills, with responses split between agreement and neutrality rather than showing clear consensus. This tension, recognizing efficiency gains while guarding against erosion of deep learning, is a recurring theme in the results.

Figure 6 presents faculty perceptions of AI as an enhancer of DS teaching (Q15). The strongest endorsement was for simplifying complex concepts through AI-enhanced visualizations, and no faculty disagreed with AI's potential for real-time tracking of student challenges. Perceptions were most mixed regarding personalizing learning paths, with the highest proportion of neutral and negative responses. These patterns suggest that faculty favor pedagogically grounded integration that preserves critical thinking and student agency. Additional detail on the equity and inclusion finding and per-item response distributions is provided in Section S1.5 of the Supplementary Materials.

4.3 Faculty Concerns Regarding AI Integration in DS Education

We explored faculty concerns regarding AI tools in teaching DS (**RQ5**). Figure 7 summarizes the responses of Q13. Faculty were unanimous in their concern about keeping up with evolving AI technologies. Concerns about biased results in student assessments, impacting faculty-student engagement, and a lack of transparency in AI outputs were also widespread (at least 75% each). Data privacy and security, while still a majority concern, were rated somewhat lower, with 8% expressing no concern, which may reflect DS faculty's generally higher comfort with technology.

Regarding ethical priorities (Q18), 84.6% of faculty cited excessive reliance on GenAI tools leading to reduced human oversight as their primary concern, and all respondents agreed that maintaining human supervision is essential. The most frequently cited integration challenges were insufficient institutional support or training, balancing AI with curriculum learning objectives,

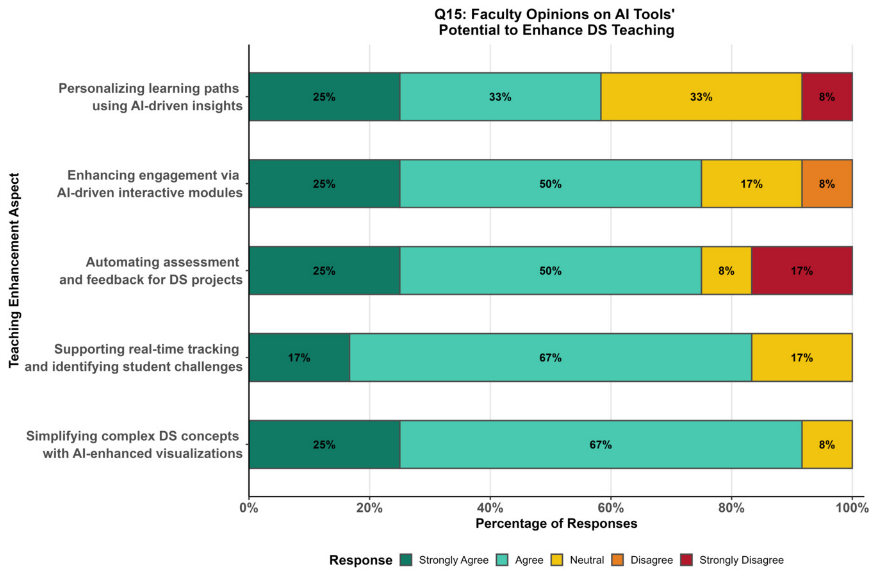

Figure 6: Faculty agreement on the potential for GenAI tools to improve DS education.

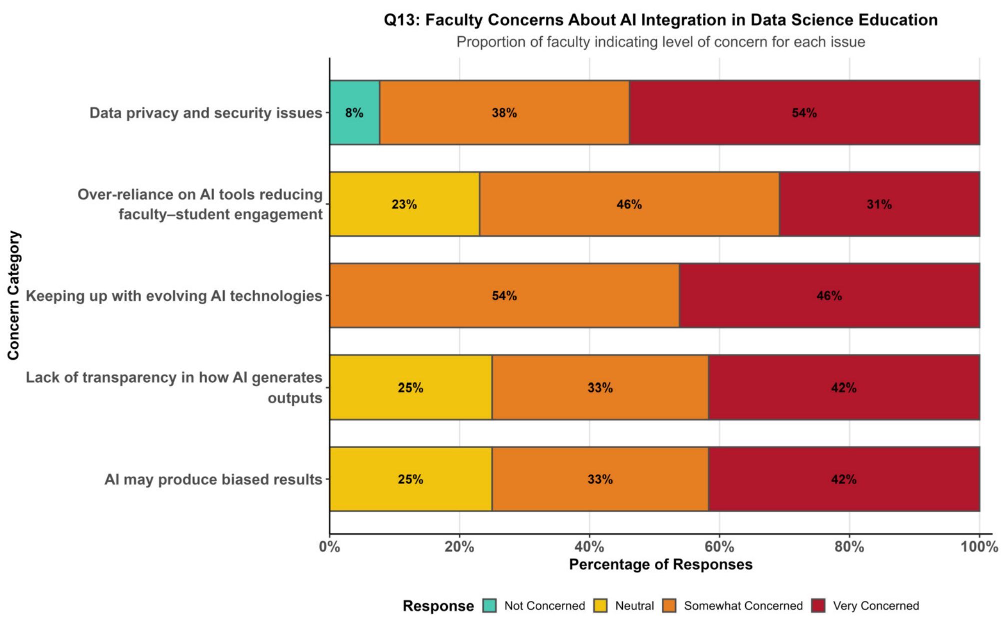

Figure 7: Faculty concerns regarding integrating GenAI tools into DS education.

and ethical concerns such as fairness and bias.

Syllabus policies varied across respondents (Q20): 83% allowed AI use for instructor-designated activities with required disclosure, while approximately 17% prohibited AI entirely; none permitted unrestricted use. More than a third of faculty were unaware of institutional AI guidelines. This absence of uniform policy within the same department risks creating confusion for students regarding the boundaries between acceptable assistance and academic dishonesty.

To address these challenges (**RQ6**), faculty identified training and support needs centered on access to advanced AI platforms, technical integration support, and professional development opportunities (Q22). A majority (53.8%) reported that their institution does not provide sufficient AI guidance or resources (Q23), highlighting a significant opportunity for institutional leadership. Complete response distributions for concerns, institutional awareness, and training needs appear in Section S1.6 of the Supplementary Materials.

5 Discussion

This study provides a descriptive quantitative account of generative AI adoption within a data science educational environment, explicitly comparing student and faculty populations. The results indicate that AI integration in DS education is neither uniform nor unidimensional. Rather, it reflects structured heterogeneity across usage intensity, interpretive confidence, and perceived risk.

From a measurement standpoint, the survey instrument operationalizes three latent dimensions: (i) familiarity with AI tools, (ii) awareness of limitations, and (iii) perceptions of GenAI tools in teaching statistics and DS content courses. Confirmatory factor analysis supports the multidimensional structure of these constructs. The moderate correlations among factors indicate related but distinct dimensions of AI engagement. Importantly, group comparisons demonstrate that AI engagement appears multidimensional and reflects interacting cognitive and attitudinal components.

Students report a significantly higher frequency of GenAI use than faculty (Figure 1 vs. Figure 5). This adoption asymmetry suggests differential exposure and incentive structures across stakeholder groups. Students appear to integrate AI tools into operational workflows, whereas faculty engagement is more cautious and pedagogically constrained. The divergence is not merely behavioral; it is accompanied by distinct confidence profiles. Students report lower confidence in interpreting AI-generated outputs, while faculty report limited confidence in using AI to support teaching tasks. The coexistence of high utilization and low interpretive confidence among students raises concerns about the acceptance of unverified outputs.

A notable empirical finding is that both faculty and students rank concerns about overreliance and skill erosion above privacy-related concerns. This prioritization suggests that perceived epistemic risk dominates technical risk in this context. In DS education, where statistical reasoning, inferential judgment, and model interpretation are core competencies, this distinction is particularly consequential. If AI systems function as probabilistic generators rather than deterministic solvers, then DS curricula must emphasize output verification, uncertainty quantification, and model interpretability.

The observed discrepancies between utilization and interpretive confidence indicate that AI adoption is outpacing the development of AI literacy. This gap underscores the importance of embedding formal instruction on the limitations of generative models, the risks of hallucination, and prompt design principles within DS coursework. Absent such instruction, students may rely on outputs without sufficient scrutiny, potentially weakening inferential rigor.

From a curricular modeling perspective, AI integration should be conceptualized as augmentation rather than substitution. Structured AI reflection assignments that require prompt disclosure, output validation, and justification of acceptance decisions may help preserve independent reasoning. Future research should empirically evaluate whether such interventions mitigate skill erosion through experimental or quasi-experimental designs.

More broadly, the results suggest that AI adoption in DS education is best approached as a socio-technical process involving competence acquisition, trust calibration, and institutional infrastructure. Access barriers, subscription costs, and limited professional development may contribute to uneven adoption patterns.

5.1 Implications for Data Science Education

The findings indicate that DS curricula must advance along two interconnected paths. Students need operational competence in using generative AI tools, including effective prompt construction, iterative refinement, and an understanding of system constraints. At the same time, programs must strengthen core statistical reasoning, particularly skills in model validation, uncertainty quantification, and the principled interpretation of results. Technical proficiency without conceptual depth risks superficial analysis.

The widespread availability of generative AI introduces a new pedagogical challenge. Students may mistake polished outputs for analytically sound conclusions. Traditional statistical software executes well-defined procedures, whereas generative systems produce responses through probabilistic processes that may not explicitly communicate assumptions or uncertainty. This distinction requires explicit instructional attention. Students should be trained to question outputs by examining assumptions, considering alternative modeling strategies, and identifying potential sources of error or instability.

Metacognition becomes central in this environment. Effective use of AI tools depends on a learner's ability to monitor their own reasoning, recognize when cognitive shortcuts are being taken, and determine whether AI assistance is enhancing or replacing understanding. Structured reflective practices can support this development. For example, students may be required to document how prompts evolved, justify acceptance or rejection of generated results, or independently verify outputs using conventional statistical methods. These activities encourage deliberate engagement rather than passive reliance.

Interpretability should also be emphasized. DS education has increasingly prioritized transparent modeling and explainable methods. The same standards should apply to generative systems. Students must understand that AI-generated outputs do not inherently guarantee correctness, reproducibility, or traceability. Assignments that require cross-validation of generated analyses, replication of results, or critical comparison between AI-assisted and independently derived solutions can preserve methodological rigor.

Historically, statistical software reduced the need for manual computation and shifted emphasis toward modeling decisions and interpretation. Generative AI represents a further shift by automating portions of analytical reasoning. The appropriate educational response is neither prohibition nor unconditional acceptance. Instead, curricula should cultivate supervisory competence. Students must learn to interrogate and evaluate AI outputs while retaining responsibility for inference and judgment.

Intentional integration of metacognitive training and interpretability principles into data science coursework is therefore essential. Without structured guidance, frequent use of GenAI tools, combined with limited interpretive confidence, may weaken foundational reasoning skills. With thoughtful instructional design, however, generative AI can promote more substantive engagement with underlying model assumptions, uncertainty, and responsibility for analytic decisions.

5.2 Limitations and Future Research

This study is limited to a single R2 HBCU in the southeastern United States. While the institutional context provides meaningful insight, external validity is limited. Multi-institutional replication across R1 universities, liberal arts colleges, and international contexts would strengthen generalizability. The reliance on self-reported survey data introduces potential response bias. Individuals with strong views regarding AI may have been more likely to participate. Additionally, concerns about academic integrity policies may have led some students to underreport their use of AI.

The analysis presented here is primarily descriptive with group comparisons. Faculty findings are especially descriptive due to the small respondent pool. Future extensions could use inferential modeling, such as regression to predict usage frequency and longitudinal designs to track changes over time. Finally, experimental studies evaluating the causal impact of structured AI literacy interventions on statistical reasoning outcomes represent an important next step. Such work would move beyond measurement of perception toward assessment of educational effectiveness.

Supplementary Material

The supplementary material contains (S1) additional findings from both the student and faculty surveys; (S2) the complete student survey instruments, including response options and item-level summaries/visualizations; and (S3) the complete faculty survey questions with item-level response summaries and visualizations.

Acknowledgments

The authors thank graduate students Stephen Agyeah, Oka Konan, Abiodun Olalere, and Chinedu Nzekwe for their support with this project. We thank the editor and referees for their constructive feedback, which improved this paper.

Funding

This study was funded by National Science Foundation Grant EES 2510214.

References

Bauer E, Richters C, Pickal AJ, Klippert M, Sailer M, Stadler M (2025). Effects of AI-generated adaptive feedback on statistical skills and interest in statistics: A field experiment in higher education. *British Journal of Educational Technology*, 56: 1735–1757.

Brynjolfsson E, Chandar B, Chen R (2025). Canaries in the coal mine? six facts about the recent employment effects of artificial intelligence. *Stanford Digital Economy Lab*.

Chan CKY, Hu W (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20(1): 43.

Duah JE, McGivern P (2024). How generative artificial intelligence has blurred notions of authorial identity and academic norms in higher education, necessitating clear university usage policies. *The International Journal of Information and Learning Technology*, 41(2): 180–193.

Glickman M, Yan J (2025). ASA members' perspectives on the use of generative AI. *Chance*, 38(2): 6–9.

Gupta V, Nyamapfene A (2025). Generative AI in universities: Practices at UCL and other institutions, and the path forward. *Internet Reference Services Quarterly*, 29(1): 131–151.

Holmes W, Miao F, et al. (2023). *Guidance for Generative AI in Education and Research.* Unesco Publishing.

Hu Lt, Bentler PM (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling. A Multidisciplinary Journal*, 6(1): 1–55.

Jauhiainen JS, Guerra AG (2023). Generative AI and ChatGPT in school children's education: Evidence from a school lesson. *Sustainability*, 15(18): 14025.

Kenny DA, Kaniskan B, McCoach DB (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research*, 44(3): 486–507.

Ko S, Chan SC (2025). A framework for the responsible integration of generative AI tools in learning. In: Çela E, Vajjhala NR, Fonkam MM (eds.), *Teachers' Roles and Perspectives on AI Integration in Schools*, 163–194. IGI Global Scientific Publishing.

Michel-Villarreal R, Vilalta-Perdomo E, Salinas-Navarro DE, Thierry-Aguilera R, Gerardou FS (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Education Sciences*, 13(9): 856.

Prilop CN, Mah DK, Jacobsen LJ, Hansen RR, Weber KE, Hoya F (2025). Generative AI in teacher education: Educators' perceptions of transformative potentials and the triadic nature of AI literacy explored through AI-enhanced methods. *Computers and Education: Artificial Intelligence*, 9: 100471.

R Core Team (2024). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Wang S, Xu L, Liu J, Zhai Y (2026). Addressing the challenges of AI-generated assignment submissions in education: Insights and strategies. *Journal of Data Science*, 24(1): 1–7.

Xie Y, Wu S, Chakravarty S (2023). AI meets AI: Artificial intelligence and academic integrity - a survey on mitigating AI-assisted cheating in computing education. In: *Proceedings of the 24th Annual Conference on Information Technology Education*, SIGITE '23, 79–83. Association for Computing Machinery, New York, NY, USA.