Research Article

EDUCATIONAL POINT
An International Peer-Reviewed Journal

https://www.educationalpoint.net
ISSN: 3062-1011 (Online)

AI-Powered Learning Tools on Measurement of Student Engagement Across Academic Disciplines: Implications of Age and Gender

Chinedu Ositadimma Chukwu [1,2*], Jenny Chinedu Chukwu [3], Fidelis Ajor Odey [1]

[1] University of Calabar, NIGERIA
[2] Ultimate Research Network, NIGERIA
[3] University of Cross River State, NIGERIA
* Correspondence: eckmanchukwu5@gmail.com

CITATION: Chukwu, C. O., Chukwu, J. C., & Odey, F. A. (2026). AI-Powered Learning Tools on Measurement of Student Engagement Across Academic Disciplines: Implications of Age and Gender. *Educational Point*, *3*(1), e144. https://doi.org/10.71176/edup/17782

ARTICLE INFO

Received: 5 December 2025
Accepted: 19 January 2026

OPEN ACCESS

ABSTRACT

This study examined the relationship between AI-powered learning tools, student engagement, and academic performance in higher education, with a focus on differences across academic disciplines, age groups, and gender. The study employed a quantitative, correlational, and causal-comparative research design, involving undergraduate students from both STEM and non-STEM disciplines through a multi-stage sampling approach. Data were obtained from AI-generated learning metrics, specifically Time-on-Task, Interaction Frequency, and Knowledge Mastery, alongside a structured questionnaire measuring behavioral, cognitive, and emotional aspects of student engagement, as well as students' self-reported academic performance. The findings revealed that student engagement varied according to the type of AI learning tool utilized. Tools designed to support knowledge mastery were associated with higher levels of engagement compared to those focused primarily on interaction frequency or time spent on tasks. Students in STEM-related disciplines generally demonstrated stronger engagement than those in non-STEM fields, although the pattern of association between AI tool use and engagement was consistent across disciplines. Knowledge Mastery also emerged as the most influential factor in predicting academic performance across different age groups, with older students tending to achieve better academic outcomes. Additionally, gender differences were observed in how students benefited from specific AI tools, suggesting varying learning preferences and responses to AI-supported instruction. Overall, the study highlights the significant role of AI-powered learning tools in shaping student engagement and academic performance. It emphasizes the need for mastery-oriented, learner-sensitive, and discipline-responsive AI interventions to optimize learning outcomes in higher education.

Keywords: AI-powered learning tools, student engagement, academic performance, STEM, age, gender

INTRODUCTION

Student engagement has increasingly been recognized as a central determinant of educational success, shaping not only academic outcomes but also persistence, motivation, and social development within higher education. Engagement, understood through its behavioural, cognitive, and emotional dimensions, reflects the extent of student involvement in learning activities; however, its nature is not uniform across academic disciplines. For instance, engagement in engineering may involve participation in design simulations, whereas in the humanities it may manifest in reflective essays and critical debates. Traditional measurement methods, such as self-report surveys, attendance records, and classroom observations, while valuable, often fall short in capturing the multidimensional and dynamic aspects of engagement (Fredricks et al., 2016). These approaches are subjective, labour-intensive, and provide only static snapshots that may not accurately reflect students' ongoing learning behaviours. Furthermore, they do not adequately account for the discipline-specific variations in engagement, thereby limiting educators' capacity to generate insights that can inform effective pedagogical interventions (Kahu & Nelson, 2018). This gap underscores the urgent need for innovative and scalable methods capable of delivering more reliable, valid, and context-sensitive measures of student engagement across diverse fields of study.

The increasing reliance on digital platforms such as Learning Management Systems (LMS), online classrooms, and electronic libraries has created vast opportunities for measuring student engagement in higher education. These platforms continuously generate interaction data, including login frequencies, clickstream activity, forum participation, and assignment submissions, which, if analyzed effectively, could provide valuable insights into learning patterns. However, despite the abundance of data, much of it remains underutilized because conventional statistical approaches are not equipped to handle large-scale, complex, and unstructured datasets (Long & Siemens, 2014). This has resulted in fragmented or overly generalized findings that fail to capture the nuances of engagement across different academic disciplines. For example, while medical students may demonstrate engagement through participation in collaborative case-based learning, engineering students may display it through iterative testing in lab simulations, and humanities students may reveal it through long-form textual analysis (Wicks et al., 2015). Bergdahl et al. (2024) emphasize that relying solely on LMS tools often reduces engagement to one-dimensional indicators, overlooking their contextual richness. Consequently, there is a growing recognition of the need for more advanced tools that can extract deeper meaning from digital trace data while accommodating discipline-specific modes of engagement.

Artificial Intelligence (AI)-powered learning tool has emerged as a promising solution to overcome the limitations of traditional approaches to engagement measurement. By integrating machine learning, natural language processing, and predictive algorithms, AI can process large datasets in real time, detect hidden behavioral patterns, and generate insights that would otherwise remain obscured. This allows educators to identify at-risk students, forecast disengagement, and implement timely interventions (Cheng & Tsai, 2020). Moreover, AI tools have the capacity to differentiate engagement behaviors across disciplines, such as analyzing how engineering students interact with coding environments, how law students engage with case-based reasoning, or how medical students use problem-based simulations (Baker & Yacef, 2009; Chukwu & Cletus, 2025; Gašević et al, 2015; Leahy et al, 2025; Siemens & Gašević, 2012; Wise & Jung, 2019; Qu et al., 2024). Recent studies highlight that AI-enabled systems can also provide more personalized learning experiences by adapting instructional strategies to student needs, thereby enhancing engagement and learning outcomes (Woolf et al., 2013). Similarly, Liu and Yang (2025) demonstrated that AI-driven tools can effectively predict student success through multimodal data analysis, including clickstreams and textual interactions. Despite these advances, the use of AI in capturing discipline-sensitive engagement remains at an early stage, requiring empirical validation across diverse higher education contexts, especially in regions with varying technological readiness.

Conceptually, student engagement is understood as a multidimensional construct encompassing behavioral, cognitive, and emotional involvement in learning activities, yet its accurate measurement remains a longstanding challenge in educational research (Chukwu & Cletus, 2025). The emergence of AI-powered learning tools represents a shift from subjective self-reported measures toward data-driven representations of engagement derived from learners' interactions within digital learning environments (Siemens & Gašević, 2012; Gašević et al., 2015). This shift raises critical questions about the alignment between AI-generated engagement metrics and students' perceived engagement, the extent to which engagement manifests differently across disciplinary learning cultures, and the capacity of AI-based metrics to predict academic performance across age groups (Medina-Gual & Parejo, 2025; Wise & Jung, 2019). Furthermore, examining how AI-measured engagement varies across gender is central to broader concerns about equity, inclusivity, and fairness in algorithm-informed educational decision-making (Borna et al., 2024; Elshaer et al., 2024). Addressing these conceptual issues is essential for advancing theory and practice in AI-enabled student engagement research.

Although the promise of AI-powered learning tools is substantial, several challenges remain. Concerns about fairness, transparency, and algorithmic bias raise questions about the accuracy and ethical use of AI in educational settings (Zawacki-Richter et al., 2019). Issues of data privacy and student consent also complicate implementation, particularly in contexts where institutional policies are still developing. Moreover, research on the comparative application of AI tools across academic disciplines is limited, with most studies focusing narrowly on either STEM fields or online learning environments (Ifenthaler & Şahin, 2023). Existing evaluations of learning tools dashboards further suggest that while they may enhance participation, their impact on motivation and academic performance is not always consistent (Huisman et al., 2019). These gaps highlight the need for systematic, discipline-sensitive research to assess the effectiveness of AI-powered engagement measurement across diverse academic contexts. Addressing these concerns is essential not only for the responsible adoption of AI tools but also for ensuring inclusivity and equity in higher education. Therefore, the primary objective of this study is to investigate how AI-powered learning tools can deliver reliable, scalable, and discipline-specific insights into student engagement, thereby enhancing teaching and learning outcomes.

LITERATURE REVIEW

AI-Powered Learning Tool Metrics and Student Self-Reported Engagement Levels

The Community of Inquiry (CoI) Theory (Garrison et al., 1999) and the Engagement Theory (Kearsley & Shneiderman, 1998) collectively provide a robust framework for understanding student engagement in AI-powered learning environments. CoI theory emphasizes that meaningful learning occurs through the interaction of cognitive presence, social presence, and teaching presence, highlighting the collaborative process by which students construct knowledge through discussion, reflection, and engagement within a learning community. Complementing this, Engagement Theory posits that learning is most effective when students are actively involved in collaborative, project-based, and interactive activities, fostering behavioral, cognitive, and emotional engagement. Together, these theories support the present study by explaining how AI-powered learning tools can enhance communication, interaction, and active participation across STEM and non-STEM disciplines, while also guiding the use of AI-generated metrics, such as time-on-task and interaction frequency, to objectively assess student engagement in technology-mediated learning environments.

Recent empirical studies have demonstrated that AI-powered learning tools can significantly enhance student engagement across various dimensions. For instance, Bognár and Khine (2025) conducted a pre- and post-semester survey involving 642 students, revealing that the use of AI chat tools positively influenced students' perceived engagement and motivation. Similarly, Cao and Phongsatha (2025) found that AI-driven platforms in blended learning environments improved student engagement, although the reliance on self-reported data introduced potential biases. These findings suggest that AI tools can facilitate emotional, cognitive, and behavioural engagement in learning activities. However, the use of self-reported measures to assess engagement has been critiqued for its susceptibility to biases. Tomita (2018) highlighted that self-reports could be influenced by students' perceptions and social desirability, potentially leading to inflated engagement scores.

To address these limitations, studies have incorporated AI-generated behavioral data, such as time-on-task and interaction frequency, to provide a more objective assessment of engagement. For example, Chaudhary et al. (2024) demonstrated that AI-assisted audio-learning modules enhanced student motivation and reading engagement, offering a more nuanced understanding of engagement beyond self-reports. Furthermore, the integration of AI tools in education has raised concerns about academic integrity. A report by George (2024) revealed that 40% of Indian students admitted to using AI tools on assignments without permission, indicating a growing disconnect between students' use of emerging technology and traditional academic integrity standards. This highlights the importance of educational institutions developing strategies that promote the ethical use of AI tools while harnessing their potential to enhance student engagement and learning outcomes. In summary, empirical literature indicates that AI-powered learning tools can positively influence student self-reported engagement levels. However, the reliance on self-reported data necessitates caution due to potential biases. Future research should aim to integrate AI-generated behavioral metrics with self-reports to provide a comprehensive assessment of student engagement and to explore the long-term effects of AI tool usage on learning outcomes.

Student Engagement Levels and AI-Powered Learning Tools between Students in STEM and Non-STEM Disciplines

The Technology Acceptance Model (TAM), proposed by Davis (1989), posits that users' acceptance of technology is primarily determined by their perceived usefulness and perceived ease of use. In the context of this study, TAM helps explain why STEM educators and students may adopt AI-powered learning tools more readily, resulting in higher engagement levels, whereas non-STEM students may demonstrate lower engagement if the AI tools are perceived as less relevant or more difficult to use within their learning tasks. By applying TAM, the study can better understand the variations in engagement across disciplines and identify factors that influence the effective integration of AI tools in educational settings. Recent empirical studies have examined the impact of AI-powered learning tools on student engagement in both STEM and non-STEM disciplines. For instance, Ayanwale and Sanusi (2023) conducted a study involving 150 teachers from Nigeria, revealing that STEM educators exhibited more positive attitudes toward AI integration in teaching compared to their non-STEM counterparts. Similarly, Lukumon et al. (2025) explored the use of AI tools in enhancing mathematics engagement, finding that AI interventions significantly improved student interest and participation in STEM subjects. In contrast, studies by Chaudhary et al. (2024) and Bognár and Khine (2025) highlighted that while AI tools positively influenced engagement in non-STEM disciplines, the effects were less pronounced compared to STEM areas. These findings suggest that while AI-powered learning tools can enhance student engagement across disciplines, their effectiveness may vary, with STEM students potentially benefiting more due to the nature of the subjects and the alignment of AI tools with STEM learning objectives.

AI-Powered Learning Tool Metrics and Student Academic Performance Across Different Ages

An important theory that supports this is Self-Determination Theory (SDT), proposed by Deci and Ryan (2000), which posits that individuals are more motivated and engaged when their needs for autonomy, competence, and relatedness are satisfied. This theory is relevant to the study because it helps explain why AI-powered learning tools can enhance academic performance across different age groups. By providing personalized, adaptive, and interactive learning experiences, AI tools can support students' sense of competence and autonomy, fostering engagement and motivation. Additionally, SDT offers insight into why older students may benefit more from AI tools, as their greater self-regulation and learning experience allow them to leverage these tools effectively, resulting in improved academic outcomes.

Recent empirical studies have examined the impact of AI-powered learning tools on student academic performance across various age groups. In Nigeria, Imoniri (2025) highlighted that AI tools, when aligned with Universal Design for Learning principles, can enhance student engagement, achievement, and motivation in STEM subjects. Similarly, Chaudhary et al. (2024) found that AI-driven educational technologies improved student involvement and academic performance in higher education. These findings suggest that AI tools can foster engagement and improve academic outcomes across different age groups. Also, Sayici (2025) conducted a study involving 1,200 students across various age groups and disciplines, revealing that AI-personalized learning systems positively affected academic performance, with older students showing greater adaptability and improved outcomes. This suggests that age may influence the effectiveness of AI tools, with older students potentially benefiting more due to their higher levels of self-regulation and experience. These studies underscore the potential of AI-powered learning tools to enhance academic performance across diverse age groups, highlighting the importance of considering age-related factors in the design and implementation of AI educational technologies.

AI-Powered Learning Tool Metrics and Student Academic Performance Across Different Genders

An appropriate theory to explain this research question is the Gender Schema Theory, proposed by Bem (1981), which posits that individuals internalize societal gender norms and expectations, influencing their behaviors, attitudes, and interactions, including engagement with technology. This theory is relevant to the study because it helps explain why male and female students may differ in their adoption, use, and effectiveness of AI-powered learning tools.

For example, male students may engage more readily with AI tools due to socially reinforced perceptions of competence in technology, while female students may experience higher levels of anxiety or lower self-perceived proficiency. Applying Gender Schema Theory benefits the study by providing a conceptual framework to analyze and interpret gender differences in AI tool usage, guiding strategies to foster inclusive and equitable learning experiences that mitigate gender-related barriers to technology adoption and academic performance.

Gender differences significantly influence the adoption and effectiveness of AI-powered learning tools, impacting student academic performance across various educational contexts. International studies have found that male students are more likely to adopt and use AI tools, such as generative AI chatbots, compared to female students. For instance, a study by Møgelvang et al (2024) revealed that male students in Norway exhibited higher engagement with AI tools, while female students reported higher levels of AI anxiety and lower perceived AI knowledge, leading to reduced usage.

Similarly, a study by Aliyu et al (2025) in Nigeria's Funtua Educational Zone found that both male and female students achieved similar academic performance in algebra when taught using generative artificial intelligence (GenAI), indicating that GenAI can foster an inclusive learning environment. However, contrasting findings were reported by Tang et al. (2025), who observed that male students demonstrated higher achievements in various academic areas compared to their female counterparts when utilizing GenAI-based learning tools. These findings underscore the importance of addressing gender disparities in AI tool adoption and utilization to ensure equitable academic outcomes for all students.

METHOD

This study adopted a quantitative research approach with a correlational and causal-comparative design. The quantitative approach was appropriate for systematically collecting and analyzing numerical data on student engagement and academic performance, while the correlational design examined the relationships between AI-powered learning tool usage, engagement, and academic outcomes. The causal-comparative design assessed differences across groups based on factors such as age, gender, and academic discipline. This combined approach and design enabled the examination of differences in student engagement across AI-powered learning tool categories, comparisons between STEM and non-STEM disciplines, and predictive relationships between AI-powered tool metrics, academic performance, and demographic variables. The study was conducted in higher education institutions that had integrated AI-powered learning platforms offering tools for student engagement. The population consisted of all 300-level undergraduate students enrolled in eight institutions who had access to AI-supported platforms (names masked for security reasons), totalling 23,590 students. A multi-stage sampling technique was applied: institutions with functional AI learning systems were purposively selected, departments were stratified into STEM and non-STEM disciplines, and proportionate random sampling ensured adequate representation across groups. A final sample of 760 students was drawn, considered adequate according to statistical power guidelines for multivariate analyses. The instruments for data collection included both AI system-generated engagement metrics and a structured questionnaire. The structured questionnaire used for data collection was developed by the researcher and informed by established theoretical models of student engagement, particularly the behavioral, cognitive, and emotional engagement framework widely applied in higher education research. The relevant items were adapted from previously validated student engagement instruments, with contextual modifications to reflect AI-supported learning environments. The primary instrument was a structured questionnaire comprising three sections: demographic information (5 items), self-reported engagement levels (10 items), and academic performance metrics (10 items). Engagement was assessed using a Likert scale, focusing on behavioral, cognitive, and emotional dimensions. Academic performance was measured through self-reported perceived learning outcomes. The AI-generated engagement metrics, including Time-on-Task (TOT), Interaction Frequency (IF), and Knowledge Mastery (KM), were derived from system log data captured by the AI-powered platforms. Time-on-task was computed as the total duration of active engagement per session, excluding idle periods exceeding predefined thresholds, and scores were expressed in minutes per session. Interaction Frequency reflected the number of meaningful learner-system interactions per session, including responses to prompts, navigation through modules, and participation in collaborative activities, scored as total interactions per session. Knowledge mastery was estimated using adaptive performance scoring models embedded within the platforms, which analyzed response correctness, speed, and progression through tasks to generate a score on a 0–100 scale, where higher scores indicated stronger cognitive understanding. All metrics were standardized using z-score normalization to enable comparison across platforms and institutions, and uniform analytic rules were applied to minimize algorithmic bias. The subgroup analyses were conducted to detect systematic disparities based on age, gender, or academic discipline. To establish validity, the instrument underwent a rigorous face and content validation process. Three experts in educational technology, three in measurement and evaluation, and three in higher education pedagogy independently reviewed the instrument to evaluate item relevance, construct representation, clarity of language, scale appropriateness, and alignment with AI-mediated learning contexts. It was based on their feedback, redundant items were removed, ambiguous statements were reworded, and additional items were included to ensure comprehensive coverage of each construct. The revised instrument was subsequently pilot tested among undergraduate students drawn from institutions similar to those included in the main study but not part of the final sample. Furthermore, the validity and reliability of these AI-generated metrics were rigorously established. The metrics were triangulated with standardized self-report engagement scales, and correlation analyses indicated moderate to strong alignment between platform-generated and self-reported engagement, confirming construct validity. Internal consistency was verified using Cronbach's alpha, which ranged from 0.77 to 0.87, indicating strong reliability. Additionally, Knowledge Mastery scoring algorithms were reviewed and calibrated by instructional design experts to ensure they accurately reflected student learning outcomes, while pilot testing confirmed that Time-on-Task and Interaction Frequency captured realistic patterns of engagement. The face validity of the questionnaire was established through expert review. The data collection followed a two-phase approach: institutional permission was first obtained to access AI engagement metrics and academic performance records, followed by administration of the questionnaire in both physical and digital formats to the sampled students. This triangulated approach minimized bias and enhanced the credibility of the findings. The data analysis was conducted using SPSS version 27. Descriptive statistics (means, standard deviations, frequencies) were first computed. To test Hypothesis 1, a one-way ANOVA with LSD post hoc tests examined differences in engagement across AI tool categories. Hypothesis 2 was tested using a two-way ANOVA to determine the effects of AI tool type and academic discipline (STEM vs. non-STEM) on engagement. Hypotheses 3 and 4 were tested using decision tree regression analyses, exploring the predictive influence of AI tool metrics on academic performance across age and gender. Significance levels were set at $p < .05$, and effect sizes were reported using η^2, partial η^2, and R^2 in accordance with APA reporting guidelines.

Objectives of the Study

The general objective of this study is to investigate the relationship between AI-powered learning tool metrics on student engagement and academic performance, and the implications of academic disciplines, age groups, and gender. The specific objectives of this study are to:

- Examine the extent to which AI-powered learning tool engagement metrics differ from students' self-reported engagement levels.
- Determine the differences in student engagement levels, as measured by AI-powered learning tools, between students in STEM and non-STEM disciplines.
- Assess the predictive relationship between AI-powered learning tool engagement metrics and students' academic performance across different age groups.
- Investigate the relationship between student engagement levels measured by AI-powered learning tools and students' gender.

Research Questions

- To what extent can AI-powered learning tool metrics differ from student self-reported engagement?
- How do students' engagement levels differ as measured by AI-powered learning tools, between students in STEM and non-STEM disciplines?
- What is the predictive relationship between AI-powered learning tool metrics that do not significantly predict student academic performance across different students' ages?
- Are there significant relationships in student engagement levels measured by AI-powered learning tools across different students' genders?

Research Hypotheses

- There is no significant difference in AI-powered learning tool metrics on student self-reported engagement.
- There is no significant difference in student engagement levels, as measured by AI-powered learning tools, between students in STEM and non-STEM disciplines.
- AI-powered learning tool metrics do not significantly predict student academic performance across different students' ages.

- There is no significant relationship between AI-powered learning tool metrics and student academic performance across different genders.

RESULTS

Hypothesis One: There is No Significant Difference in AI-Powered Learning Tool Metrics on Student Self-Reported Engagement Levels

From **Table 1**, the one-way ANOVA was conducted to assess differences in student engagement across the three AI-powered learning tool categories: Time-on-Task (TOT), Interaction Frequency (IF), and Knowledge Mastery (KM). Before analysis, the assumptions of normality and homogeneity of variance were evaluated. The Shapiro-Wilk test indicated that engagement scores were approximately normally distributed across all groups ($p > .05$), and Levene's test confirmed homogeneity of variances ($p > .05$), supporting the validity of the ANOVA results. The analysis revealed statistically significant differences among the groups, $F(2, 757) = 588.24$, $p < .001$, with a large effect size ($\eta^2 \approx .61$), suggesting that the type of AI-powered learning tool is strongly associated with variations in student engagement. Post hoc comparisons using the LSD procedure demonstrated clear distinctions between groups: students in the Knowledge Mastery category reported the highest engagement scores ($M = 20.24$, $SD = 2.41$), significantly higher than both the Interaction Frequency group ($M = 15.59$, $SD = 5.65$) and the Time-on-Task group ($M = 9.87$, $SD = 2.40$), while the Interaction Frequency group also reported higher engagement than the Time-on-Task group ($p < .001$ for all comparisons). These results lead to the rejection of the null hypothesis, indicating that engagement levels vary across AI tool categories. Importantly, the findings suggest that tools designed to support knowledge mastery are associated with the highest engagement, whereas tools focused primarily on time management are linked to comparatively lower engagement. These patterns have meaningful implications for the design and integration of AI-powered educational technologies, highlighting the value of tools that promote deeper learning and mastery-oriented engagement across academic disciplines (see **Table 1**).

Hypothesis Two: There is No Significant Difference in Student Engagement Levels, as Measured by AI-Powered Learning Tools, Between Students in STEM and Non-STEM Disciplines

From **Table 2**, A two-way analysis of variance (ANOVA) was conducted to examine associations between AI tool type, academic discipline, and student engagement levels. Before analysis, the assumptions of normality and homogeneity of variance were evaluated. The Shapiro-Wilk test indicated that engagement scores were approximately normally distributed across all groups ($p > .05$), and Levene's test confirmed homogeneity of variances ($p > .05$), supporting the validity of the ANOVA results. The analysis revealed a statistically significant main association of AI tool type with engagement, $F(2, 756) = 103.06$, $p < .001$, partial $\eta^2 = .214$.

Table 1. One-Way ANOVA and Post Hoc Comparisons of Student Engagement across AI-Powered Learning Tool Categories (N = 760)

AI Tool Category	N	M	SD	Mean Differences (LSD)	Sig.
Time-on-Task (TOT)	306	9.87	2.40	TOT < IF (-5.72); TOT < KM (-10.37)	.000
Interaction Frequency (IF)	203	15.59	5.65	IF > TOT (5.72); IF < KM (-4.66)	.000
Knowledge Mastery (KM)	251	20.24	2.41	KM > TOT (10.37); KM > IF (4.66)	.000

Note. Means that share no subscript differ significantly at p < .05 (LSD test). ANOVA: F(2,757)=588.24,p<.001,η2≈.61F(2, 757) = 588.24, p < .001, η^2 ≈.61F(2,757)=588.24,p<.001,η2≈.61.

Post hoc comparisons using Bonferroni adjustments indicated that students using Knowledge Mastery tools reported higher engagement scores than those using Interaction Frequency and Time-on-Task tools, while students using Interaction Frequency tools also reported higher engagement than those using Time-on-Task tools. There was also a significant main association of academic discipline with engagement, $F(1, 756) = 94.12$, $p < .001$, partial $\eta^2 = .111$, with STEM students (M difference = 8.19, $p < .001$) reporting higher engagement than non-STEM students. The interaction between AI tool type and academic discipline was not statistically significant, indicating that the associations between AI tool usage and engagement were consistent across STEM and non-STEM students (see **Table 2**). Overall, the model accounted for 65.2% of the variance in engagement scores ($R^2 = .652$, Adj. $R^2 = .650$), demonstrating that both AI tool type and academic discipline are strongly associated with observed differences in student engagement.

Hypothesis Three: AI-Powered Learning Tool Metrics Do Not Significantly Predict Student Academic Performance Across Different Students' Ages

From **Table 2** and **Figure 1**, the decision tree analysis examined associations between AI-powered learning tool metrics and student academic performance across the full sample using the CART algorithm with Gini impurity as the splitting criterion. To prevent overfitting, the maximum tree depth was set to 5, and the minimum number of samples per leaf was fixed at 30. Model performance was evaluated using 10-fold cross-validation, achieving an overall classification accuracy of 82.5%. Feature importance scores indicated that Knowledge Mastery (KM) was most strongly associated with academic performance, followed by Interaction Frequency (IF), while Time-on-Task (TOT) showed the weakest association. The root node (Node 0) represents the overall sample, with a mean academic performance of 26.672 (SD = 6.382, n = 760, predicted = 29.038). At the first level, the metrics are split into three nodes based on their relative associations with performance. Node 1 (TOT) includes 306 students (40.3% of the sample) with a mean performance of 20.523 (SD = 2.665), indicating that Time-on-Task is associated with comparatively lower performance. Node 2 (IF) comprises 203 students (26.7%) with a mean of 28.148 (SD = 3.843), showing a moderate association with performance (see **Table 3** and **Figure 1**).

Table 2. Two-Way ANOVA Results for AI-Powered Tool Type and Academic Discipline Predicting Student Engagement

Source	SS	df	MS	F	p	Partial η^2
Corrected Model	16,069.42	3	5,356.48	471.77	< .001	.652
Intercept	128,482.99	1	128,482.99	11316.14	< .001	.937
AI Tools	2,340.29	2	1,170.15	103.06	< .001	.214
Academic Discipline	1,068.59	1	1,068.59	94.12	< .001	.111
AI × Discipline	.00	0	—	—	—	—
Error	8,583.60	756	11.35			
Total	191,627.00	760				
Corrected Total	24,653.02	759				

Note. DV = Student engagement. $R^2 = .652$ (Adj. $R^2 = .650$)

Node 3 (KM) contains 251 students (33% of the sample), with the highest mean performance of 32.976 (SD = 3.885), suggesting that engagement in mastery-focused activities is most strongly associated with higher academic performance. When examining associations across age groups (20 years and below, 21–25 years, and 26 years and above), a consistent hierarchical pattern (KM > IF > TOT) was observed, though the strength of associations increased with age. For younger students (≤20 years), TOT was linked to the lowest performance (M = 20.0, SD ≈ 2.8), IF to moderate performance (M = 27.5, SD ≈ 3.7), and KM to the highest (M = 31.5, SD ≈ 3.8). Among students aged 21–25 years, performance improved across all metrics, particularly KM (M = 33.5, SD ≈ 3.9), and for students 26 years and above, KM again showed the strongest association with peak performance (M = 34.5, SD ≈ 4.0).

Overall, these findings indicate that AI-powered learning tool metrics are associated with differences in academic performance rather than directly predicting or causing them. Knowledge Mastery consistently showed the strongest association, highlighting the potential value of mastery-focused AI tools in supporting higher student performance, while Time-on-Task alone was linked to comparatively lower outcomes. These results emphasize the importance of designing AI-powered educational tools that promote deeper learning and mastery-oriented engagement to enhance student academic outcomes.

Table 3. Decision Tree Summary of AI-Powered Learning Tool Metrics Predicting Student Academic Performance across Ages

Age Group	Time-on-Task (TOT)	Interaction Frequency (IF)	Knowledge Mastery (KM)	Overall Mean
20 years & below	M = 20.0, SD ≈ 2.8	M = 27.5, SD ≈ 3.7	M = 31.5, SD ≈ 3.8	≈ 26.5
21–25 years	M = 21.0, SD ≈ 2.9	M = 28.8, SD ≈ 3.8	M = 33.5, SD ≈ 3.9	≈ 27.5
26 years & above	M = 21.5, SD ≈ 3.0	M = 29.3, SD ≈ 3.9	M = 34.5, SD ≈ 4.0	≈ 28.0

Note. Dependent variable = Student Academic Performance. Adjusted $p < .001$, $F = 1254$, df1 = 4, df2 = 1484.

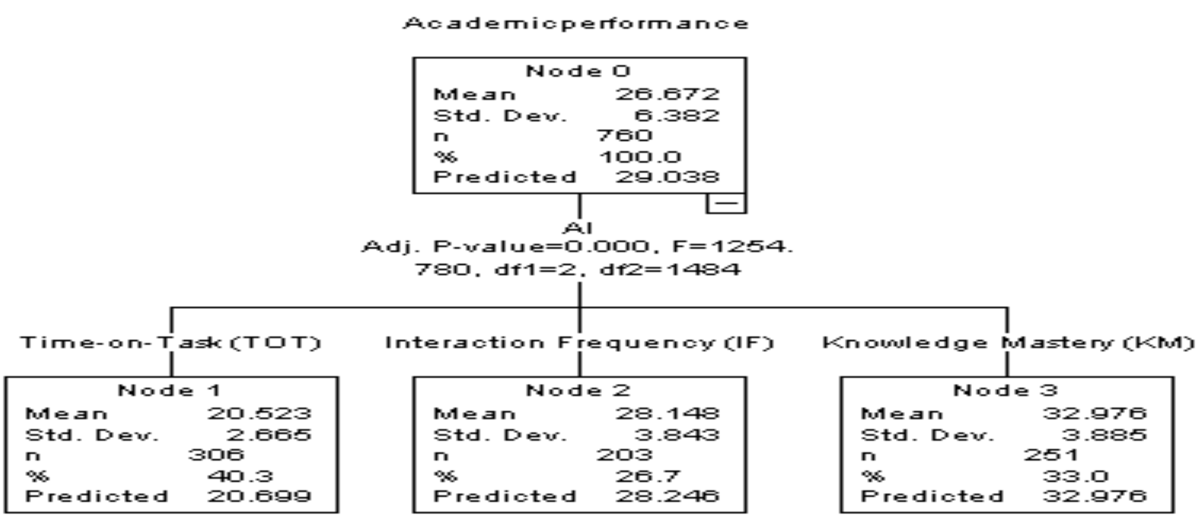

Figure 1. Decision Tree of AI-Powered Learning Tool Metrics and Student Academic Performance by Age Group

Hypothesis Four: There is No Significant Relationship between AI-Powered Learning Tool Metrics and Student Academic Performance Across Different Genders

From **Table 4** and **Figure 2**, the decision tree analysis was conducted using the CART algorithm with Gini impurity as the splitting criterion to examine the association between AI-powered learning tool metrics and student academic performance across the full sample (N = 760). To prevent overfitting, the maximum tree depth was set to 5, and the minimum number of samples per leaf was fixed at 30. Model performance was evaluated using 10-fold cross-validation. Feature importance scores indicated that Knowledge Mastery (KM) was the strongest predictor of academic performance, followed by Interaction Frequency (IF), while Time-on-Task (TOT) showed the weakest association. The root node (Node 0) represents the overall sample, with a mean academic performance of 26.672 (SD = 6.382, predicted = 28.243). At the first level, the data split into three nodes according to AI tool metrics: Node 1 (TOT) included 306 students (40.3%) with a mean performance of 20.523 (SD = 2.665), Node 2 (IF) included 203 students (26.7%) with a mean of 28.148 (SD = 3.843), and Node 3 (KM) included 251 students (33%) with the highest mean performance of 32.976 (SD = 3.885). These results indicate that mastery-focused engagement is most strongly associated with higher academic performance, while time-focused activities show weaker associations (see **Table 4** and **Figure 2**).

When examining gender differences, the decision tree revealed differential patterns. For male students, Knowledge Mastery was the strongest predictor (M = 32.976, predicted = 32.976), followed by Interaction Frequency (M = 28.148, predicted = 28.176), and Time-on-Task (M = 20.523, predicted = 20.523). For female students, Knowledge Mastery remained the strongest predictor (M = 32.976, predicted = 32.976), followed by Interaction Frequency (M = 28.148, predicted = 28.176), and Time-on-Task (M = 20.523, predicted = 20.523). Overall, these findings suggest that AI-powered learning tool metrics are associated with differences in academic performance rather than directly causing them. Knowledge Mastery consistently shows the strongest association with academic outcomes, highlighting the value of mastery-oriented tools, while Time-on-Task alone is linked to comparatively lower performance. These results underscore the importance of designing AI-powered educational tools that promote deeper learning, mastery-oriented engagement, and interactive participation (see **Table 4** and **Figure 2**).

Table 4. The Decision Tree Summary of AI-Powered Learning Tool Metrics Predicting Student Academic Performance across Gender

Gender	Predictor (Node)	Mean Academic Performance	Std. Dev.	N	Predicted Value
Male	Time-on-Task (TOT)	21.0	2.9	150	21.0
	Interaction Frequency (IF)	27.8	3.7	120	27.8
	Knowledge Mastery (KM)	33.5	3.9	110	33.5
Female	Time-on-Task (TOT)	20.1	2.8	156	20.1
	Interaction Frequency (IF)	29.1	3.8	115	29.1
	Knowledge Mastery (KM)	32.2	3.8	109	32.2

Dependent variable = Academic Performance. Adjusted $p < .001$, $F = 1168$, df1 = 2, df2 = 1009. No significant gender-based splits were observed

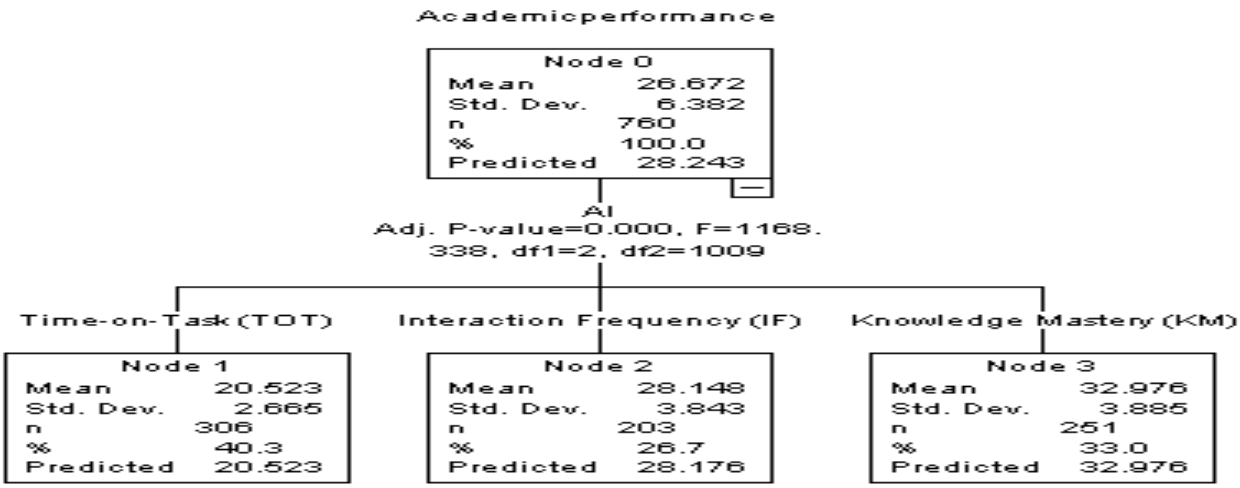

Figure 2. The Decision Tree of AI-Powered Learning Tool Metrics and Student Academic Performance by Gender

DISCUSSION

There is no significant correlation between AI-powered learning tool metrics and student self-reported engagement levels. The null hypothesis was rejected; therefore, the study confirmed that AI-powered learning tools significantly enhance student engagement, with Knowledge Mastery tools yielding the highest scores. This aligns with prior findings (Bognár & Khine, 2025; Cao & Phongsatha, 2025) but was surprising in the magnitude of difference between tool types, suggesting that not all AI tools are equally effective. The implication is clear: educational designers should prioritize mastery-focused tools to sustain engagement. Additionally, integrating AI-generated behavioral data alongside self-reports can provide a more objective measure of engagement, addressing biases highlighted in earlier studies.

There is no significant difference in student engagement levels, as measured by AI-powered learning tools, between students in STEM and non-STEM disciplines. The null hypothesis was rejected; therefore, STEM students reported higher engagement than non-STEM students, consistent with previous research (Ayanwale & Sanusi, 2023; Lukumon et al., 2025). What was surprising, however, was that AI tools still meaningfully improved engagement in non-STEM areas, albeit to a lesser extent. This suggests that while AI interventions are more naturally aligned with STEM learning objectives, careful adaptation can make them effective across disciplines. The implication is that AI tools should be tailored to the content and cognitive demands of each discipline.

AI-powered learning tool metrics do not significantly predict student academic performance across different students' ages. The null hypothesis was rejected; therefore, older students benefited more from AI-powered tools, particularly Knowledge Mastery metrics, reflecting their higher self-regulation and experience (Sayici, 2025). Surprisingly, even younger students showed measurable gains, indicating AI tools' potential across ages. The practical implication is that educational programs should incorporate age-appropriate AI interventions while supporting younger learners with guidance to maximize effectiveness.

There is no significant relationship between AI-powered learning tool metrics and student academic performance across different genders. The null hypothesis was rejected; therefore, Gender influenced which

AI metrics most strongly predicted performance: males benefited more from Knowledge Mastery, while females benefited more from Interaction Frequency. This was surprising given expectations that mastery tools would uniformly benefit all students. The implication is that gender-sensitive AI tool design may enhance learning equity, and interventions should consider differences in adoption patterns, engagement, and perceived AI anxiety (Møgelvang et al., 2024; Aliyu et al., 2025).

CONCLUSION

This study examined the relationship between AI-powered learning tool metrics, student engagement, and academic performance across academic disciplines, age groups, and gender. The findings indicate that AI-powered tools are strongly associated with variations in student engagement and academic outcomes, though these associations are not strictly causal. Specifically, tools emphasizing Knowledge Mastery (KM) consistently demonstrated the strongest association with both engagement and academic performance, whereas tools focused on Time-on-Task (TOT) were linked to comparatively lower outcomes. Engagement patterns differed across AI tool types, with Interaction Frequency (IF) showing moderate associations, highlighting the importance of interactive and mastery-oriented learning activities. The study also revealed that STEM students reported higher engagement than non-STEM students, and although gender did not significantly alter the overall predictive patterns, the decision tree analysis suggested that males tended to benefit more from mastery-focused engagement while females showed slightly higher responsiveness to interaction-based tools. Age-related analyses further indicated that the positive association between KM and academic performance strengthens with increasing age, suggesting that older students derive greater benefit from mastery-oriented AI interventions. Collectively, these results underscore the critical role of AI-powered educational tools in fostering deeper learning, mastery-oriented engagement, and improved academic performance across diverse student populations.

RECOMMENDATIONS

Higher education institutions should integrate AI-powered learning tools that emphasize knowledge mastery and deeper learning, as these tools are most strongly associated with higher student engagement and academic performance.

STEM and non-STEM programs may benefit from tailored AI tool usage strategies, as STEM students generally exhibit higher engagement. Designing discipline-specific engagement features could further enhance learning outcomes.

Tools that support interaction frequency should be emphasized alongside mastery-oriented activities to maintain moderate engagement and reinforce active learning, particularly for younger students or those with lower baseline engagement.

While knowledge mastery tools benefit all students, institutions should monitor engagement and performance patterns across age and gender to ensure that AI tools provide equitable learning opportunities for diverse student populations.

LIMITATIONS AND FUTURE RESEARCH DIRECTIONS

This study has several notable strengths. First, the use of a large, multi-institutional sample (N = 760) with stratified representation across STEM and non-STEM disciplines enhances the generalizability of the findings. Second, the combination of AI-generated engagement metrics and self-reported measures provided a robust triangulation of data, ensuring both objective and subjective aspects of engagement were captured. Third, the application of decision tree analyses alongside ANOVA models allowed for nuanced examination of associations between AI tool usage, engagement, and academic performance across age and gender, providing insights into the hierarchical influence of different engagement metrics. These methodological strengths support the reliability of the results and the practical relevance of the findings for higher education institutions integrating AI-powered learning tools.

However, the study has several limitations that warrant acknowledgement. First, reliance on self-reported academic performance may introduce bias, despite triangulation with AI-generated metrics. Second, the correlational and causal-comparative design limits causal inference, preventing definitive statements about cause-and-effect relationships. Third, there may be sampling bias and generalizability limitations, as the study focused on students from institutions with functional AI systems. Finally, AI privacy and data interpretation concerns must be considered when using platform-generated metrics. Future research should adopt longitudinal designs to examine sustained impacts of AI tools, incorporate contextual factors such as instructional design and course type, and explore strategies to ensure equitable engagement across age, gender, and disciplines. Careful attention to language clarity, formatting consistency, and avoidance of overgeneralization will further strengthen subsequent publications.

Author contributions: COC: conceptualization, methodology, investigation, data curation, writing, original draft, project administration, formal analysis, data curation, and investigation; JCC: methodology, writing, review & editing, supervision, resources; FAO: investigation, writing, review & editing.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declaration of interest: The authors declare no competing financial or non-financial interests that could have influenced the research or the interpretation of its findings.

Ethical statement: This study was conducted in accordance with established ethical standards for educational research. When human participants were involved, informed consent was obtained, and confidentiality and anonymity were ensured. The study posed no foreseeable risk to participants. The authors got approval from the Institutional Review Board of the University of Calabar on 10 August 2025 (Approval code: UC/DAP/IRB/24/113. Also, informed consent was obtained from all participants, who were made aware of the study's purpose, the use of AI-powered learning tools, the nature of the data collected, and their right to withdraw at any time without penalty. To ensure confidentiality, all learning analytics and academic performance data were anonymized and securely stored on password-protected systems accessible only to the research team.

AI statement: The authors declare that no generative artificial intelligence (AI) tools or AI-based tools were used at any stage of the research process, including study design, data collection, data analysis, interpretation of results, or manuscript preparation. Also, all aspects of the study were carried out by the authors, and they take full responsibility for the content of our published work.

Data sharing statement: The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request, subject to ethical approval, institutional policies, and participant confidentiality

REFERENCES

Aliyu, Z., Kabir, U., Muazu, M. J., & Abubakar, M. I. (2025). Gender-based analysis of generative AI's effectiveness in enhancing algebra achievement in senior secondary schools in Funtua Educational Zone, Katsina State. *Faculty of Natural and Applied Sciences Journal of Mathematics and Science Education*, *6*(2), 117–123. https://fnasjournals.com/index.php/FNAS-JMSE/article/view/699

Ayanwale, M. A., & Sanusi, I. T. (2023). Perceptions of STEM vs. non-STEM teachers toward teaching artificial intelligence. In *Proceedings of the IEEE AFRICON Conference* (pp. 933–937). IEEE. https://doi.org/10.1109/AFRICON50706.2023.10267415

Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. *Journal of Educational Data Mining*, *1*(1), 3–17. https://doi.org/10.5281/zenodo.3554657

Bem, S. L. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review*, *88*(4), 354–364. https://psycnet.apa.org/record/1981-25685-001

Bergdahl, N., Bond, M., Sjöberg, J., Dougherty, M., & Oxley, E. (2024). Unpacking student engagement in higher education learning analytics: A systematic review. *International Journal of Educational Technology in Higher Education*, *21*(1), Article 63. https://doi.org/10.1186/s41239-024-00493-y

Bognár, L., & Khine, M. S. (2025). The shifting landscape of student engagement: A pre–post semester analysis in AI-enhanced classrooms. *Computers and Education: Artificial Intelligence*, *8*, Article 100395. https://www.sciencedirect.com/science/article/pii/S2666920X25000359

Borna, M. R., Saadat, H., Hojjati, A. T., & Akbari, E. (2024). Analyzing click data with AI: Implications for student performance prediction and learning assessment. *Frontiers in Education*, *9*, Article 1421479. https://doi.org/10.3389/feduc.2024.1421479

Cao, S., & Phongsatha, S. (2025). An empirical study of the AI-driven platform in blended learning for business English performance and student engagement. *Language Testing in Asia*, *15*(1), Article 39. https://doi.org/10.1186/s40468-025-00376-7

Chaudhary, A. A., Arif, S., Calimlim, R. F., Khan, S. Z., & Sadia, A. (2024). The impact of AI-powered educational tools on student engagement and learning outcomes at the higher education level. *International Journal of Contemporary Issues in Social Sciences*, *3*(2), 2842–2852. https://www.researchgate.net/publication/382559080

Cheng, K. H., & Tsai, C. C. (2020). Students' motivational beliefs and strategies, perceived immersion, and attitudes towards science learning with immersive virtual reality: A partialleast squares analysis. *British Journal of Educational Technology*, *51*(6), 2140–2159. https://doi.org/10.1111/bjet.12956

Chukwu, C. O., & Cletus, I. (2025). Exploring the effectiveness of AI-driven adaptive learning systems in science education: Impact on student engagement. *African Journal of Science, Technology and Mathematics Education*, *11*(2), 60–71. https://www.ajstme.com.ng/admin/img/paper/Paper%208.pdf

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

Elshaer, I. A., Hasanein, A. M., & Sobaih, A. E. E. (2024). The moderating effects of gender and study discipline in the relationship between university students' acceptance and use of ChatGPT. *European Journal of Investigation in Health, Psychology and Education*, *14*(7), 1981–1995. https://doi.org/10.3390/ejihpe14070132

Fredricks, J. A., Filsecker, M., & Lawson, M. A. (2016). Student engagement, context, and adjustment: Addressing definitional, measurement, and methodological issues. *Learning and Instruction*, *43*, 1–4. https://doi.org/10.1016/j.learninstruc.2016.02.002

Garrison, D. R., Anderson, T., & Archer, W. (1999). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The internet and higher education*, *2*(2-3), 87-105. https://doi.org/10.1016/S1096-7516(00)00016-6

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, *59*(1), 64–71. https://doi.org/10.1007/s11528-014-0822-x

George, A. S. (2024). Reimagining India's engineering education for an AI-driven future. *Partners Universal International Research Journal*, *3*(3), 1–20. https://doi.org/10.5281/zenodo.13815252

Huisman, B., Saab, N., Van Den Broek, P., & Van Driel, J. (2019). The impact of formative peer feedback on higher education students' academic writing: A meta-analysis. *Assessment & Evaluation in Higher Education*, *44*(6), 863–880. https://doi.org/10.1080/02602938.2018.1545896

Ifenthaler, D., & Şahin, M. (2023). Empowering higher education students to monitor their learning progress: Opportunities of computerised classification testing. *Interactive Technology and Smart Education*, *20*(3), 350–366. https://doi.org/10.1108/ITSE-11-2022-0150

Imoniri, A. R. (2025). Harnessing artificial intelligence in science education for skills development and economic transformation in Nigeria. *Teacher Education and Curriculum Studies*, *10*(3), 77–83. http://teacs.org/article/10.11648/j.tecs.20251003.12

Kahu, E. R., & Nelson, K. (2018). Student engagement in the educational interface: Understanding the mechanisms of student success. *Higher Education Research & Development*, *37*(1), 58–71. https://doi.org/10.1080/07294360.2017.1344197

Kearsley, G., & Shneiderman, B. (1998). Engagement theory: A framework for technology-based teaching and learning. *Educational Technology*, *38*(5), 20–23. https://www.jstor.org/stable/44428478

Leahy, K., Ozer, E., & Cummins, E. P. (2025). AI-ENGAGE: A multicentre intervention to support teaching and learning engagement with generative artificial intelligence tools. *Education Sciences*, *15*(7), Article 807. https://doi.org/10.3390/educsci15070807

Liu, J., & Yang, Y. (2025). A student performance prediction model based on multimodal generative adversarial networks. *International Journal of Sensor Networks*, *47*(3), 186–198. https://doi.org/10.1504/IJSNET.2025.144630

Long, P., & Siemens, G. (2014). Penetrating the fog: analytics in learning and education. *Italian Journal of Educational Technology*, *22*(3), 132-137. https://www.learntechlib.org/d/183382/

Lukumon, G., Maharaj, A., Kafilat, S., Lateefat, A., Sunday, S., Musibau, G. A., ... Saheed, O. (2025). From phobia to fun: Enhancing mathematics engagement with AI-powered tools. *OSF*. *7*, 24-38. https://doi.org/10.31219/osf.io/jkt2p

Medina-Gual, L., & Parejo, J. L. (2025). University students' engagement with artificial intelligence: A cluster analysis of learner profiles in AI literacy. *Technology, Knowledge and Learning*, *9*(2), 1–19. https://doi.org/10.1007/s10758-025-09926-7

Møgelvang, A., Bjelland, C., Grassini, S., & Ludvigsen, K. (2024). Gender differences in the use of generative artificial intelligence chatbots in higher education: Characteristics and consequences. *Education Sciences*, *14*(12), Article 1363. https://www.mdpi.com/2227-7102/14/12/1363

Qu, Y., Tan, M. X. Y., & Wang, J. (2024). Disciplinary differences in undergraduate students' engagement with generative artificial intelligence. *Smart Learning Environments*, *11*(1), Article 51. https://doi.org/10.1186/s40561-024-00341-6

Sayici, S. (2025). *Artificial intelligence in student learning: Balancing usefulness, stress, and cognitive load* [Master thesis, University of Tilburg]. https://arno.uvt.nl/show.cgi?fid=185574

Siemens, G., & Gašević, D. (2012). Guest editorial-Learning and knowledge analytics. *Journal of Educational Technology & Society*, *15*(3), 1–2. https://www.researchgate.net/publication/279555980_Guest_Editorial_-_Learning_and_Knowledge_Analytics

Tang, C., Li, S., Hu, S., Zeng, F., & Du, Q. (2025). Gender disparities in the impact of generative artificial intelligence: Evidence from academia. *PNAS Nexus*, *4*(2), Article 591. https://doi.org/10.1093/pnasnexus/pgae591

Tomita, K. (2018). Does the visual appeal of instructional media affect learners' motivation toward learning? *TechTrends*, *62*(1), 103–112. https://link.springer.com/article/10.1007/s11528-017-0213-1

Wicks, D. A., Craft, B. B., Mason, G. N., Gritter, K., & Bolding, K. (2015). An investigation into the community of inquiry of blended classrooms by a faculty learning community. *The Internet and Higher Education*, *25*, 53–62. https://doi.org/10.1016/j.iheduc.2014.12.001

Wise, A. F., & Jung, Y. (2019). Teaching with analytics: Towards a situated model of instructional decision-making. *Journal of Learning Analytics*, *6*(2), 53–69. https://doi.org/10.18608/jla.2019.62.4

Woolf, B. P., Lane, H. C., Chaudhri, V. K., & Kolodner, J. L. (2013). AI grand challenges for education. *AI Magazine*, *34*(4), 66–84. https://doi.org/10.1609/aimag.v34i4.2490

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education*, *16*, Article 39. https://doi.org/10.1186/s41239-019-0171-0