

An Investigation of the Use of Generative AI Chatbots Among English Major Students at Kabul University

Morshedullah Abdul Noorzai^{*1} & Abdullah Noori² 

Kabul University, English Language and Literature Department, Kabul, Afghanistan

ABSTRACT

Generative AI (GenAI) chatbots are increasingly used in higher education, yet little is known about how English major students, particularly in Afghanistan, perceive and use them. This study examines perceptions, perceived benefits, concerns, and the frequency of GenAI chatbot use among English major students at Kabul University, and tests whether use frequency relates to academic year. The study used a descriptive quantitative survey design, with 113 closed-ended questionnaires distributed via convenience sampling. Data were analyzed in SPSS using descriptive statistics, and a one-way ANOVA with Tukey HSD post hoc tests compared frequency of use across academic years. Results showed students held positive but cautious attitudes toward chatbots ($M = 3.17$, $SD = 0.82$) and perceived the greatest benefits for speaking, writing, and listening skills. Students used chatbots mainly for everyday academic tasks rather than high-stakes work such as exams ($M = 3.36$, $SD = 0.64$), and they expressed notable concern about the accuracy and reliability of chatbot-generated content. Third-year students used chatbots significantly more frequently than first- and second-year students. These findings suggest that university policymakers, teachers, and students should promote responsible, guided chatbot use, particularly by addressing accuracy concerns and supporting engagement across academic years.

ARTICLE HISTORY

Revised 15 Aug 2026
Accepted 27 Aug 2026

KEYWORDS

Concerns; English major students; Frequency of use; Generative AI chatbots; Perceptions; Perceived benefits

Cite as: Abdul Noorzai, M. & Noori, A. (2026). An investigation of the use of generative AI chatbots among English major students at Kabul University. *Aureviaj Journal of Social Sciences and Humanities*. 1(1), 39-16. <https://aureviaj.com/ajssh/article/view/18>



This work is licensed under a Creative Commons Attribution-Non-Commercial 4.0 International License.

INTRODUCTION

Generative artificial intelligence (GenAI) chatbots, because they can generate human-like text, answer questions, translate languages, and summarize long passages, have rapidly become one of the most researched technologies in education (Khurma et al., 2024). Since ChatGPT and similar tools were introduced to the public in late 2022, their user base has reached nearly one billion worldwide, and the global chatbot market is projected to grow exponentially by 2029 (Kujundziski et al., 2025). A large-scale meta-analysis of 35 experimental studies with over 4,000 participants found that ChatGPT has a moderately positive effect on students' overall learning experience (Wu et al., 2026), and across educational settings, these tools are consistently linked to motivation, engagement, and skill development (Khurma et al., 2024; Kujundziski et al., 2025; Wu et al., 2026). Built on large language models, chatbots such as

*CONTACT ¹Morshedullah Abdul Noorzai morshedullah.abdul.noorzai@gmail.com Department of English Language and Literature, Kabul University, Kabul, Afghanistan

ChatGPT, Co-pilot, Gemini, and Claude function as flexible, on-demand assistants rather than the fixed-response educational software of earlier decades, offering learners personalized help and instant feedback that lets them evaluate their own understanding without needing a teacher's constant intervention (Badger et al., 2026; Elnaffar et al., 2025; Mittal et al., 2024; Ng et al., 2025; Noroozi et al., 2024).

For English major students, these tools support core language tasks, including writing, translation, and conversation practice (Mittal et al., 2024; Ng et al., 2025; Noroozi et al., 2024). Chatbots can act as writing assistants that help students overcome writing barriers (Ng et al., 2025), provide tailored explanations of mistakes, and generate interactive, conversation-based exercises for skill practice (Mittal et al., 2024), while simulations and role-plays delivered through chatbots have been linked to gains in motivation, creativity, and critical thinking (Michel-Villarreal et al., 2023). In English Language Teaching specifically, chatbots are valued for creating a low-anxiety space where students can interact with language material at their own pace, instead of facing the pressure of face-to-face conversation with an instructor (Peña-Acuña & Corga Fernandes Durão, 2024; LIMIYA & Kumar, 2025; Zaim et al., 2025). This support extends to speaking, fluency, pronunciation, and vocabulary, although the benefit depends more on how attentively students engage with a chatbot than on how frequently they use it (Al-Khresheh et al., 2026; Lee et al., 2025; Lopez & Arabadzhy, 2026).

Despite these documented benefits, researchers consistently raise two related concerns. The first is overreliance: students who lean too heavily on chatbots report weakened independent thinking, superficial understanding of language structure, reduced face-to-face interaction, and diminished research skills (Athanasopoulos et al., 2026; Elnaffar et al., 2025; LIMIYA & Kumar, 2025; Zaim et al., 2025; Zhong & Rosli, 2025). The second is academic integrity, since chatbots make it easy for students to submit AI-generated text without fully comprehending it (Athanasopoulos et al., 2026; Elnaffar et al., 2025). A further, more field-specific concern is linguistic: because chatbots are trained predominantly on standard English, they may limit learners' exposure to the varieties of English used globally and marginalize non-standard forms (Lee et al., 2025; LIMIYA & Kumar, 2025; Zaim et al., 2025). Taken together, scholars argue that chatbots support learning only when used in a guided, purposeful, and critically evaluated way (Al-Khresheh et al., 2026; Lee et al., 2025; Zhong & Rosli, 2025).

The pace of GenAI adoption varies widely across contexts, and developing countries face distinct structural barriers (Simiyu et al., 2026). Unequal access to electricity, stable internet, and affordable devices means AI-driven tools risk widening rather than closing existing educational inequalities (Landa-Blanco, 2026; Liu et al., 2025; Nguyen & Perkins, 2026); Liu et al. (2025), for instance, report that ChatGPT usage among internet users falls from 24% in high-income countries to under 1% in low-income countries. Institutions in these contexts also lack the trained personnel, clear data-privacy policies, and content-detection tools available elsewhere (Cai, 2025; Landa-Blanco, 2026; Nguyen & Perkins, 2026), and because chatbots are built on majority languages and Western content, their responses may carry cultural bias that further marginalizes local and minority varieties of English (Landa-Blanco, 2026; Lee et al., 2025; Nguyen & Perkins, 2026).

These same barriers shape GenAI adoption in Afghan higher education. Afghan students and researchers report overreliance that weakens critical thinking and raises academic dishonesty concerns, alongside doubts about AI accuracy, poor integration with existing academic systems, a lack of formal training, weak infrastructure, unclear ethical guidance, unstable internet, unreliable electricity, and skepticism toward AI outside Kabul (Ahmadi & Danishyar, 2025; Sadat, 2025; Sadat et al., 2025; Zgham et al., 2025). Fear of plagiarism, loss of academic integrity, and concern that AI might eventually replace human teachers remain widespread nationally (Sadat, 2025).

Despite this growing body of research, most existing studies either survey students across multiple fields without focusing on English majors specifically, or examine populations such as STEM students, general education cohorts, or mixed-discipline samples (Al Mashagbeh et al., 2025; Almassaad et al., 2024; He et al., 2024; Johnston et al., 2024; Nelson et al., 2025; Saleh & ElSayary, 2026; Wang & Deng, 2026). Where perceptions, benefits, and concerns are studied, they are typically treated as secondary outcomes within broader technology-acceptance models, and only a handful of studies examine the relationship between academic year and frequency of chatbot use (Al-Khresheh et al., 2026; Badger et al., 2026; He et al., 2024; Mansoor et al., 2026; Wang & Deng, 2026). Within Afghanistan specifically, existing research has focused on general barriers to GenAI adoption in higher education, such as infrastructure limitations, low digital literacy, and ethical concerns, without examining English major students' perceptions, perceived benefits, frequency of use, or concerns in a focused, discipline-specific way (Ahmadi & Danishyar, 2025; Sadat, 2025; Zgham et al., 2025). Because English major students rely heavily on language-based tasks such as writing, translation, and conversation practice, their relationship with GenAI chatbots may differ from that of students in other disciplines. However, this population remains underexamined at Kabul University, leaving administrators, teachers, and policymakers without the evidence needed to design informed guidelines for responsible GenAI integration in English language programs.

To address this gap, this descriptive quantitative study examines English major students' perceptions of GenAI chatbots at Kabul University, their perceived benefits, frequency of use, and concerns, and tests whether frequency of use is related to academic year. Specifically, the study pursues five objectives: to explore English major students' perceptions toward GenAI chatbots; to identify their perceived benefits; to determine their frequency of use; to identify their concerns; and to examine whether a relationship exists between academic year and frequency of use. By collecting first-hand data from this specific and previously underexamined population, the study contributes field-specific evidence that responds directly to calls in the literature for research into how GenAI use varies across academic disciplines and levels of study (Al Mashagbeh et al., 2025; Badger et al., 2026; Mansoor et al., 2026), and provides the English Language and Literature Department of Kabul University, its teachers, and policy makers with evidence to guide the responsible integration of GenAI chatbots into English language teaching and learning.

RESEARCH METHODOLOGY

This study followed a descriptive quantitative survey design. According to Creswell (2019), quantitative research describes trends and explains relationships among variables by using clear, focused research questions, appropriate instruments for collecting numerical data, and statistical analysis interpreted against prior research; survey research specifically involves identifying a population and sample, collecting data through questionnaires or interviews, and drawing conclusions about the population from the sample (Creswell, 2019). This design was chosen to measure students' perceptions, perceived benefits, frequency of use, and concerns regarding GenAI chatbots, and to test whether a relationship exists between academic year and frequency of use.

The population comprised 148 students from the Department of English Language and Literature at Kabul University. Based on the Krejcie and Morgan table for a finite population of 150, a minimum sample of 108 was required; to guard against faulty or incomplete responses, 113 questionnaires were distributed, and all were completed and usable. Participants were selected through convenience sampling and represented all four academic years, and therefore varied in their experience with GenAI chatbots and English proficiency level. All participants were informed of the study's purpose, participated voluntarily, and were actively enrolled in the Department at the time of data collection.

Data were collected using a five-part closed-ended questionnaire adapted from Xuyen (2024) and Serhan and Welcome (2026): demographics, perceptions (4 items), perceived benefits (12 items), concerns (6 items), and frequency of use (7 items). All items aligned with the study's research questions and were reviewed for clarity and comprehensibility before distribution. Following departmental approval, the researcher, a senior student in the Department, administered the questionnaires in person over one week, explained unclear items as needed, and assured participants of confidentiality and their right to withdraw at any time; this yielded a 100% completion rate with no faulty responses.

Data were analyzed in SPSS (2024) using descriptive statistics (frequencies, percentages, means, and standard deviations) to summarize demographic and questionnaire responses, and a one-way analysis of variance (ANOVA) with Tukey HSD post hoc tests to examine whether GenAI chatbot use frequency differed significantly across academic years. All questionnaires were numerically coded prior to analysis to protect respondent confidentiality.

RESULTS

A total of 113 English major students participated in the study. Table 1 below presents the sample's demographic profile by age, academic year, and English proficiency level. Most respondents (66.4%) were aged 21–23 years, and fourth-year students made up the largest year group (38.1%), followed by third-year (24.8%), first-year (19.5%), and second-year students (17.7%). In terms of self-reported proficiency, 69.9% identified as intermediate, 24.8% as advanced, and 5.3% as pre-intermediate.

Table 1. *Demographic Characteristics of the Respondents*

Variable	Category	Frequency	Percent
Age	18–20 Years	11	9.7
	21–23 Years	75	66.4
	24–26 Years	26	23.0
	27 years or above	1	0.9
	Total	113	100.0
Year of Study	First Year	22	19.5
	Second Year	20	17.7
	Third Year	28	24.8
	Fourth Year	43	38.1
	Total	113	100.0
English Proficiency	Pre-Intermediate	6	5.3
	Intermediate	79	69.9
	Advanced	28	24.8
	Total	113	100.0

Perceptions, Benefits, Concerns, and Frequency of Use

Table 2 presents respondents' general perceptions toward GenAI chatbots. Respondents reported feeling comfortable using chatbots in class ($M = 3.35$, $SD = 0.96$) and enjoying their use ($M = 3.24$, $SD = 1.03$), with more moderate agreement that they would like to use chatbots in all classes ($M = 3.12$, $SD = 0.99$). Respondents rated the belief that chatbots would positively affect their overall university learning experience lowest ($M = 2.98$, $SD = 1.10$; overall $M = 3.17$, $SD = 0.82$).

Table 2. *Respondents' Perceptions of the Use of GenAI Chatbots*

Item	M	SD
I feel comfortable using GenAI chatbots in the class.	3.35	0.96
I enjoy using GenAI chatbots in class.	3.24	1.03
I would like to use GenAI chatbots in all classes.	3.12	0.99
I think using GenAI chatbots will positively affect my university learning experience.	2.98	1.10
Overall	3.17	0.82

Table 3 presents perceived benefits across twelve statements. Respondents rated improvement in English-speaking skills highest ($M = 3.95$, $SD = 1.23$), followed by writing ($M = 3.91$, $SD = 1.39$) and listening ($M = 3.87$, $SD = 1.36$); immediate feedback ($M = 3.30$, $SD = 1.37$) and

unlimited practice opportunities ($M = 3.27$, $SD = 1.47$) were rated lowest, though still above the scale midpoint (overall $M = 3.65$, $SD = 0.76$).

Table 3. *Respondents' Perceived Benefits of GenAI Chatbots*

Item	M	SD
GenAI chatbots help me improve my English-speaking skills.	3.95	1.23
GenAI chatbots help me improve my English-writing skills.	3.91	1.39
GenAI chatbots help me improve my English-listening skills.	3.87	1.36
GenAI chatbots provide immediate feedback and assessment, enabling me to adjust my learning right away.	3.30	1.37
GenAI chatbots provide me with unlimited opportunities to practice the target language.	3.27	1.47
Overall (12 items)	3.65	0.76

Table 4 summarizes respondents' concerns. The greatest concern was the reliability and accuracy of chatbot-generated information ($M = 3.83$, $SD = 1.30$), followed by reduced human interaction ($M = 3.78$, $SD = 1.40$) and bias in AI responses ($M = 3.62$, $SD = 1.39$); concerns about overdependence and its effect on critical thinking were rated lowest but remained above the scale midpoint ($M = 3.44$ – 3.50 ; overall $M = 3.62$, $SD = 0.81$).

Table 4. *Respondents' Concerns about GenAI Chatbots*

Item	M	SD
I am worried about the reliability and accuracy of the information provided by GenAI chatbots.	3.83	1.30
I am worried that overusing GenAI chatbots can lead to lack of human interaction.	3.78	1.40
I am worried that the information provided by GenAI tools such as ChatGPT can be biased.	3.62	1.39
I am worried that I will become overdependent on GenAI tools in my English language learning.	3.50	1.38
I am worried that my overdependence on GenAI chatbots can negatively affect my critical thinking skills.	3.44	1.43
Overall (6 items)	3.62	0.81

Table 5 below presents the frequency of use across seven statements. Respondents reported the highest frequency for general academic studies ($M = 3.58$, $SD = 1.05$), followed by completing English-language assignments ($M = 3.48$, $SD = 0.96$) and improving writing skills ($M = 3.40$, $SD = 0.83$); the lowest frequencies were for exam or presentation preparation ($M = 3.22$, $SD = 1.04$) and direct reliance on AI-generated answers ($M = 3.20$, $SD = 1.06$; overall $M = 3.36$, $SD = 0.64$).

Table 5. *Respondents' Frequency of Use of GenAI Chatbots*

Item	M	SD
How often do you use GenAI tools (e.g., ChatGPT) for your academic studies?	3.58	1.05
How often do you use GenAI chatbots to complete English language assignments?	3.48	0.96
How often do you use GenAI chatbots to improve your writing skills (e.g., paragraphs or essays)?	3.40	0.83
How often do you use GenAI chatbots to prepare for examinations, quizzes, or presentations?	3.22	1.04
How often do you rely on AI-generated responses when completing academic tasks?	3.20	1.06
Overall (7 items)	3.36	0.64

Relationship between Academic Year and GenAI Frequency of Use

A one-way ANOVA tested whether chatbot use frequency differed by academic year Table 6. The results indicate that Levene's test indicated that homogeneity of variance was approached but not clearly established, $F(3, 109) = 2.662$, $p = .052$ (Table 7), and the ANOVA revealed a statistically significant difference across academic years, $F(3, 109) = 6.136$, $p = .001$.

Table 6. *One-Way ANOVA for the Frequency of GenAI Chatbot Use*

Source	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	6.70	3	2.23	6.13	.001
Within Groups	39.70	109	0.36		
Total	46.40	112			

Tukey HSD post hoc comparisons (Table 7) showed third-year students used chatbots significantly more often than first-year students (mean difference = 0.67, $p = .01$), with a difference approaching significance for second-year students (mean difference = 0.57, $p = .09$); no significant difference was found between third- and fourth-year students (mean difference = 0.35, $p = .07$) or between first- and fourth-year students (mean difference = -0.31 , $p = .19$).

Table 7. *Post Hoc Tukey HSD Multiple Comparisons by Academic Year*

Year (I)	Year (J)	Mean Diff. (I–J)	SE	Sig.	95% CI
First Year	Second Year	–0.10	0.18	.94	[–0.59, 0.38]
	Third Year	–0.67*	0.17	.01	[–1.12, –0.22]
	Fourth Year	–0.31	0.15	.19	[–0.72, 0.09]
Second Year	First Year	0.10	0.18	.94	[–0.38, 0.59]
	Third Year	–0.57*	0.17	.09	[–1.03, –0.10]
	Fourth Year	–0.21	0.16	.57	[–0.63, 0.21]
Third Year	First Year	0.67*	0.17	.01	[0.22, 1.12]

Year (I)	Year (J)	Mean Diff. (I–J)	SE	Sig.	95% CI
Fourth Year	Second Year	0.57*	0.17	.09	[0.10, 1.03]
	Fourth Year	0.35	0.14	.07	[−0.02, 0.74]
	First Year	0.31	0.15	.19	[−0.09, 0.72]
	Second Year	0.21	0.16	.57	[−0.21, 0.63]
	Third Year	−0.35	0.14	.07	[−0.74, 0.02]

Note. The mean difference is significant at the 0.05 level (indicated by *).

The homogeneous-subsets analysis (Table 8) placed first-, second-, and fourth-year students in one overlapping subset (Sig. = .238) and fourth- and third-year students in a second (Sig. = .146), with third-year students occupying the highest subset alone.

Table 10. Homogeneous Subsets for Frequency of Use by Academic Year (Tukey HSD)

Academic Year	N	Subset 1	Subset 2
First Year	22	3.05	—
Second Year	20	3.16	—
Fourth Year	43	3.37	3.37
Third Year	28	—	3.73
Sig.		.238	.146

Note. Uses harmonic mean sample size = 25.90; group sizes are unequal.

DISCUSSION

The study's primary hypothesis—that academic year is related to frequency of GenAI chatbot use—was supported: the one-way ANOVA showed a statistically significant difference across academic years, with third-year students using chatbots more frequently than first-year students. This rise in use from the first to the third year, followed by a slight decline in the fourth year, was not fully anticipated; a plausible post hoc explanation is that third-year students face a heavier load of research-based and writing-intensive coursework than first- and second-year students, prompting greater reliance on chatbots for academic support, while fourth-year students, closer to graduation and often occupied with practicum, thesis supervision, or job preparation, may substitute other resources or rely on skills already developed in earlier years.

The finding that students hold comfortable but cautious perceptions toward chatbots, rating their effect on overall learning experience lowest among the four perception items, confirms Wu et al.'s (2026) meta-analytic conclusion that ChatGPT has only a moderate effect on learning outcomes, and refines it by showing that this caution centers specifically on doubts about learning impact rather than on comfort or enjoyment, which were rated more favorably. The pattern of perceived benefits, concentrated in speaking, writing, and listening skills, confirms Mittal et al. (2024) and Ng et al. (2025), who describe chatbots as writing assistants that help students overcome language barriers, while extending this literature by showing that English majors in this sample valued chatbots most for productive and receptive skill practice

rather than for feedback speed or unlimited practice, which were rated comparatively lower. The concerns findings confirm Athanassopoulos et al. (2026), Elnaffar et al. (2025), and Zhong and Rosli (2025) on overreliance and academic-integrity risk, and align with Afghan-specific findings from Ahmadi and Danishyar (2025) and Sadat (2025) on accuracy doubts, but refine this literature by identifying accuracy and reliability, rather than overdependence itself, as the single greatest concern among this population. The relationship between academic year and frequency of use confirms He et al. (2024), who found that ChatGPT use frequency differs by field and level of study, and extends this literature into a population and analytic approach, one-way ANOVA with Tukey HSD post hoc comparisons, that prior studies such as Almassaad et al. (2024) and Wang and Deng (2026) did not apply.

Several methodological parameters warrant consideration when interpreting these results. Because all data were collected through a self-report questionnaire, responses are subject to social desirability bias, particularly for the concerns and perceptions items, where students may have understated overreliance or overstated caution. The use of convenience sampling, while yielding a full response rate, also introduces the possibility of self-selection bias, as students more engaged with or opinionated about GenAI chatbots may have been more willing to participate. On statistical grounds, Levene's test approached but did not clearly fall below the conventional significance threshold ($p = .052$), indicating variances across academic-year groups were only marginally homogeneous; this was accounted for by reporting Tukey HSD comparisons, which are relatively robust to minor variance heterogeneity, but the closeness of this result to the threshold warrants some caution in generalizing the ANOVA findings. In addition, several pairwise comparisons in the post hoc analysis, including the differences between third- and second-year students ($p = .09$) and between third- and fourth-year students ($p = .07$), approached but did not reach conventional significance, suggesting the true pattern of change across years may be more gradual than the significant overall F-test alone implies.

The target population was the full population of 148 English major students at Kabul University, and the achieved sample of 113 exceeded the minimum required by the Krejcie and Morgan table, supporting reasonable representativeness within this population. However, generalizability beyond this context is limited in several respects. The sample was drawn from a single department at a single university, and Sadat (2025) and Zgham et al. (2025) both note that skepticism toward AI and infrastructure access vary considerably between Kabul and other Afghan provinces, meaning students at provincial universities may report different patterns of use or concern. The sample was also unevenly distributed across academic years, with fourth-year students comprising more than a third of respondents and second-year students the smallest group, which may have influenced the precision of the year-by-year comparisons. Because data collection occurred at a single point in time, the findings describe a cross-sectional snapshot rather than a developmental trend, and the apparent rise and slight decline in use across academic years should be interpreted as a between-group difference rather than confirmed evidence of within-student change over time.

Taken together, these findings address the gap identified in the Introduction. While prior research on GenAI in Afghan higher education has examined general adoption barriers without a discipline-specific focus, and research on English major students elsewhere has rarely tested

academic year as a predictor of frequency, this study provides field-specific evidence on both fronts for Kabul University. The results validate the broader proposition, established in general higher-education literature, that GenAI chatbot engagement is neither uniform nor static across a student population, and they challenge any assumption that concerns about GenAI use in Afghanistan are limited to infrastructure and access; even among students situated at the country's largest and best-resourced university, concerns about the accuracy and reliability of chatbot output remain prominent. An unresolved question that follows directly from these findings is why third-year use peaks and fourth-year use recedes. Whether this reflects a genuine developmental pattern in how students integrate chatbots into their academic work, a cohort-specific difference unrelated to year of study, or task-specific demands unique to the third-year curriculum is a question this cross-sectional design cannot resolve, and that future longitudinal research should address.

CONCLUSION

This descriptive quantitative study examined the perceptions, perceived benefits, frequency of use, and concerns of 113 English major students at Kabul University toward GenAI chatbots, and tested whether academic year relates to frequency of use. Students held a comfortable yet cautious attitude toward chatbots, perceived the greatest benefits for speaking, writing, and listening skills, used chatbots more for everyday academic tasks than for high-stakes work such as examinations, and expressed strong concern about the accuracy and reliability of chatbot-generated content. Academic year was significantly related to frequency of use, with third-year students engaging most frequently.

The central contribution of this study is field-specific evidence, previously unavailable, on how English major students at Kabul University perceive, use, and evaluate GenAI chatbots, and on how that engagement varies across academic years. This evidence can inform more targeted guidance from teachers and policy makers, particularly around accuracy and academic integrity, than general adoption research from other disciplines or contexts could provide.

The study is bounded by its use of convenience sampling within a single department at one Afghan university, an unevenly distributed sample across academic years, reliance on self-reported data, and a cross-sectional design that captures only a single point in time; these boundaries limit how far the findings can be generalized and mean the year-by-year differences observed should be read as a snapshot rather than confirmed developmental change. Future research should use qualitative or mixed-methods designs to explore students' concerns and usage patterns in greater depth, conduct comparative studies across multiple Afghan universities and disciplines, and use longitudinal designs that track individual students' chatbot use across a full course of study.

REFERENCES

- Ahmadi, R., & Danishyar, R. (2025). The impact of artificial intelligence on the educational activities of Bamyan University students. *Journal of Effective Teaching Methods*, 3(4), 112–124. <https://doi.org/10.59652/9vf86350>

- Al Mashagbeh, M., Alsharqawi, M., Tudevdagva, U., & Khasawneh, H. J. (2025). Student engagement with artificial intelligence tools in academia: A survey of Jordanian universities. *Frontiers in Education*, 10, Article 1550147, 1–12.
<https://doi.org/10.3389/educ.2025.1550147>
- Al-Khresheh, M. H., Almayez, M., & Alruwaili, S. F. (2026). Working memory, attention control, and vocabulary retention in AI (ChatGPT)-assisted foreign language learning: A structural cognitive modelling approach. *Journal of Intelligence*, 14(4), Article 62, 1–31. <https://doi.org/10.3390/jintelligence14040062>
- Almassaad, A., Alajlan, H., & Alebaikan, R. (2024). Student perceptions of generative artificial intelligence: Investigating utilization, benefits, and challenges in higher education. *Systems*, 12(10), Article 385, 1–16.
<https://doi.org/10.3390/systems12100385>
- Athanassopoulos, S., Tzavara, A., Aravantinos, S., Lavidas, K., Komis, V., & Papadakis, S. (2026). Teacher education students' practices, benefits, and challenges in the use of generative AI tools in higher education. *Education Sciences*, 16(2), 228.
<https://doi.org/10.3390/educsci16020228>
- Badger, M., López, N., & Nissen, C. F. R. (2026). The impact of GenAI chatbots on student learning in higher education: A literature review. *International Journal of Technology in Education (IJTE)*, 9(1), 43–69. <https://doi.org/10.46328/ijte.5111>
- Cai, Y. (2025). Navigating the impact of generative AI in English language education. *Forum for Education Studies*, 3(3), 2143, 1–7. <https://doi.org/10.59400/fes2143>
- Creswell, J. W., & Guetterman, T. C. (2019). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (6th ed.). Pearson.
- Elnaffar, S., Rashidi, F., & Abualkashik, A. Z. (2026). Teaching with AI: A systematic review of chatbots, generative tools, and tutoring systems in programming education. *International Journal of Learning, Teaching and Educational Research*, 25(1), 1–28.
<https://doi.org/10.26803/ijlter.25.1.1>
- He, A. J., Zhang, Z., Anand, P., & McMinn, S. (2024). Embracing generative artificial intelligence tools in higher education: A survey study at the Hong Kong University of Science and Technology. *Journal of Asian Public Policy*. Advance online publication, 1–26. <https://doi.org/10.1080/17516234.2024.2447195>
- Johnston, H., Wells, R. F., Shanks, E. M., Boey, T., & Parsons, B. N. (2024). Student perspectives on the use of generative artificial intelligence technologies in higher education. *International Journal for Educational Integrity*, 20, Article 2, 1–21.
<https://doi.org/10.1007/s40979-024-00149-4>
- Khurma, O. A., Albahti, F., Ali, N., & Bustanji, A. (2024). AI ChatGPT and student engagement: Unraveling dimensions through PRISMA analysis for enhanced learning experiences. *Contemporary Educational Technology*, 16(2), 1–22.
<https://doi.org/10.30935/cedtech/14334>

- Kujundziski, A. P., Ademi, N., & Rahmani, D. (2025). AI chatbots in education – the future of teaching and learning. *E-mentor*, 4(111), 4–12.
<https://www.doi.org/10.15219/em111.1724>
- Landa-Blanco, M. (2026). Artificial intelligence in education: Applications and limitations for teachers in low- and middle-income countries. *Frontiers in Education*, 10, 1681836, 1–6. <https://doi.org/10.3389/feduc.2025.1681836>
- Lee, S., Choe, H., Zou, D., & Jeon, J. (2025a). Generative AI (GenAI) in the language classroom: A systematic review. *Interactive Learning Environments*, 34(1), 335–359.
<https://doi.org/10.1080/10494820.2025.2498537>
- Lee, S., Jeon, J., McKinley, J., & Rose, H. (2025b). Generative AI and English language teaching: A global Englishes perspective. *Annual Review of Applied Linguistics*, 45, 85–108. <https://doi.org/10.1017/S0267190525100184>
- Limiya, A. V. G., & Kumar, P. N. S. (2025). Generative artificial intelligence in enhancing English language skills: A systematic review. *Studies in Media and Communication*, 13(4), 256–267. <https://doi.org/10.11114/smc.v13i4.7763>
- Liu, Y., Huang, J., & Wang, H. (2025). *Who on earth is using generative AI? Global trends and shifts in 2025* (Policy Research Working Paper No. 11231). World Bank Digital Transformation Vertical. <https://reproducibility.worldbank.org/index.php/catalog/332>
- Lopez, L., & Arabadzhy, G. (2026). GenAI chatbots in language learning: A systematic review of emerging affordances. *Language Learning & Technology*, 30(1), 1–25.
<https://doi.org/10.64152/10125/73700>
- Mansoor, H. S., Sumardjoko, B., & Sutopo, A. (2026). External variables influencing the attitudes of students toward AI acceptance in improving English writing: A systematic review. *Frontiers in Artificial Intelligence*, 8, Article 1719955, 1–16.
<https://doi.org/10.3389/frai.2025.1719955>
- Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Education Sciences*, 13(9), 856, 1–18.
<https://doi.org/10.3390/educsci13090856>
- Mittal, U., Sai, S., Chamola, V., & Sangwan, D. (2024). A comprehensive review on generative AI for education. *IEEE Access*, 12, 142733–142759.
<https://doi.org/10.1109/ACCESS.2024.3468368>
- Nelson, A. S., Santamaría, P. V., Javens, J. S., & Ricaurte, M. (2025). Students' perceptions of generative artificial intelligence (GenAI) use in academic writing in English as a foreign language. *Education Sciences*, 15(5), Article 611, 1–31.
<https://doi.org/10.3390/educsci15050611>
- Ng, D. T. K., Chan, E. K. C., & Lo, C. K. (2025). Opportunities, challenges and school strategies for integrating generative AI in education. *Computers and Education: Artificial Intelligence*, 8, 100373, 1–12. <https://doi.org/10.1016/j.caeai.2025.100373>

- Nguyen, A. T., & Perkins, M. (2026). Perceptions of generative AI in the Global South: A scoping review. *Journal of University Teaching and Learning Practice*. Advance online publication. <https://doi.org/10.53761/f22j6648>
- Noroozi, O., Soleimani, S., Farrokhnia, M., & Banihashem, S. K. (2024). Generative AI in education: Pedagogical, theoretical, and methodological perspectives. *International Journal of Technology in Education*, 7(3), 373–385. <https://doi.org/10.46328/ijte.845>
- Peña-Acuña, B., & Corga Fernandes Durão, R. (2024). Learning English as a second language with artificial intelligence for prospective teachers: A systematic review. *Frontiers in Education*, 9, 1490067, 1–7. <https://doi.org/10.3389/educ.2024.1490067>
- Sadat, M. M. N. (2025). E-learning and artificial intelligence: A solution to the limitations of educational infrastructure in Afghanistan. *International Journal of Educational Research*, 2(3), 17–24. <https://doi.org/10.62951/ijer.v2i3.394>
- Sadat, M. M. N., Tokhi, M., Azizi, M. A., Esmati, S., & Emynorane, R. H. (2025). Evaluating the impact of AI-enabled cloud communication platforms on student engagement in higher education: A case of Afghanistan private universities. *Global Education Trends*, 3(2), 113–124.
- Saleh, D., & ElSayary, A. (2026). Generative AI use and self-learning in higher education: The role of learning difficulties. *Frontiers in Education*, 10, Article 1737928, 1–15. <https://doi.org/10.3389/educ.2025.1737928>
- Serhan, D., & Welcome, N. (2026). Students & AI tools: Attitudes & perceptions. *International Journal of Research in Education and Science (IJRES)*, 12(2), 332–345. <https://doi.org/10.46328/ijres.6888>
- Simiyu, C. N., Malanga, K. N., & Wabwoba, F. (2026). Bridging the AI chasm: A strategic framework for generative AI in higher education. *Sustainable and Pedagogical Technologies*, 2(1), 1–20. <https://doi.org/10.33902/SPT.202642185>
- Wang, L., & Deng, P. (2026). Determinants of university students' intention to use generative AI tools for personalized English learning: Mediating effect of flow experience and moderating effect of personal innovativeness. *Frontiers in Psychology*, 17, Article 1728820, 1–17. <https://doi.org/10.3389/fpsyg.2026.1728820>
- Wu, X., Zhu, P., Zhang, J., Yin, M., & Wang, Y. (2026). ChatGPT's impact on student learning outcomes: A meta-analysis of 35 experimental studies. *Humanities and Social Sciences Communications*, 1–17. <https://doi.org/10.1057/s41599-026-07019-z>
- Xuyen, N. T. (2024). English majors' perceptions of AI tool application in English language learning at tertiary level in Vietnam. *Journal of Knowledge Learning and Science Technology*, 3(1), 179–193. <https://doi.org/10.60087/jklst.vol3.n1.p193>
- Zaim, M., Arsyad, S., Waluyo, B., Ardi, H., Al Hafizh, M., Zakiyah, M., Syafitri, W., Nusi, A., & Hardiah, M. (2025). Generative AI as a cognitive copilot in English language learning in higher education. *Education Sciences*, 15(6), 686, 1–29. <https://doi.org/10.3390/educsci15060686>

- Zgham, W., Danishmal, M., & Haidary, M. H. (2025). Artificial intelligence in research: Benefits, challenges, and strategies for implementing AI tools in research in Afghanistan. *Artificial Intelligence and Applications*, Advance online publication, 1–10. <https://doi.org/10.47852/bonviewAIA52024855>
- Zhong, Y., & Rosli, M. S. B. (2025). Generative artificial intelligence in higher education: Opportunities, challenges, and future directions. *International Journal of Academic Research in Progressive Education and Development*, 14(3), 1–20. <https://doi.org/10.6007/IJARPED/v14-i3/25813>