INTERNATIONAL SCIENTIFIC CONFERENCE
19-21 November 2026
GABROVO

GENERATIVE ARTIFICIAL INTELLIGENCE IN HIGHER EDUCATION: A NARRATIVE REVIEW ACROSS STUDENTS, FACULTY, AND TEACHING–LEARNING PROCESSES

Eylem Seyhan[1*], Berfin Kılınç[1], Deniz Mertkan Gezgin[2]

[1]Trakya University, Department of Computer Engineering, Edirne, Türkiye
[2]Trakya University, Department of Instructional Technologies, Edirne, Türkiye
**Corresponding author: eylemseyhan@trakya.edu.tr*

Abstract
Generative artificial intelligence (GAI) is becoming increasingly embedded in higher education, yet its educational value remains difficult to interpret because existing evidence is dispersed across student use, faculty adoption, and teaching–learning practices. This narrative review brings these three perspectives together by examining 20 peer-reviewed publications from 2023 to 2026. Rather than treating GAI use as a single educational phenomenon, the review considers how it is used, what benefits are reported, and what limitations emerge across different actors and contexts. Students primarily use GAI for information seeking, idea generation, academic writing, and problem solving, whereas faculty adoption is associated more strongly with perceived usefulness, trust, and institutional support than with ease of use alone. At the instructional level, GAI is still used mainly to enhance existing tasks rather than to fundamentally redesign learning. Although positive effects on learning outcomes have been reported, their magnitude varies across outcomes and becomes considerably smaller when publication bias is considered. Limited evidence also suggests that structured and guided use may be more beneficial than access alone. Across all three levels, concerns persist regarding accuracy, over-reliance, academic integrity, and whether AI-supported products genuinely reflect students' own learning. Taken together, the findings suggest that the educational significance of GAI lies less in its mere presence than in how its use is pedagogically structured, supported, and made visible within assessment.

Keywords: Higher education, generative artificial intelligence, GAI, academic integrity, narrative review

INTRODUCTION

With the widespread adoption of chat tools based on large language models, generative artificial intelligence (GAI) has become increasingly visible in higher education. GAI is no longer regarded merely as a new technological tool but as a force that may reshape how academic production and teaching–learning processes are carried out (Qian, 2025). University students draw on GAI for purposes such as accessing information, academic writing, idea generation, and learning support, whereas faculty members employ it to prepare course materials, assess student work, provide feedback, and support various academic tasks (Wang et al., 2025). This diversity calls for examining GAI in higher education not as a single tool with a single purpose, but in terms of the functions it assumes across different academic roles and processes.

The spread of GAI in higher education has brought with it notable educational opportunities. By offering immediate and personalised support, GAI may contribute to a range of learning outcomes, including motivation, self-regulated learning, self-efficacy, higher-order thinking, and academic performance (Chen & Cheung, 2025). Students particularly value support for academic writing, idea generation, research, and learning (Chan & Hu, 2023),

This is an open access article licensed under Creative Commons Attribution 4.0 International
Print ISSN: 1313-230X, Online ISSN: 3033-1404

while for faculty members, time savings and reduced workload are regarded as important benefits (D. Lee et al., 2024). However, these positive outcomes do not mean that GAI use is educationally beneficial in every situation. Risks such as ethical concerns, over-reliance, academic integrity problems, and the weakening of independent thinking are also increasingly debated (Chan & Hu, 2023; Chen & Cheung, 2025; McGrath et al., 2025; Perkins, 2023; Wang et al., 2025).

The decisive factor, therefore, is not merely whether GAI is used but how it is used. Completing an academic task more quickly does not mean that deeper learning has taken place, just as a reduction in faculty workload does not automatically indicate an improvement in pedagogical quality. Indeed, GAI use aimed at understanding, structuring, and developing knowledge has been associated with stronger learning outcomes, whereas use aimed merely at completing the task or directly adopting the generated output has been linked to more limited learning gains (Pallant et al., 2026). Similarly, although students' increasing delegation of cognitive tasks such as thinking, remembering, and problem solving to GAI may offer short-term convenience, it may undermine independent thinking and original work (Abbas et al., 2024; Chen & Cheung, 2025; Wang et al., 2025). Among faculty members, GAI adoption appears to be associated with factors such as perceived usefulness, trust, and the social or institutional environment, while accuracy, the detectability of plagiarism, and the reliability of assessment processes are among the main concerns (D. Lee et al., 2024; Lyu et al., 2025; Shata & Hartley, 2025).

The effects of GAI are not confined to the student and faculty levels. GAI is also employed in instructional practices such as feedback, collaborative learning, problem-based learning, simulation, and reflective learning (Wang et al., 2025). Accordingly, how students use GAI, how faculty members adopt this technology, and which functions GAI assumes in teaching–learning processes should be considered interrelated dimensions of the same higher education ecosystem. Nevertheless, because early research on GAI chatbots in higher education is dispersed across different disciplines, samples, and methodological approaches, integrated knowledge addressing these three levels together remains limited (McGrath et al., 2025).

This study aims to integrate this fragmented body of literature across students, faculty members, and teaching–learning processes, and to examine the functions of GAI in higher education, the opportunities it offers, and the risks it entails within a comparative framework. Accordingly, the following research questions are addressed:

1. What functions does GAI assume among students, faculty members, and in teaching–learning processes, and how are these functions reflected in academic practice?
2. What are the main benefits and opportunities of GAI use at these three levels?
3. What risks and limitations are associated with GAI use at these three levels?

METHOD

Research design

This study was designed as a narrative review (Green et al., 2006; Sukhera, 2022). This design was chosen mainly because the research questions aim not to determine the effect of GAI on a single outcome, but to compare functions and uses, benefits and opportunities, and risks and limitations across the levels of students, faculty members, and teaching–learning processes. The reviewed literature has a heterogeneous evidence base comprising survey, qualitative, experimental, and quasi-experimental studies as well as reviews and conceptual papers. The study therefore focuses on integrating findings from different research designs in an interpretive and comparative manner across three levels of analysis (Greenhalgh et al., 2018).

Although the literature includes systematic reviews focusing on specific dimensions (Belkina et al., 2025; Qian, 2025; Wang et al., 2025) and meta-analyses (Chen & Cheung, 2025), the aim of this study is not to replicate existing syntheses but to consider them together with primary studies in order to reveal common and divergent patterns across students, faculty members, and teaching–learning processes. Accordingly, the review prioritises interpretive synthesis over comprehensive and reproducible systematic coverage. The process by which the reviewed studies were identified and selected, together with the limitations of this approach, is set out explicitly below.

Search strategy

The literature search was conducted in the Web of Science (WoS) database. The following query was applied to the Topic (TS) field in the WoS Core Collection:

TS= ("generative artificial intelligence" OR "generative AI" OR GenAI OR ChatGPT OR "large language model*") AND TS= ("higher education" OR university* OR college*) AND TS= (university student* OR faculty OR instructor* OR educator* OR teaching* OR learning*)

The search was limited to the 2023–2026 period following the public release of ChatGPT, and only Article and Review Article document types published in English were considered. The search was performed on 19 September 2026.

Study selection process

The search criteria yielded 5717 records. These records were not screened using a systematic screening protocol. Publications were purposively selected based on their direct relevance to GAI use in higher education and to at least one of the research questions. In practice, the WoS results were sorted by relevance, the titles and abstracts of the records were reviewed in light of the research questions, and the full texts of publications deemed relevant were examined.

No citation threshold, ranking-based cut-off point, or predetermined number of studies was used in the selection. Instead, care was taken to ensure that all three levels of analysis, namely students, faculty members, and teaching–learning processes, were represented in the study set. The selection also sought to represent quantitative, qualitative, mixed-methods, experimental, and quasi-experimental designs, as well as both primary and synthesis studies. The reference lists of the included reviews were not additionally searched to identify further studies.

The final study set consisted of 14 empirical studies, three systematic reviews, one systematic review and meta-analysis, one review, and one conceptual paper. Reviews and meta-analyses were included because they help to identify broader patterns in the field. Perkins (2023) was included for its conceptual contribution regarding academic integrity and GAI use policies. This paper was used not as empirical evidence on student or teaching outcomes, but as a source providing a normative and conceptual framework.

All 20 selected publications were published in English in peer-reviewed journals between 2023 and 2026 and were retrieved through the WoS search. The study set should be regarded not as a complete or representative sample of the relevant literature, but as a selected body of literature that allows the research questions to be examined comparatively.

Data extraction and analysis

Information on author(s) and year of publication, study context, participant group, aim, sample characteristics, research design, GAI tool or model used, areas of use, reported benefits and opportunities, and risks and limitations was extracted from each publication. For reviews and meta-analyses, the number of included studies was also recorded; for the meta-analysis, information on effect sizes, heterogeneity, and publication bias was additionally extracted.

The synthesis was carried out using an analytical framework derived in advance from the research questions. This framework consists of three levels of analysis (students, faculty members, and teaching–learning processes) and three dimensions of inquiry (functions, uses, and applications; benefits and opportunities; and risks and limitations). The categories were therefore not generated inductively during the analysis but were predetermined on the basis of the research questions. The findings were examined comparatively under the relevant levels and dimensions and synthesised narratively. Studies relating to more than one level of analysis were considered under all dimensions to which they were relevant.

Findings from reviews and the meta-analysis were used to describe general patterns, whereas primary studies were used to illustrate these patterns in concrete contexts; primary studies were not counted as independent evidence in addition to the findings of the reviews. Self-reported perceptions and directly measured learning outcomes were not treated as equivalent, and findings were interpreted with the research designs of the studies in mind.

Findings on teaching–learning processes were interpreted within the SAMR model (Hamilton et al., 2016), which was also used by two of the included systematic reviews (Belkina et al., 2025; Wang et al., 2025). The 20 studies included in this review were not separately recoded according to SAMR levels; SAMR findings are reported as presented in these two reviews.

The main characteristics, evidence types, outcome sources, and levels of analysis of the studies are presented in Table 1.

Table 1. *Main characteristics of the synthesised studies*

Studies	Design / sample	Evidence type / outcome source	Level of analysis	Main contribution
Chan (2023)	Mixed; 457 students + 180 faculty/staff	Primary; perception (self-report)	Process	Pedagogical, governance, and operational policy framework for GAI integration
Chan & Hu (2023)	Survey; Hong Kong, 6 universities, 399 students	Primary; perception (self-report)	Student	Students' perceptions, benefits, and concerns regarding GAI use
Chan & Lee (2023)	Comparative survey; 399 students + 184 faculty	Primary; perception (self-report)	Student / faculty	Differences between students' and faculty members' GAI adoption tendencies and concerns
Chaudhry et al. (2023)	Case study; undergraduate programmes	Primary; product scoring	Process	Impact of ChatGPT on current assessment and academic integrity practices
Perkins (2023)	Conceptual / perspective paper	Conceptual	Process	Evaluation of LLM use from the perspective of academic integrity policies
Abbas et al. (2024)	Quantitative; 165 + 494 university students	Primary; self-report (longitudinal)	Student	Relationships between ChatGPT use, procrastination, reported memory problems, and academic performance
Ivanov et al. (2024)	Quantitative; 168 students + 130 faculty	Primary; perception (self-report)	Student / faculty	Attitudes, subjective norms, and perceived behavioural control influencing GAI adoption
D. Lee et al. (2024)	Mixed; Australia, 30 faculty, 8 interviews	Primary; perception (self-report)	Faculty / process	Faculty members' GAI use experiences, benefits, concerns, and assessment practices

Studies	Design / sample	Evidence type / outcome source	Level of analysis	Main contribution
H.-Y. Lee et al. (2024)	Randomised experiment; 61 undergraduate students	Primary; performance + self-report	Process	Effect of guidance-based ChatGPT use on self-regulation, higher-order thinking, and knowledge construction
Mohamed (2024)	Qualitative; 10 EFL faculty members	Primary; perception (qualitative)	Faculty	Benefits of ChatGPT in language teaching and concerns about critical thinking and reliability
Sun et al. (2024)	Quasi-experimental; 82 university students	Primary; performance, behaviour + perception	Student / process	Effect of ChatGPT-assisted programming on students' behaviours, performance, and perceptions
Belkina et al. (2025)	Systematic review; 21 empirical case studies	Secondary synthesis	Student / process	Practical integration of GAI in higher education and pedagogical frameworks
Chen & Cheung (2025)	Systematic review and meta-analysis; 57 studies, 5389 students	Secondary synthesis (meta-analysis)	Student	Effect of GAI use on different learning outcomes
Kim et al. (2025)	Qualitative; 20 higher education students	Primary; perception (qualitative)	Student	Experiences, benefits, and challenges in GAI-assisted academic writing
Lyu et al. (2025)	Mixed; USA, 178 faculty	Primary; perception (self-report)	Faculty	Faculty members' GAI use, perceptions, and levels of trust and distrust
McGrath et al. (2025)	Review; 23 empirical studies	Secondary synthesis	Student / process	Conceptual and pedagogical overview of early higher education research on GAI chatbots
Qian (2025)	Systematic review; 37 empirical studies	Secondary synthesis	Process	Pedagogical applications and opportunities of GAI and risks such as over-reliance
Shata & Hartley (2025)	Quantitative; USA, 294 faculty	Primary; perception (self-report)	Faculty	Usefulness, trust, and social factors influencing GAI adoption
Wang et al. (2025)	Systematic review; 139 studies	Secondary synthesis	Process	Use of GAI across instructional contexts, roles, and teaching strategies
Pallant et al. (2026)	Content analysis; 192 student reflections and course/assignment grades	Primary; performance (grades) + reflection	Student	Comparison of procedural use and knowledge-construction use in terms of learning outcomes

Note. Studies are listed chronologically by year of publication and, within the same year, alphabetically by the first author's surname. Most of the reviewed studies focus on ChatGPT or other text-based GAI tools. Chan and Hu (2023) and Chan and Lee (2023) were conducted by the same research group during the same period and report student samples of the same size (n = 399); the student data of the two studies therefore very likely overlap, and they were not treated as independent evidence in this respect. In the evidence type column, primary studies are classified according to the source of the outcome data: perception (self-report or qualitative views), performance (test, grade, or product score), and behaviour. Secondary syntheses were not counted as independent evidence additional to the primary studies.

FINDINGS

The findings are presented at three levels of analysis: students, faculty members, and teaching–learning processes.

GAI use among students

Students use GAI to obtain information and explanations, generate ideas, support

their academic writing, and solve problems. In the study by Chan and Hu (2023), students most frequently cited personalised learning, writing, brainstorming, research, and analysis. Students also position GAI not merely as an auxiliary tool but in roles such as writing assistant, virtual tutor, and digital peer (Kim et al., 2025). However, some of these findings are based on actual use, while others are based on students' expectations and experiences (Chan & Hu, 2023; Kim et al., 2025).

It is important to distinguish perceived benefits from directly measured learning outcomes. A meta-analysis covering 57 studies and 97 effect sizes found a large overall effect (g = 0.804). The largest effect was observed for language skills (g = 2.331), and positive effects were also reported for academic achievement (g = 0.633), affective-motivational state (g = 0.617), and higher-order thinking (g = 0.580). No significant effect was detected for metacognition (g = 0.078). However, the high heterogeneity across studies and the fact that correction for publication bias reduced the overall effect estimate from 0.804 to 0.321 indicate that the magnitude of these effects should be interpreted with caution (Chen & Cheung, 2025).

The benefit derived from GAI use also appears to vary according to the manner of use. Use aimed at structuring knowledge has been associated with higher course and assignment grades than use based solely on task completion (Pallant et al., 2026). However, because patterns of use were not randomly assigned, this finding cannot be interpreted as a causal relationship.

The main risks reported at the student level are accuracy, privacy, ethical concerns, and over-reliance on the tool (Chan & Hu, 2023; Kim et al., 2025). In a three-wave observational study, higher ChatGPT use was associated with more procrastination, self-reported memory problems, and lower reported grade point average (Abbas et al., 2024). However, these findings do not demonstrate causality either. Accordingly, student-level evidence suggests that the effect of GAI is not unidirectional and that benefits and risks may vary according to the manner of use, the outcome measured, and the research design of the study.

GAI use among faculty members

Faculty members use GAI to generate assessment questions, test whether assignments can be completed with GAI, and obtain support with writing and administrative tasks (D. Lee et al., 2024). In the context of language teaching, benefits such as additional practice, feedback, and time savings also stand out; however, because these findings are based on faculty members' views, they should not be regarded as directly measured teaching effectiveness (Mohamed, 2024).

Nevertheless, high awareness of GAI does not translate into regular use. In a study conducted in the United States, although 82.6% of participants reported being familiar with GAI, 64.0–75.8% reported never having used the tool, depending on the task examined. In the same study, 55.1% of participants considered their training and experience with GAI to be insufficient (Lyu et al., 2025). Similarly, in the Australian study, 78% of the faculty members who answered the relevant question reported a need for additional support (D. Lee et al., 2024). These findings suggest that adoption is related not only to familiarity with the technology but also to competence in its use and the availability of support.

Perceived usefulness, attitude, trust, social environment, and perceived behavioural control emerge as key factors in the adoption of GAI (Ivanov et al., 2024; Shata & Hartley, 2025). In particular, perceived usefulness has been reported to be more decisive than ease of use (Shata & Hartley, 2025). By contrast, concerns about accuracy, academic integrity, and student dependence persist (D. Lee et al., 2024; Lyu et al., 2025; Mohamed, 2024). While students generally approach GAI use more positively, faculty members appear to be more cautious about accuracy and pedagogical consequences; however, this difference cannot be explained solely by age

or generational differences (Chan & Lee, 2023).

GAI use in teaching–learning processes

GAI is used for various instructional purposes, such as feedback, assessment, learning support, content generation, collaborative work, and the development of higher-order skills (Belkina et al., 2025; Qian, 2025; Wang et al., 2025). To assess the extent to which these applications transform the teaching process, Wang et al. (2025) and Belkina et al. (2025) drew on the SAMR model. In this model, substitution refers to performing an existing task through technology; augmentation, to functionally improving the task; modification, to redesigning the task; and redefinition, to creating a learning activity that was previously not possible. In the review by Wang et al. (2025), which covered 139 studies, most applications were concentrated at the augmentation level, whereas applications at the redefinition level remained more limited. In contrast, Belkina et al. (2025), who focused on case studies, reported notable examples at the modification and redefinition levels as well. These findings indicate that GAI integration does not occur at a single pedagogical level and may vary depending on how it is implemented.

The pedagogical grounding of this integration, however, appears to be more limited. Of the 23 studies examined by McGrath et al. (2025), only three explicitly drew on learning theories. This suggests that the incorporation of GAI into teaching processes is not always based on an explicit pedagogical framework.

Experimental and quasi-experimental findings also indicate that access to GAI alone is not sufficient and that how its use is structured may be important. In the study by H.-Y. Lee et al. (2024), in which 61 students were individually randomly assigned to two groups (n = 31 and n = 30), ChatGPT use that encouraged students to produce their own solution first and provided hints rather than direct answers was associated with higher post-test and delayed test results than the control group using standard ChatGPT. The same study also reported positive outcomes in terms of cognitive and behavioural engagement, self-efficacy, and higher-order thinking. By contrast, in the programming study by Sun et al. (2024), no significant difference in performance was found between the ChatGPT-assisted group and the group engaged in self-directed programming ($p = .204$). Although these two studies are based on different contexts and comparison conditions, taken together they suggest that the pedagogical effect of GAI may depend not only on access to the tool but also on the instructional structure within which it is used.

In assessment, a central problem is the growing uncertainty about the extent to which a submitted product reflects the student's own knowledge and skills. ChatGPT outputs have been reported to receive scores similar to those of high-scoring student work on some tasks, while the detection tools of the time were insufficient to determine authorship (Chaudhry et al., 2023). However, this finding does not indicate an increase in student learning; rather, it shows that the relationship between the source of a performance product and the student's actual learning needs to be evaluated more carefully. It is therefore recommended that permitted GAI use be clearly defined, that GAI assistance be declared, and that the student's original contribution be made visible (Perkins, 2023). At the institutional level, it has been argued that GAI policies should cover not only pedagogical practices but also training, infrastructure, data security, and governance. Nevertheless, the effectiveness of these proposed policy frameworks has not yet been empirically tested (Chan, 2023).

DISCUSSION

When the three levels of analysis are considered together, a common pattern emerges that is not directly visible in individual studies: at each level, the most visible indicator does not necessarily

correspond to the educational outcome of interest. At the student level, a product generated with GAI does not by itself show what the student has learned. At the faculty level, being familiar with the tool does not mean using it effectively in teaching. At the level of teaching–learning processes, the use of GAI does not mean that teaching has been transformed. The central interpretation offered by this review is that these three mismatches may not be independent of one another. A lack of institutional support and training may limit faculty members' ability to integrate GAI into their courses in a structured manner (D. Lee et al., 2024; Lyu et al., 2025). When integration remains at the level of merely supporting existing tasks, students may use GAI in a more individual and unguided manner, and the associated risks of over-reliance and threats to assessment validity may consequently become more pronounced. However, this relationship between levels has not been directly tested; the connection proposed here rests on a joint interpretation of the existing findings and should be tested empirically in the future.

In making this interpretation, differences in the strength of the evidence must also be taken into account. While meta-analytic and experimental findings are available at the student level, the evidence at the faculty level is largely based on perception, attitude, and self-report data. Consequently, conclusions regarding learning outcomes appear to rest on a comparatively stronger empirical basis. Nevertheless, this evidence also has important limitations: heterogeneity across studies is high, corrections for publication bias markedly reduce effect estimates, and some of the experimental findings are based on small samples. For this reason, inferences concerning faculty practices and institutional outcomes in particular should be treated with greater caution.

One point on which the findings converge is that the educational value of GAI cannot be explained by access to the tool alone. While guided and structured use may yield positive results for some learning outcomes, access to GAI alone does not produce a consistent gain in achievement (Chen & Cheung, 2025; H.-Y. Lee et al., 2024; Pallant et al., 2026; Sun et al., 2024). This points to the importance of structuring GAI so that it supports, rather than replaces, the student's cognitive activity. In this respect, activities that require students first to produce their own solution and then to question GAI output, compare alternatives, and justify their decisions may offer a more productive framework for use.

Similarly, access to technology alone does not appear to be sufficient for faculty members' adoption of GAI. In addition to perceived usefulness, trust, attitude, and social environment, training, clear rules, and technical or institutional support are also associated with the decision to use it (Ivanov et al., 2024; D. Lee et al., 2024; Lyu et al., 2025; Shata & Hartley, 2025). The differing expectations and concerns of students and faculty members regarding GAI further require institutions to develop support mechanisms that take the needs of both groups into account rather than adopting a one-size-fits-all approach (Chan & Lee, 2023).

The findings further indicate that assessment practices need to be reconsidered. In an environment where GAI use has become widespread, assessments based solely on the final submitted product may carry greater uncertainty about the extent to which they reflect the student's own knowledge and skills (Chaudhry et al., 2023). Institutions may therefore consider assessment approaches that make visible not only the student's final product but also their drafts, decisions, justifications, and interactions with GAI (Perkins, 2023). Such practices should be supported by clear rules of use, data security, faculty training, and institutional responsibilities (Chan, 2023).

Limitations

This narrative review has several limitations. First, the classification of the 20 included publications was carried out by a single researcher, and no standardised

quality appraisal tool was used to assess the methodological quality of the studies. In addition, because there is a natural overlap between the student, faculty, and teaching–learning process axes, some studies were considered at more than one level of analysis. These axes should therefore not be interpreted as entirely independent groups of studies.

A further limitation concerns the way in which studies were selected. Publications were identified through researcher-led selection on the basis of their relevance to the research questions rather than according to a predefined systematic screening protocol. Consequently, the possibility of selection bias cannot be ruled out, and the resulting synthesis represents not the entire body of relevant literature, but a portion selected in line with the research questions. Moreover, restricting the search to the 2023–2026 period, to studies published in English, and to peer-reviewed journal articles excluded conceptual debates predating ChatGPT as well as book chapters, conference papers, and other publication types.

The reviewed studies differ considerably in terms of research design, sample, context, and measurement methods. Owing to this heterogeneity, self-reported perceptions and directly measured learning outcomes were not treated as equivalent evidence. Furthermore, since some primary studies may also have been included in the reviews examined here, the findings of reviews and meta-analyses were used to identify general patterns, whereas primary studies were not treated as independent evidence in addition to these syntheses. The high heterogeneity reported in meta-analytic studies was also taken into account when interpreting the findings (Chen & Cheung, 2025; McGrath et al., 2025).

Finally, the rapidly changing nature of the GAI field constitutes an important limitation with regard to the currency and generalisability of this review. Most of the included studies focused on ChatGPT and other text-based tools, whereas multimodal GAI systems that generate images, audio, and video were only marginally represented. It therefore remains unclear whether the findings are specific to text-based GAI tools or can be generalised to broader GAI applications. The rapid pace of technological development also makes it necessary to re-evaluate the current findings at regular intervals.

Recommendations for future research

Future research would benefit from focusing primarily on the long-term effects of GAI use on students' learning outcomes. In particular, outcomes such as skill development, knowledge retention, and the later recall of what has been learned could be examined through longitudinal and experimental designs. Comparing traditional instruction, direct GAI use, and GAI use guided by faculty members may help to disentangle the effects of the tool itself from those of pedagogical structuring. Conducting similar research across different disciplines would also clarify whether the effects of GAI vary by field.

From a pedagogical perspective, it is important to develop instructional practices that require students to question, verify, and evaluate information generated by GAI on the basis of their own reasoning rather than simply using it, and to test the effectiveness of these practices. Studies focusing on faculty members, in turn, should not rely solely on attitude, perception, and self-report data; classroom practices, instructional designs, and their consequences for student learning should also be examined directly.

Finally, a considerable proportion of the existing literature focuses on ChatGPT and other text-based tools. Extending future studies to multimodal GAI systems that generate images, audio, and video would contribute to a broader picture of GAI use in higher education. These research directions are broadly consistent with the limitations and recommendations for future research reported in the reviewed studies (Chen & Cheung, 2025; D. Lee et al., 2024; H.-Y. Lee et al., 2024; Lyu et al., 2025; Qian, 2025; Sun et al., 2024).

CONCLUSION

This review suggests that the impact of GAI in higher education cannot be evaluated solely on the basis of the extent to which it is used. The manner of use for students, pedagogical and institutional support for faculty members, and the quality of integration for teaching–learning processes appear to be decisive. Approaches aimed at enhancing the educational value of GAI in higher education should therefore go beyond expanding access to the tool and focus on supporting guided use, preparing faculty members, and developing assessment processes that make students' own cognitive contributions visible.

REFERENCES

Abbas, M., Jam, F. A., & Khan, T. I. (2024). Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students. *International Journal of Educational Technology in Higher Education, 21*, Article 10. https://doi.org/10.1186/s41239-024-00444-7

Belkina, M., Daniel, S., Nikolic, S., Haque, R., Lyden, S., Neal, P., Grundy, S., & Hassan, G. M. (2025). Implementing generative AI (GenAI) in higher education: A systematic review of case studies. *Computers and Education: Artificial Intelligence, 8*, Article 100407. https://doi.org/10.1016/j.caeai.2025.100407

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education, 20*, Article 38. https://doi.org/10.1186/s41239-023-00408-3

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education, 20*, Article 43. https://doi.org/10.1186/s41239-023-00411-8

Chan, C. K. Y., & Lee, K. K. W. (2023). The AI generation gap: Are Gen Z students more interested in adopting generative AI such as ChatGPT in teaching and learning than their Gen X and millennial generation teachers? *Smart Learning Environments, 10*, Article 60. https://doi.org/10.1186/s40561-023-00269-3

Chaudhry, I. S., Sarwary, S. A. M., El Refae, G. A., & Chabchoub, H. (2023). Time to revisit existing student's performance evaluation approach in higher education sector in a new era of ChatGPT: A case study. *Cogent Education, 10*(1), Article 2210461. https://doi.org/10.1080/2331186X.2023.2210461

Chen, S., & Cheung, A. C. K. (2025). Effect of generative artificial intelligence on university students learning outcomes: A systematic review and meta-analysis. *Educational Research Review, 49*, Article 100737. https://doi.org/10.1016/j.edurev.2025.100737

Green, B. N., Johnson, C. D., & Adams, A. (2006). Writing narrative literature reviews for peer-reviewed journals: Secrets of the trade. *Journal of Chiropractic Medicine, 5*(3), 101–117. https://doi.org/10.1016/S0899-3467(07)60142-6

Greenhalgh, T., Thorne, S., & Malterud, K. (2018). Time to challenge the spurious hierarchy of systematic over narrative reviews? *European Journal of Clinical Investigation, 48*(6), Article e12931. https://doi.org/10.1111/eci.12931

Hamilton, E. R., Rosenberg, J. M., & Akcaoglu, M. (2016). The Substitution Augmentation Modification Redefinition (SAMR) model: A critical review and suggestions for its use. *TechTrends, 60*(5), 433–441. https://doi.org/10.1007/s11528-016-0091-y

Ivanov, S., Soliman, M., Tuomi, A., Alkathiri, N. A., & Al-Alawi, A. N. (2024). Drivers of generative AI adoption in higher education through the lens of the Theory of Planned Behaviour. *Technology in Society, 77*, Article 102521. https://doi.org/10.1016/j.techsoc.2024.102521

Kim, J., Yu, S., Detrick, R., & Li, N. (2025). Exploring students' perspectives on generative AI-assisted academic writing. *Education and Information Technologies, 30*, 1265–1300. https://doi.org/10.1007/s10639-024-12878-7

Lee, D., Arnold, M., Srivastava, A., Plastow, K., Strelan, P., Ploeckl, F., Lekkas, D., & Palmer, E. (2024). The impact of generative AI on higher education learning and teaching: A study of educators' perspectives. *Computers*

and Education: Artificial Intelligence, 6, Article 100221. https://doi.org/10.1016/j.caeai.2024.100221

Lee, H.-Y., Chen, P.-H., Wang, W.-S., Huang, Y.-M., & Wu, T.-T. (2024). Empowering ChatGPT with guidance mechanism in blended learning: Effect of self-regulated learning, higher-order thinking skills, and knowledge construction. *International Journal of Educational Technology in Higher Education, 21*, Article 16. https://doi.org/10.1186/s41239-024-00447-4

Lyu, W., Zhang, S., Chung, T. R., Sun, Y., & Zhang, Y. (2025). Understanding the practices, perceptions, and (dis)trust of generative AI among instructors: A mixed-methods study in the U.S. higher education. *Computers and Education: Artificial Intelligence, 8*, Article 100383. https://doi.org/10.1016/j.caeai.2025.100383

McGrath, C., Farazouli, A., & Cerratto-Pargman, T. (2025). Generative AI chatbots in higher education: A review of an emerging research area. *Higher Education, 89*, 1533–1549. https://doi.org/10.1007/s10734-024-01288-w

Mohamed, A. M. (2024). Exploring the potential of an AI-based chatbot (ChatGPT) in enhancing English as a foreign language (EFL) teaching: Perceptions of EFL faculty members. *Education and Information Technologies, 29*, 3195–3217. https://doi.org/10.1007/s10639-023-11917-z

Pallant, J. L., Blijlevens, J., Campbell, A., & Jopp, R. (2026). Mastering knowledge: The impact of generative AI on student learning outcomes. *Studies in Higher Education, 51*(4), 714–735. https://doi.org/10.1080/03075079.2025.2487570

Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice*, *20*(2), Article 07. https://doi.org/10.53761/1.20.02.07

Qian, Y. (2025). Pedagogical applications of generative AI in higher education: A systematic review of the field. *TechTrends, 69*, 1105–1120. https://doi.org/10.1007/s11528-025-01100-1

Shata, A., & Hartley, K. (2025). Artificial intelligence and communication technologies in academia: Faculty perceptions and the adoption of generative AI. *International Journal of Educational Technology in Higher Education, 22*, Article 14. https://doi.org/10.1186/s41239-025-00511-7

Sukhera, J. (2022). Narrative reviews: Flexible, rigorous, and practical. *Journal of Graduate Medical Education, 14*(4), 414–417. https://doi.org/10.4300/JGME-D-22-00480.1

Sun, D., Boudouaia, A., Zhu, C., & Li, Y. (2024). Would ChatGPT-facilitated programming mode impact college students' programming behaviors, performances, and perceptions? An empirical study. *International Journal of Educational Technology in Higher Education, 21*, Article 12. https://doi.org/10.1186/s41239-024-00446-5

Wang, P., Jing, Y., & Shen, S. (2025). A systematic literature review on the application of generative artificial intelligence (GAI) in teaching within higher education: Instructional contexts, process, and strategies. *The Internet and Higher Education, 65*, Article 100996. https://doi.org/10.1016/j.iheduc.2025.100996.