TechTrends (2025) 69:1105–1120
https://doi.org/10.1007/s11528-025-01100-1

ORIGINAL PAPER

Pedagogical Applications of Generative AI in Higher Education: A Systematic Review of the Field

Yufeng Qian[1]

Accepted: 23 May 2025 / Published online: 11 June 2025

Abstract
The release of ChatGPT in late 2022 marked the beginning of a rapid transformation in higher education, soon followed by the development of multimodal generative AI programs. As this technology becomes increasingly integrated into teaching and learning, it is crucial to evaluate its current use and impact. This systematic literature review captures the initial academic response to generative AI, providing insights into how higher education has adopted this transformative technology in its first two years. The findings indicate that while some themes from the pre-ChatGPT era persist, new and emerging trends—particularly in fostering creativity, critical thinking, learning autonomy, and prompt literacy—are now taking shape. This shift underscores a growing emphasis on the pedagogical integration of generative AI. However, the review also highlights a key tension: while generative AI enhances efficiency, it raises concerns about overreliance, potentially leading to the outsourcing of critical cognitive and metacognitive skills. To address these challenges and fully harness the potential of generative AI, future research should focus on exploring multimodal generative AI tools and fostering student–teacher-AI collaboration.

Keywords Generative AI · Teaching and learning · Higher education · Systematic literature review

Introduction

The release of OpenAI's ChatGPT-3.5 in November 2022 marked a transformative moment for education, sparking the development of generative AI (GenAI) tools that are reshaping teaching, learning, and assessment practices (EDUCAUSE, 2023). GenAI, powered by Large Language Models (LLMs), can summarize and generate content across various modalities, including text, image, audio, and video (MIT News, 2023). Initially, single-modal GenAI tools like DALL-E for image, Suno for music, and Google's Imagen for video gained traction (McKinsey & Company, 2024). By late 2023, the advent of multimodal programs, such as GPT-4, Google's Gemini, and Meta's ImageBind, marked a new phase, enabling simultaneous integration and generation across media types — a development that expands the reach and impact of GenAI (Meta, 2023a).

GenAI has immense potential to advance pedagogical approaches and the experiences of teaching, learning, and assessment in higher education. Yet, as a technology that is both "transformative" and "disruptive," its future development demands a nuanced understanding of its current applications and the untapped potential it holds (McCormack, 2023; Robert, 2024). Since late 2022, researchers and practitioners have increasingly integrated GenAI into their teaching and learning, leading to a rapid increase in studies on its applications. Yet, despite this growing body of research, significant gaps remain in understanding which GenAI tools are used, how they are applied, and for what learning tasks. A comprehensive exploration is therefore essential to map the current landscape, uncover GenAI's unrealized potential, and address its associated challenges.

Background

Before undertaking the systematic review, it is essential to address two foundational questions:

What exactly is GenAI, and which GenAI programs are currently available? Additionally, what prior systematic reviews on this topic are available?

✉ Yufeng Qian
jqian@lsu.edu

[1] Lutrill & Pearl Payne School of Education, College of Human Sciences & Education, Louisiana State University, 221 Peabody Hall, Baton Rouge, LA 70803, USA

Development of GenAI

GenAI's roots extend back to the 1960 s and 1970 s, during which time early AI-based text generation tools, such as MIT's ELIZA (Weizenbaum, 1966), SCHOLAR (Carbonell, 1970), and MYCIN (Shortliffe et al., 1975), were developed. These early systems relied on rule-based pattern matching and scripted responses, which limited their ability to generate novel content or comprehend conversational context in depth (Natale, 2019). As a result, these early programs do not meet the criteria for modern GenAI systems.

The emergence of GenAI as we know it today occurred with the development of LLMs, which utilize extensive datasets (e.g., books, web crawls, encyclopedias) to generate contextually appropriate and coherent content (IEEE Spectrum, 2024). OpenAI's launch of ChatGPT-3.5 in November 2022 was a significant milestone, popularizing GenAI's text generation capabilities and sparking interest in single-modal tools. These early GenAI systems, including DALL-E (image generation), Midjourney (image manipulation), and others, set the stage for the rapid development that followed (Medium, 2024; MIT Management, 2024).

By late 2023, multimodal GenAI tools began to emerge, capable of processing and generating content across multiple media types. Notable examples include Google's Gemini and Meta's Movie Gen. These systems enable users to integrate diverse media, such as text, images, audio, and video, to create richer, more immersive educational experiences (Meta, 2023b). Google's Notebook LM, which integrates LLMs with structured data and interactive interfaces, is an example of the evolving sophistication in multimodal systems (Google, 2024). These advancements have opened up new possibilities for education by combining text-based learning with dynamic visuals, sounds, and interactivity.

From an educational perspective, GenAI is highly adaptable, offering personalized learning experiences tailored to individual student needs and preferences. By adjusting content complexity and providing real-time feedback, GenAI helps foster personalized learning environments that support diverse learning styles and needs (Owan et al., 2023; Wang et al., 2024). Its multimodal resources also make complex concepts more accessible by combining visual, auditory, and interactive elements (Lee et al., 2023).

Extant Systematic Reviews on AI and ChatGPT in Higher Education

Several prior systematic reviews have examined the applications of AI in education. Zawacki-Richter et al. (2019) analyzed 146 publications on AI from 2007 to 2018, revealing that AI research in education was primarily conducted in the fields of Computer Science and STEM. The review categorized AI applications into four primary areas: profiling and prediction, assessment and evaluation, adaptive systems and personalization, and intelligent tutoring systems. Similarly, Crompton and Burke (2023) reviewed AI research trends between 2016 and 2022 and highlighted a surge in publications on AI applications in education, particularly focusing on language learning and intelligent tutoring systems.

Since the release of ChatGPT, systematic reviews have increasingly explored its applications within higher education. Crompton and Burke (2024) reviewed 44 studies, emphasizing ChatGPT's potential to enhance teaching support, automate tasks, and foster professional development. They also identified critical limitations, such as inaccuracies, biases, and risks of misuse. Similarly, Mohebi (2024) analyzed 32 studies, highlighting ChatGPT's capability to enhance personalized learning and collaborative activities while noting significant challenges related to its pedagogical integration.

Regarding broader explorations of GenAI, both Preiksaitis and Ross (2023) and Ogunleye et al. (2024) offered general examinations of GenAI without a specific focus on pedagogical applications. Preiksaitis and Ross (2023) conducted a scoping review of 68 studies in medical education published between January 2022 and June 2023, mapping opportunities and limitations of GenAI tools in this specialized domain. Similarly, Ogunleye et al. (2024) analyzed 355 studies on GenAI in higher education spanning 2018–2023, but their work prioritized metadata classification and trends in AI adoption over a detailed exploration of teaching and learning methodologies.

Despite widespread interest in AI, particularly Gen AI tools like ChatGPT, no comprehensive systematic review to date has focused on the pedagogical applications of GenAI in higher education, specifically targeting studies published in 2023 and 2024 – the first two years post ChatGPT. This review addresses that gap by examining how GenAI tools beyond ChatGPT have been used pedagogically in higher education, highlighting their existing applications and the areas where their potential remains underexplored.

Research Questions

This study's primary research question is: What is the current state of integrating GenAI in teaching and learning within higher education? The study will address specific sub-questions related to the types of GenAI employed, their specific usages, and the challenges faced in implementation.

1. What patterns characterize the geographic, disciplinary, and technological distribution of GenAI adoption in higher education?
2. How are GenAI tools being pedagogically deployed in higher education teaching and learning?
3. What challenges emerge from current GenAI implementations in higher education teaching and learning?

Methods

This study follows a systematic review approach, which is a rigorous research method used to synthesize and evaluate existing studies on a specific topic. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were followed to ensure transparency, reproducibility, and reliability in the process of article identification, screening, and selection (Page et al., 2021).

Identification

Three primary search terms were used to conduct this systematic review:"generative artificial intelligence"(or its variants"GenAI,""Gen AI,""GAI"),"higher education,"and"teaching and learning."These terms were combined using Boolean operators (AND, OR) to ensure comprehensive coverage of the topic. The search was conducted across three major academic databases: ERIC, Web of Science, and ScienceDirect. These databases were selected because they provide access to high-quality, peer-reviewed articles in the field of educational technology, including journals such as *Computers and Education*, *Computers and Education: Artificial Intelligence*, and the *International Journal of Educational Technology in Higher Education*. The search focused on articles that appeared in the title, abstract, and keywords.

Only peer-reviewed journal articles published in 2023 and 2024 were included in this review. This time frame was chosen to focus on the post-launch period of ChatGPT-3.5 in November 2022, which marked a significant increase in research on GenAI applications in higher education. Conference papers, theoretical or conceptual articles, and systematic review studies were excluded to maintain the empirical focus of the review.

Screening

A total of 262 articles were initially identified from ERIC, Web of Science, and ScienceDirect. After screening the articles for relevance, a significant portion (n = 121, 46%) was excluded because these studies relied primarily on surveys and interviews to explore student and faculty perceptions of GenAI, rather than its direct application in teaching and learning. Additionally, review articles and editorials were excluded, as these do not provide original research data. Articles that did not explicitly focus on the use of GenAI in higher education teaching and learning were also excluded. After removing duplicates, 37 studies remained for inclusion in this review. Figure 1 presents a flowchart illustrating the article identification and screening process.

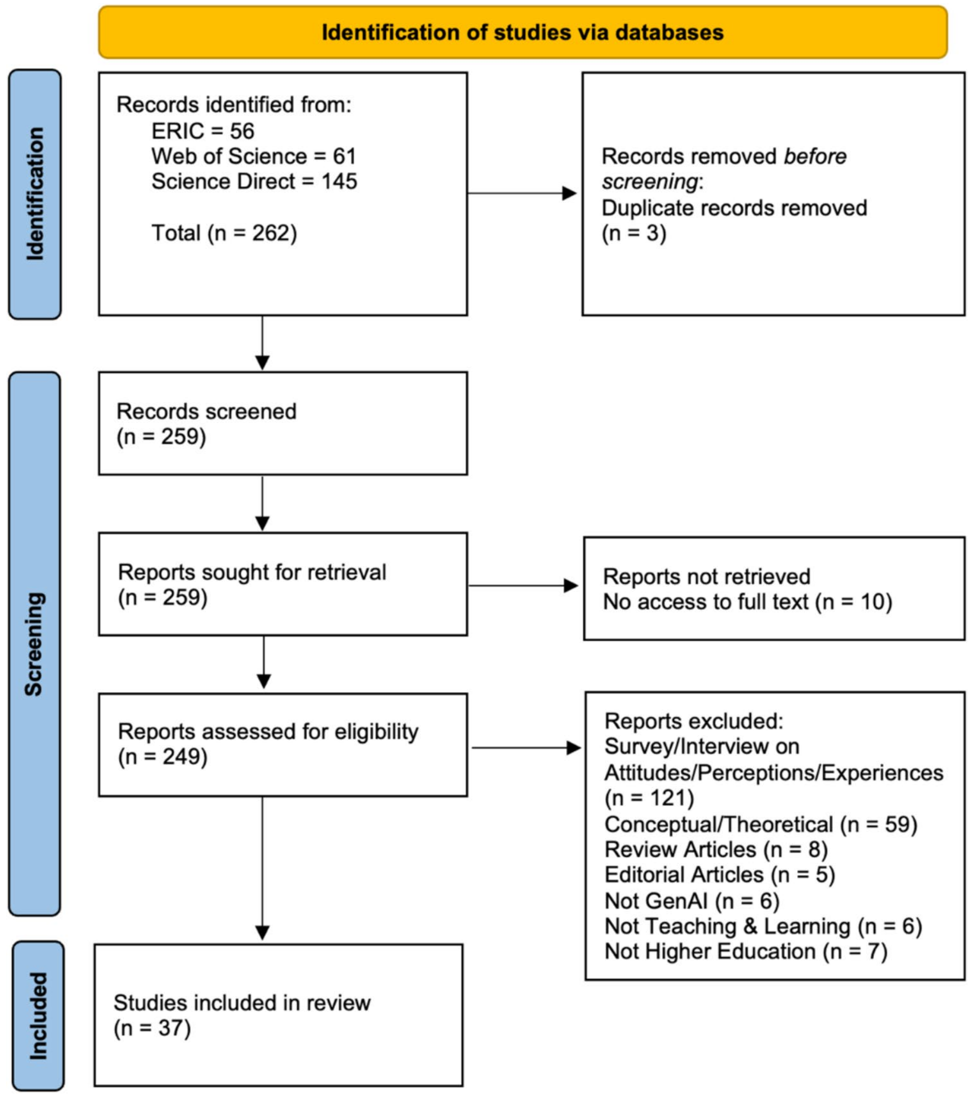
Coding and Thematic Analysis

Each article was initially coded for descriptive metadata, including authorship, publication year, and the first author's country or region, to identify temporal and geographic trends in GenAI research. Articles were further categorized by the types of GenAI tools discussed, creating an inventory of established and emerging technologies. To address the study's central questions—pedagogical applications of GenAI in higher education and challenges in implementation—a grounded coding approach (Corbin & Strauss, 2014) was employed. This inductive process involved three iterative phases: open coding, focused coding, and theoretical coding. During open coding, a line-by-line analysis of all articles identified initial concepts (e.g., automation of routine Q&A tasks), which were later synthesized into broader codes (e.g., automated feedback/assessment) during focused coding. Theoretical coding then linked these codes to the overarching research questions, resulting in 42 codes for pedagogical applications and 25 codes for challenges.

Thematic analysis (Thomas & Harden, 2008) was conducted to interpret findings in relation to the research questions. The 42 pedagogical codes were grouped into three central themes, while the 25 challenge codes coalesced into four thematic areas. This phase emphasized deriving meaning from patterns rather than imposing preconceived categories. This dual approach (i.e., grounded coding and thematic analysis) ensured findings remained anchored in the dataset but also provided a structured framework to map an emerging field like GenAI in education.

To enhance methodological rigor, NVivo 14, an AI-powered qualitative analysis program, was used as a secondary analytical tool. The software's machine learning algorithms generated preliminary coding suggestions during open coding, which the author reviewed and refined. After manual coding of the full dataset, NVivo's Coding Comparison Query tool cross-verified a subset of 4 articles (10% of the dataset) to assess inter-rater reliability between the author's codes and the AI's suggestions. Discrepancies, such as mismatched code assignments, were resolved through iterative calibration, where codebook definitions were refined and NVivo's AI was retrained on updated criteria. After two rounds of revision, 100% consensus was achieved, ensuring consistency across code definitions and thematic groupings.

Fig. 1 PRISMA flow chart of article identification and screening (Page et al., 2021)



Findings and Discussion

Research Question One: What Patterns Characterize the Geographic, Disciplinary, and Technological Distribution of GenAI Adoption in Higher Education?

Geographical Distribution

Research on GenAI in education is predominantly concentrated in China (including Hong Kong, which is administered as a Special Administrative Region with a high degree of autonomy), with a combined total of 11 publications highlighting a strong regional focus on GenAI-enhanced learning. The U.S. follows with eight publications. Beyond these regions, contributions span multiple continents, including **Europe** (Finland, Germany, Netherlands, Serbia, Sweden, UK), **Asia** (Singapore, Taiwan, Thailand, Turkey), **Africa** (South Africa, Ghana), **North America** (Canada), **South America** (Brazil), and **Oceania** (Australia).

China's national strategy, exemplified by the Next Generation Artificial Intelligence Development Plan (2017), has prioritized AI innovation, directing substantial funding toward educational applications and encouraging universities to align their research with national objectives. This approach has accelerated empirical studies on GenAI in higher education (Knox, 2023; State Council of China, 2017). Additionally, Chinese higher education institutions, including those in Hong Kong, place a strong emphasis on STEM disciplines, facilitating the rapid integration of GenAI tools in fields such as coding and engineering. These policy and cultural factors contribute to China and Hong Kong's current dominance in applied GenAI research publications in education Figure 2.

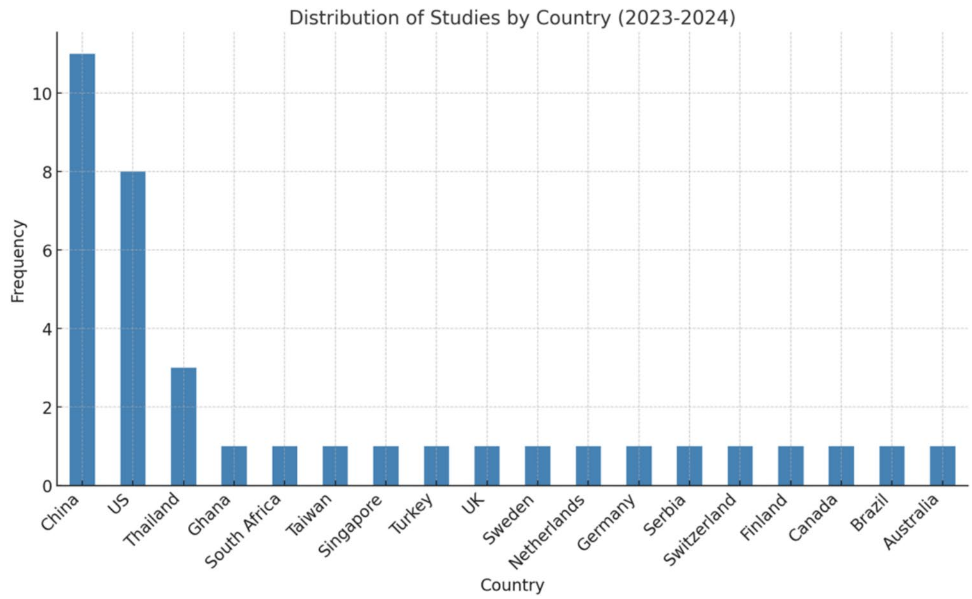


Fig. 2 Geographical distribution of studies

GenAI Tools Used

Figure 3 illustrates the distribution of GenAI tools, highlighting ChatGPT as the dominant choice by a substantial margin. Nevertheless, customized GPT-based applications are gaining traction, indicating a growing interest in AI tools tailored to specific educational contexts. Notable examples include AnatomyGPT for anatomy education (Collins et al., 2024), AI Learning Companion Systems (LCS) designed to enhance students'self-efficacy in information literacy (Hu et al., 2024), and retrieval-augmented generation (RAG) chatbots, which blend retrieval and generative capabilities for more sophisticated AI interactions (Guo et al., 2024). Additionally, beyond text-based outputs, tools for image and video generation are emerging, expanding the scope of GenAI applications (Cummings et al., 2024; Koh et al., 2024; Netland et al., 2025; Tsao & Nogues, 2024 Furthermore, writing support tools like Wordtune, Rytr, Wordtune, and specialized platforms such as Rytr assist students in improving their academic writing (Cummings et al., 2024; Koh et al., 2024; Tsao and Nogues, 2024). GenAI is also being increasingly

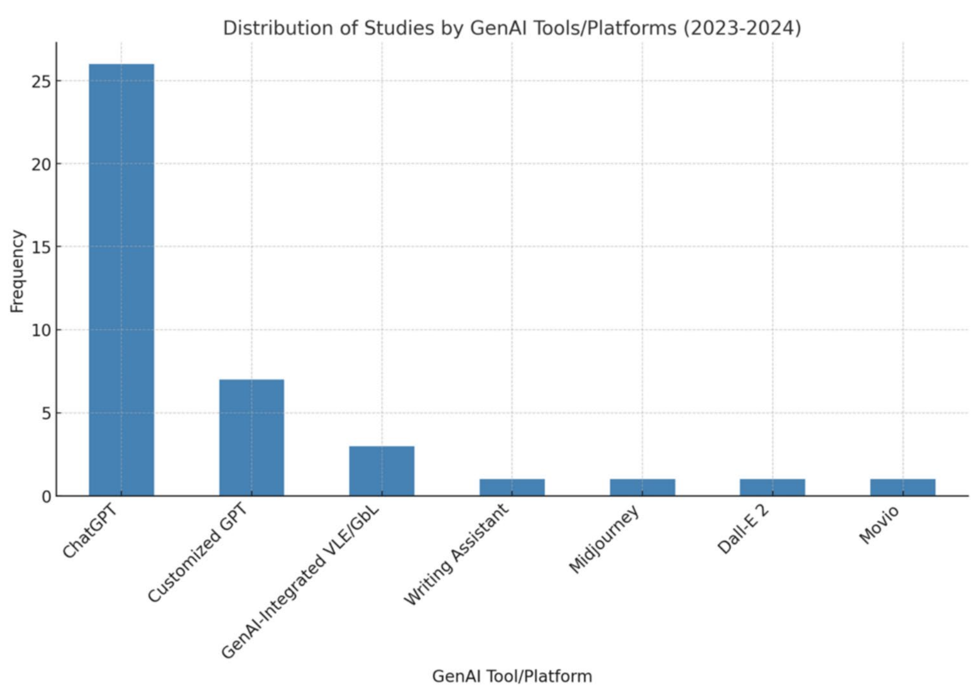


Fig. 3 Distribution of GenAI tools/platforms

integrated into game-based learning environments to facilitate adaptive learning as an embedded feature (Song et al., 2024).

Despite the diversity of available tools, ChatGPT continues to dominate due to its ease of use (especially version 3.5), its effectiveness in text-based academic tasks (e.g., writing, critical thinking), and its general suitability for higher education contexts (Kasneci et al., 2023). However, specialized GPT tools, like AnatomyGPT, have emerged to overcome specific limitations associated with general-purpose platforms such as ChatGPT, Google Gemini, and Claude. These specialized GPTs provide tailored capabilities for discipline-specific teaching and learning needs. However, tools involving visual media—such as image and video generation applications—often face higher barriers, including substantial technical requirements, limited applicability in non-visual disciplines, and significant ethical concerns around misinformation and copyright (Bender, 2021). Hybrid AI applications, such as virtual reality (VR) combined with AI (Muengsan & Chatwattana, 2024), face even greater hurdles, demanding interdisciplinary collaboration and substantial investment in resources.

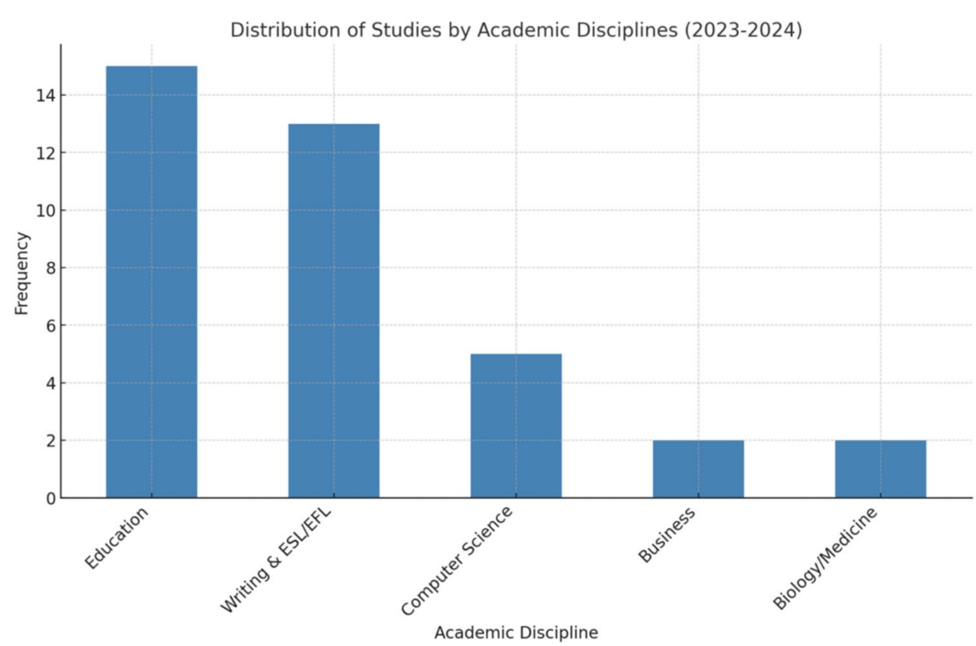
Overall, the current trend suggests a preference for GenAI tools that prioritize simplicity, accessibility, and alignment with core educational tasks, such as writing and critical thinking (Kasneci et al., 2023). Nonetheless, the exploration of more specialized, immersive, and hybrid AI experiences highlights an important avenue for future development—one that promises richer, more interactive learning opportunities, but requires careful attention to practical, ethical, and collaborative considerations.

Academic Disciplines

Figure 4 illustrates the distribution of GenAI application studies by academic discipline. To categorize these 37 studies, each was coded based on its primary educational context. Studies with disciplinary overlaps—such as those combining Education with ESL or Computer Science—were classified according to their predominant pedagogical focus. Note that the category"Education"broadly includes studies addressing the process of improving or expanding educational systems, practices, and outcomes, including general or liberal education, teacher preparation, and critical skill development (e.g., creativity, critical thinking, learner autonomy, and prompt literacy).

The 37 empirical studies represent a diverse range of academic disciplines, with Education, Writing and English Language Learning (including ESL/EFL), and Computer Science emerging as the most prominent. Education accounts for 15 studies, reflecting its broad applicability in pedagogical innovation and learning enhancement. For example, Yang et al. (2024) examined student agency facilitated by GenAI, while Van den Berg and du Plessis (2023) explored ChatGPT's use in lesson planning within teacher education. Writing and English Language Learning followed closely with 13 studies, divided into Writing/Composition (5 studies)—as exemplified by Bedington et al.'s (2024) research on AI-supported writing processes—and ESL/EFL (8 studies), such as Moorhouse et al. (2024), who investigated first-language (L1) use in second-language (L2) classrooms, highlighting GenAI's role in linguistic and compositional skill development. Computer Science, comprising 5 studies, emphasizes technical

Fig. 4 Distribution of studies by academic disciplines



innovation and practical applications, including Zhong and Kim (2024) utilization of ChatGPT for R programming in business analytics and Allen et al.'s (2024) development of the specialized"Q-Module-Bot."Other disciplines, such as Business (2 studies), Biology (1 study), and Medicine (1 study), are comparatively underrepresented.

The dominance of Education, Writing and English Language Learning, and Computer Science reflects GenAI's strong alignment with pedagogical needs, technical suitability, and practical demands within these fields. Education's prominence can be attributed to its versatility, using GenAI to foster essential cross-disciplinary skills like creativity, critical thinking, and digital literacy (Selwyn, 2022). Moreover, its application in teacher education aligns closely with contemporary educational priorities, including technological fluency and pedagogical advancement (Prensky, 2012; U.S. Department of Education, 2023).

Similarly, the prominence of Writing/Composition (Cummings et al., 2024; Farazouli et al., 2024; Nguyen et al., 2024; Tsao and Nogues, 2024) and ESL/EFL studies (Escalante et al., 2024; Guo & Li, 2024; Guo et al., 2024; Waluyo & Kusumastuti, 2024; Wiboolyasarin et al., 2024) stems from GenAI's inherent strengths in text-based tasks, including grammar correction, writing support, and conversational practice (Hwang and Chang, 2023). This emphasis aligns closely with Crompton and Burke's (2023) systematic review findings, reaffirming ChatGPT's central role in language acquisition and writing support within higher education contexts.

Computer Science's representation (Allen et al., 2024; Groothuijsen et al., 2024; Song et al., 2024; Yilmaz and Yilmaz, 2023) aligns with previous findings by Zawacki-Richter et al. (2019) on pre-GenAI programs, reflecting the discipline's natural synergy with AI capabilities in programming, data analysis, and technical problem-solving, as well as institutional readiness to adopt innovative technological solutions. In contrast, fields such as Biology (Collins et al., 2024), Business (Milić et al., 2024; Netland et al. 2025), and Medicine (Hudon et al., 2024) remain underexplored, largely due to challenges in adapting general-purpose GenAI models to specialized or niche contexts.

Overall, the existing literature prioritizes text-centric and technically aligned disciplines, emphasizing ChatGPT's prevalent use. Moving forward, there is a clear need for future research to expand the exploration of domain-specific GenAI applications, extending beyond the current dominance of ChatGPT to uncover richer, more diverse possibilities for GenAI-supported learning.

Research Question Two: How Are GenAI Tools Being Pedagogically Deployed in Higher Education Teaching and Learning?

Figure 5 maps the pedagogical applications of GenAI into three distinct yet interrelated themes: (1) automated feedback and assessment, (2) learning supports, and (3) development of critical skills. These themes were designed to

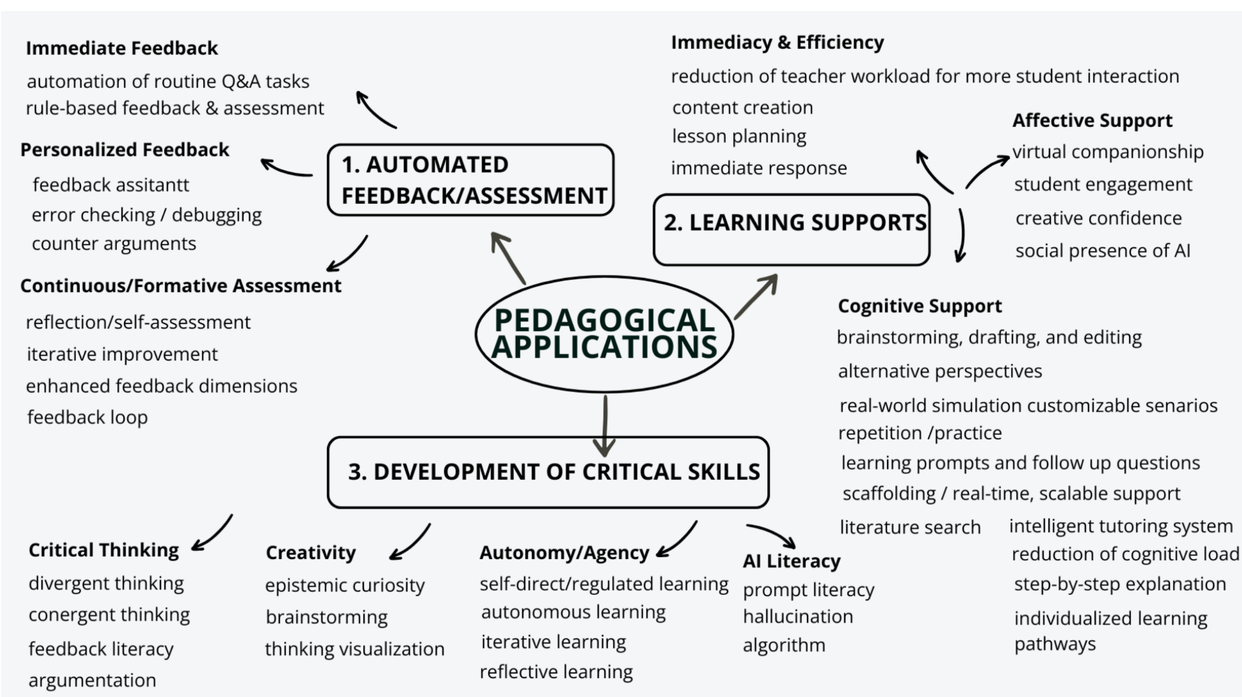


Fig. 5 Themes and codes of pedagogical applications

ensure mutual exclusivity and collective exhaustiveness by categorizing 42 codes based on their primary instructional purpose. Specifically, automated feedback/assessment focuses on efficiency-driven tasks (e.g., rule-based grading, immediate responses) that reduce educator workload. Learning supports encompass scaffolded interventions (e.g., cognitive aids, affective tools like virtual companionship) that directly assist learners in real time. Development of critical skills prioritizes long-term competencies (e.g., critical thinking, creativity, reflective learning) that transcend immediate task support, fostering metacognitive growth and learner autonomy.

While studies like Tang et al. (2024) explored overlapping applications (e.g., using ChatGPT for peer feedback analysis, automated feedback generation, and learning engagement), the thematic framework distinguishes between operational functions (automation), interventional supports (scaffolding), and transformative outcomes (skill development). For instance, reflective learning was categorized under critical skills rather than learning supports because its purpose is to cultivate self-regulated agency—a higher-order competency—rather than provide transient assistance. This tripartite structure clarifies how GenAI's roles align with distinct pedagogical objectives, even when applied concurrently.

Automated Feedback and Assessment

Automated feedback and assessment are a cornerstone of GenAI's pedagogical utility. The nine codes related to these functions can be grouped into three key themes*: immediacy, personalization, and continuous or formative evaluation*. For example, Collins et al. (2024) demonstrated the power of this approach with AnatomyGPT, a customized AI tool for anatomical sciences education. AnatomyGPT assesses students'mastery using National Board of Medical Examiners sample items, and provides immediate, detailed feedback complete with rationales and citations. Similarly, Hudon et al. (2024) showed that ChatGPT can generate medical education assessments comparable to expert-designed Script Concordance Tests, highlighting the tool's reliability in creating valid evaluation instruments. Tang et al. (2024) extended these capabilities by harnessing ChatGPT to evaluate critical thinking skills in online peer feedback, automating the assessment process, and delivering immediate insights that guide student improvement. Meanwhile, Txirides et al. (Tzirides et al., 2024) employed GenAI to offer rapid feedback on writing tasks, enabling students to iteratively refine their work in real time. Additionally, Beddington et al. (2024) and Du et al. (2024) used ChatGPT to generate formative feedback, such as executive summaries and counterarguments, that students then use to polish their compositions. These studies illustrate how GenAI not only streamlines the assessment process but also enriches learning by delivering timely, personalized feedback. This dual role helps reduce the burden on instructors while enhancing student engagement and supporting continuous improvement.

Learning Support

Learning support is the second foundational application of GenAI. The 19 codes related to learning support can be grouped into three types: *immediacy/efficiency, affective support, and cognitive support.* Immediacy and efficiency are key advantages of GenAI due to its ability to provide rapid, targeted assistance. For example, Allen et al. (2024) demonstrated this with Q-Module-Bot—a Q&A bot tailored for biology education that delivers immediate, module-specific responses to student queries. Similarly, Zhong and Kim (2024) extended this approach to business education by using ChatGPT to generate R code that simplifies logistic regression for students with limited programming skills. Van den Berg and du Plessis (2023) further amplified this support by employing ChatGPT to create lesson plans, worksheets, and visual presentations for teacher education.

Beyond efficiency, GenAI tools provide significant affective support by fostering an encouraging and patient learning environment. Hu et al. (2024) illustrated this with an AI Learning Companion that offers tailored assistance, helping students navigate course material independently. Studies also reveal that the social presence of GenAI tools, being consistently encouraging and patient, enhances students' overall learning experiences. GenAI also plays a crucial role in cognitive support, acting as a scaffold for brainstorming, practice, and real-world scenario applications. For instance, Yilmaz and Yilmaz (2023) showcased how ChatGPT assists in programming education by providing clear code examples and detailed explanations, which reinforce key computational concepts and facilitate deeper understanding. Together, these examples underscore GenAI's role as a versatile scaffold—bridging gaps in understanding and empowering students with instant, context-specific support across various dimensions of learning.

Development of Critical Skills

The third pedagogical application of GenAI—development of critical skills—encompasses 19 codes targeting higher-order competencies, including creativity, critical thinking, learner autonomy, and emergent prompt literacy. This theme is distinguished by its focus on metacognitive growth and adaptive skill-building, rather than immediate task support. For example, codes such as"divergent thinking"and"reflective learning"were grouped here because they prioritize fostering intellectual independence and problem-solving agility, while"prompt literacy"reflects the growing necessity for learners to strategically engage with AI as a collaborator. Unlike transactional supports (e.g., cognitive aids), these codes emphasize foundational capacities that enable learners to navigate complex, evolving educational landscapes, aligning with GenAI's transformative potential to cultivate lifelong, self-directed learners.

Creativity flourishes as GenAI inspires novel outputs across disciplines. Bedington et al. (2024) demonstrated this in a professional writing course where ChatGPT generates social media posts and summaries, prompting students to refine drafts into polished work. Tsao and Nogues (2024) extended this to creative writing, integrating multimodal tools like Midjourney and Stable Diffusion with ChatGPT to support storytelling and graphic narratives. Similarly, Essel et al. (2024) highlighted AI-assisted brainstorming in research methods courses, enabling students to generate original research ideas. Huang et al. (2024) leveraged ChatGPT's generative capabilities to enhance creative ideation in a product design course. Muengsan and Chatwattana (2024) further illustrated this in game-based learning, where GenAI designed educational content, fostering imaginative engagement. These examples showcase GenAI's role as a creative scaffold.

Critical thinking is strengthened as students analyze and refine GenAI outputs. Tzirides et al. (2024) demonstrated this through students evaluating AI-generated feedback for accuracy and bias, enhancing analytical skills. In teacher education, Van den Berg and du Plessis (2023) had instructors critique ChatGPT-generated lesson plans, such as an ESL prepositions lesson, adapting them to specific classroom needs, promoting reflective judgment. Tang et al. (2024) used ChatGPT to assess critical thinking in peer feedback, requiring students to scrutinize AI evaluations. Pinochet et al. (2023) took this further, encouraging ethical critiques of AI in collaborative tasks and fostering deeper critical perspectives and active evaluation.

Learning autonomy is supported by GenAI's immediacy and efficiency, reducing instructor dependence. Yang et al. (2024) enhanced student agency in a postgraduate course, where learners independently explored tasks using chat logs and journals. Students who took an active role in their learning (e.g., resourceful and reflective approaches) were more likely to benefit from GenAI, while those who were passive/receptive or resistant may not fully leverage its potential. Hu et al. (2024) introduced an AI Learning Companion that provides tailored, autonomous support, while Allen et al. (2024) presented Q-Module-Bot, enabling students to seek answers independently. Guo et al. (2024) offered EFL students AI-generated exercises for self-paced learning. These studies have underscored GenAI's potential to empower learning autonomy and learner agency.

Prompt literacy is emerging as a critical skill in the AI era, as students increasingly refine their ability to craft effective queries for AI systems. Yilmaz and Yilmaz (2023) underscored its importance in programming, demonstrating how precise prompts enable ChatGPT to generate accurate code. Similarly, Guo et al. (2024) explored its application in EFL instruction, where tailored prompts enhance language learning exercises. Bedington et al. (2024) further advanced the concept by introducing the "rhetoric of prompting," a framework for refining AI-generated outputs. Meanwhile, Zhong and Kim (2024) highlighted its role in R code generation, showing how prompt literacy can deepen analytics education.

Although these critical skills have been studied individually, future research should examine their interconnections to maximize GenAI's impact. Creativity sparks ideas, critical thinking refines them, autonomy drives exploration, and prompt literacy unlocks GenAI's full potential—together fostering a more cohesive and transformative skill development approach.

The 42 codes, which coalesced into three overarching themes, have been central to pedagogical adoptions of GenAI in its initial two years of use within teaching and learning. They draw on GenAI's core strengths—natural language processing, adaptability, and scalability (Kasneci et al., 2023), while simultaneously addressing key institutional priorities such as improving efficiency, enhancing student engagement, and fostering skill development. This synergy positions GenAI as an essential tool in the evolving landscape of higher education (Partnership for 21 st Century Learning, 2019). In addition, these applications align with broader educational imperatives frequently discussed in scholarly literature.

Research Question Three: What Challenges Emerge from Current GenAI Implementations in Higher Education Teaching and Learning?

Through a thematic analysis of challenges reported in 37 empirical studies, this section explores critical barriers to integrating GenAI into higher education practices. Twenty-five distinct codes were grouped into four areas: **(1) Technical, Usability, and Scalability; (2) Quality and Ethical Concerns; (3) Pedagogical Challenges; and (4) AI Literacy and Dependency.** See Fig. 6 for the themes and their codes.

Technical, Usability, and Scalability Challenges

The most commonly cited challenges in integrating GenAI into teaching and learning revolve around technical limitations, usability barriers, and scalability constraints. These challenges are well-documented in empirical studies. Technical issues, such as unreliable query handling and system glitches, affect tools like Q-Module-Bot (Allen et al., 2024) and Fermat (Cummings et al., 2024), while self-made RAG chatbots often generate inaccurate content (Guo et al., 2024). Similarly, the AI Learning Companion's dependence on ChatGPT's inconsistent accuracy

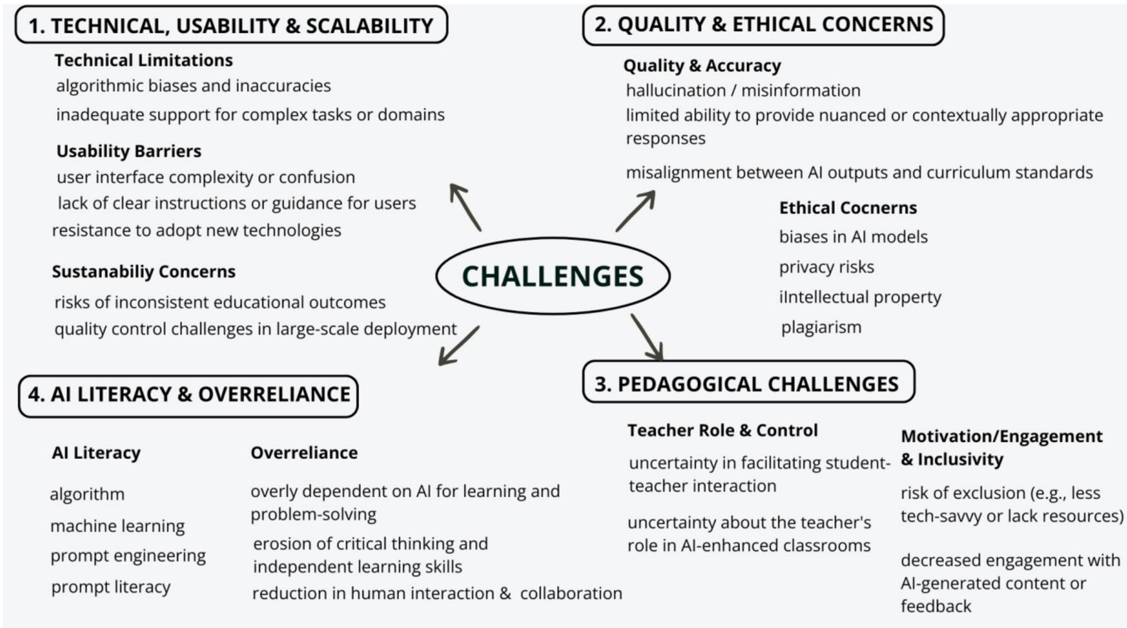


Fig. 6 Themes and codes of challenges in GenAI implementation

(Hu et al., 2024) underscores the need for more robust technical infrastructure. Usability barriers further hinder adoption, with 50% of users finding Q-Module-Bot's interface non-intuitive (Allen et al., 2024), while Fermat's confusing spatial canvas frustrates students (Cummings et al., 2024). Additionally, students struggle with the steep learning curve of crafting effective prompts (Guo et al., 2024), emphasizing the need for user-friendly design and better training (Hu et al., 2024). Sustainability remains another major concern, particularly regarding scalability and resource constraints. The Q-Module-Bot's reliance on local CSV storage limits its expansion (Allen et al., 2024), and highly customized RAG chatbots restrict broader applications (Guo et al., 2024). Furthermore, resource integration challenges seen in the AI Learning Companion (Hu et al., 2024) and tools like Elicit and Wordtune (Cummings et al., 2024) highlighted the need for scalable, resource-efficient solutions to support GenAI's long-term viability in education.

These issues mirror the early adoption phases of previous technologies, such as classroom computers and virtual learning environments, where initial glitches, user unfamiliarity, and scalability constraints required iterative refinement and investment (Rogers, 2003). Addressing these challenges demands a multifaceted approach, including robust technical development, user-centered design, and scalable infrastructure (Sillitoe, 2017). Without such efforts, the potential of GenAI to transform education will remain slow to be realized.

Quality and Ethical Concerns

Concerns about GenAI's reliability—particularly misinformation, bias, and unreliable feedback—persist across recent research. Escalante et al. (2024) found that GPT-4 occasionally generates inaccurate or harmful content, emphasizing the need for stringent oversight. Farazouli et al. (2024) highlighted ChatGPT's tendency toward repetitive and incoherent responses, undermining its educational utility. Van den Berg and du Plessis (2023) linked inaccuracies to outdated training data and the absence of real-time information, while Hudon et al. (2024) stressed the need for AI-generated medical content to align with expert-validated knowledge, especially in fields demanding high precision.

Ethical challenges—including bias, privacy risks, and academic integrity threats—further complicate GenAI's adoption. Bedington et al. (2024) and Habib et al. (2024) documented biases in ChatGPT that misrepresented human experiences or fabricated information, requiring users to evaluate outputs critically. Escalante et al. (2024) exposed weaknesses in safeguards against misuse, while Van den Berg and du Plessis (2023) raised concerns about copyright infringement and plagiarism in AI-generated texts. Smerdon (2024) and Waluyo and Kusumastuti (2024) pointed to academic integrity risks, noting that AI enables students to outsource assignments, challenging traditional notions of authorship and intellectual ownership.

Addressing these challenges requires comprehensive training for educators and students to recognize and mitigate

GenAI's limitations, alongside strong ethical frameworks to guide responsible use (Bond et al., 2024; Francis et al., 2025; Yan et al., 2024). The U.S. National Institute of Standards and Technology's AI Risk Management Framework (2023) provides structured approaches for detecting and correcting biases and inaccuracies, offering a model for educational applications. Similarly, the U.S. Department of Education (2023) emphasizes equitable implementation and ongoing professional development to ensure GenAI enhances learning without exacerbating disparities or enabling misconduct. While ethical concerns have long been part of educational technology, their scale and complexity have expanded with GenAI, making proactive measures more critical than ever.

Pedagogical Challenges

As discussed earlier, automated feedback and assessment are the primary pedagogical applications, but they also present significant challenges. Bedington et al. (2024) observed that while ChatGPT could summarize drafts, it missed critical points students deemed essential, necessitating human revision to ensure feedback addressed rhetorical intent. Escalante et al. (2024) suggested that while AI can generate feedback on writing, its output is not inherently reliable and requires scrutiny to ensure accuracy and appropriateness in an educational context. These findings suggest that while GenAI provides accessible, scalable feedback, studies suggest that human oversight is necessary to ensure nuanced understanding and support. This balance is particularly crucial in areas requiring ethical judgment or deep critical analysis, as noted in Farazouli's (2024) work on AI in teacher assessment practices.

Cheating risks and assessment validity were recurring concerns. Pinochet et al. (2023) warned of students outsourcing assignments to ChatGPT, threatening traditional evaluation methods. Bedington et al. (2024) critiqued AI detection tools for generating false positives and unfairly penalizing students. Waluyo and Kusumastuti (2024) reported difficulties in assessing originality in AI-assisted English writing as students might use AI as "an easy way to finish assignments quickly," bypassing deep engagement, a concern echoed by teachers' call for "judicious" use to maintain integrity (p. 8). Such issues demand rethinking assessment design, feedback, and academic integrity policies.

These challenges—AI feedback's lack of depth, obscured assessment of student work, and the need for human oversight—underscore a critical pedagogical tension: GenAI's efficiency can enhance scalability but risks diluting the personalized, critical engagement central to higher education. Addressing this requires innovative assessment strategies and instructor training to integrate AI effectively while preserving educational rigor.

Student AI Literacy and Dependency

A lack of AI literacy among students, coupled with overreliance, emerged as a significant challenge to effectively leveraging GenAI in higher education. Knoth et al. (2024) emphasized that non-expert users—those without formal AI training—struggled with unsystematic and trial-and-error approaches to crafting prompts for large language models, often overgeneralizing expectations from human interactions, which hampered their ability to elicit high-quality outputs. Similarly, Song et al. (2024) addressed this literacy gap in their development of LearningverseVR, noting the difficulty of prompt writing and the risk of"prompt word attacks"that could destabilize AI responses, necessitating a nested prompt engineering design to reduce user burden and prevent misuse (p. 4). Together, these studies underscore how inadequate AI literacy undermines GenAI's educational potential, requiring structured support to enhance prompt engineering proficiency among students and educators.

Compounding this literacy deficit, overreliance on GenAI threatens potentially the development of foundational skills, particularly in critical and creative domains. Gao et al. (2024) observed that business students overly dependent on ChatGPT saw their independent problem-solving abilities diminish, a pattern echoed in programming education by Groothuijsen et al. (2024), where AI support reduced peer collaboration and weakened students' capacity to tackle challenges independently. Xie et al. (2024) found that excessive AI interaction reduced social presence and learning autonomy, while Yilmaz and Yilmaz (2023) found AI tools did not mitigate motivation gaps in programming courses, leaving less active students sidelined in group tasks. Faculty concerns amplify these findings. Essel et al. (2024) reported adverse cognitive effects from students outsourcing critical thinking, and Habib et al. (2024), alongside Tsao and Nogues (2024), warned that AI's generic outputs might foster cognitive fixation, stifling creative agency.

Collectively, these studies highlight a tension between GenAI's efficiency and the risk of overreliance, which can undermine deep learning if not thoughtfully managed. While GenAI offers rapid support, unchecked dependence may diminish critical thinking, creativity, and autonomy—core pillars of higher education. Addressing this challenge requires more than technical solutions; it demands a pedagogical shift. The National AI Initiative (2021), a U.S. policy framework, advocates embedding AI literacy into curricula, suggesting that courses on prompt design and AI limitations could help mitigate overreliance. Similarly, EDUCAUSE (2023) underscores the importance of faculty development programs that model balanced AI integration, ensuring tools like ChatGPT serve as supplements rather than substitutes for intellectual effort. Without such strategies, GenAI's potential may be compromised, fostering dependency instead of enhancing the very skills it aims to support.

Implications

The findings of this systematic review reveal three central insights that directly inform the path forward for GenAI in higher education: (1) a predominance of text-centric tools like ChatGPT has limited exploration of multimodal and discipline-specific GenAI applications; (2) while GenAI's efficiencies (e.g., automated feedback, scalability) enhance pedagogy, they risk fostering overreliance that undermines critical thinking and creativity; and (3) institutional strategies to balance AI's utility with ethical safeguards remain underdeveloped. Building on these insights, this section advances two targeted implications to address these gaps and tensions, ensuring GenAI integration aligns with both pedagogical innovation and core educational values.

Leveraging Advances in Multimodal GenAI

Unlike traditional multimedia tools that rely on static, predesigned content, multimodal GenAI can dynamically generate personalized explanations, visualizations, and interactive elements in real time. One example of this is Google's Notebook LM, which creates personalized study guides by combining text, visual diagrams, and audio explanations. This offers tailored support for diverse learning styles, making learning more engaging and accessible. Additionally, such tools can assist educators by automating tasks such as lesson planning, quiz generation, and multimedia presentations, saving time while ensuring that content remains aligned with curriculum standards (Google Notebook, n.d.). Although still in its early stages, multimodal GenAI is poised to play a significant role in education by 2025 and beyond (John, 2023; Nayak, 2025).

Despite its promise, the impact of multimodal GenAI in education is currently constrained by the text-centric nature of most curricula and assessments. Richard Mayer's Cognitive Theory of Multimedia Learning (2005) suggests that integrating verbal and visual information enhances learning by reducing cognitive overload and improving understanding. However, many current GenAI tools focus primarily on text-based tasks, often overlooking the significant benefits of multimodal learning. To fully leverage this technology, educational institutions should reconsider their text-based assessment strategies and curricular designs. Multimodal approaches not only have the potential to deepen students'understanding but also offer more effective ways to assess it. For example, a medical school could use multimodal GenAI tools to create interactive study guides that combine 3D anatomical models, narrated explanations, and text-based descriptions. This would allow students to explore human anatomy from multiple perspectives, enhancing both understanding and learning retention (Chheang et al., 2024).

Promoting Teacher-Student-GenAI Collaboration

Currently, the pedagogical applications of GenAI primarily reflect traditional instructional approaches and skills, heavily relying on pre-AI educational models. Most of these applications focus on automating existing teaching tasks (e.g., writing, language learning) and pedagogical goals (e.g., feedback/assessment, practice). To realize the transformative potential of GenAI, educators must reimagine the role of AI, transitioning from mere automation to meaningful learner-teacher-GenAI interaction and collaboration. However, this shift has not yet been extensively explored in existing research.

Unlike traditional AI-driven educational tools, which primarily emphasize efficiency and task completion, GenAI facilitates dynamic, co-constructive learning experiences. For instance, a study conducted at the University of Pennsylvania's Wharton School of Business illustrates how ChatGPT was used in debate coaching, actively engaging students through iterative argumentation. In this process, students articulated arguments, received immediate AI-generated critiques, and refined their positions based on feedback from both the AI system and instructors, fostering deeper engagement, critical thinking, and active participation (Mollick & Mollick, 2023; Tegmark, 2017). Similarly, Liu et al. (2024) found that students who performed best using GenAI tools adhered to a structured interaction framework, effectively leveraging GenAI-generated feedback to enhance their understanding. This structured interaction demonstrates that GenAI can facilitate not only individual learning but also meaningful dialogue and reflection, reinforcing deeper engagement and higher-order thinking skills.

In the era of GenAI, education must reconsider existing interaction models to include GenAI, thereby forming a teacher-student-AI triadic relationship explicitly. This collaborative model can help address the current challenges of low AI literacy and high AI dependency by ensuring that students develop a nuanced understanding of AI's capabilities while avoiding overreliance on automated systems. Historically, classroom teachers have mediated learning interactions, and their role becomes even more critical within this new three-way collaboration (Li et al., 2024). Early efforts are already pioneering this exploration, with structured interaction models emerging in disciplines such as business (Mollick & Mollick, 2023) and teacher education (Liu et al., 2024). These studies underscore the pivotal role of educators (i.e., human-in-the-loop) in orchestrating and guiding these complex interactions when treating AI as a collaborative partner (Mollick & Mollick, 2023). Moving forward, it is essential that future research builds on these foundations to further explore and refine the dynamics of this triadic relationship.

Limitations

This review has several limitations that contextualize its findings. First, the geographical distribution of the analyzed studies reveals a pronounced concentration in China (including its Special Administrative Region, Hong Kong) and the United States. This regional imbalance limits the generalizability of findings, as cultural and institutional contexts, such as curriculum design priorities, technological infrastructure, or ethical frameworks, may uniquely shape GenAI adoption across regions. Second, while excluding 121 survey-based studies (46% of the initial pool) allowed this review to maintain a strict focus on pedagogical applications, these excluded studies represent a critical complementary research avenue. A dedicated systematic review of surveys exploring stakeholder attitudes, perceptions, and experiences could illuminate barriers to adoption, ethical dilemmas, and human-AI collaboration dynamics, enriching this review's findings with qualitative insights into implementation challenges. Finally, excluding non-peer-reviewed sources (e.g., conference proceedings, dissertations, preprints) and conceptual/theoretical articles ensured methodological rigor but risks omitting emerging innovations or frameworks not yet validated through peer review. For example, preliminary findings on novel GenAI tools often debut in conferences or preprints. This trade-off between depth and breadth underscores the need for iterative updates to capture the rapidly evolving GenAI landscape.

Conclusion

This study contributes to the growing body of research on GenAI in higher education by systematically reviewing peer-reviewed empirical studies that focus exclusively on direct teaching and learning applications. By excluding studies on perceptions and attitudes, this review provides a focused synthesis of evidence-based practices. Additionally, by analyzing studies published in 2023 and 2024, this review uniquely captures the initial academic response to ChatGPT's release in November 2022, offering insights into how higher education has adapted to this transformative technology in its first two years.

The findings reveal that while some recurring themes from the pre-ChatGPT era persist—such as the use of AI for writing support, English language learning, automated feedback, and assessment automation—new and emerging themes have also surfaced. Notably, GenAI is now being explored as a tool for fostering critical skills such as learning autonomy and prompt literacy, reflecting a significant shift in how AI is integrated into pedagogy. This shift may also explain why education has emerged as a leading field in AI adoption within higher education. Furthermore, while this study confirms previously identified challenges, such as concerns **over** quality, hallucination, bias, and ethical considerations, it also highlights AI dependency, specifically the tendency of students to outsource cognitive effort to AI tools. As AI becomes increasingly embedded in learning environments, concerns about over-reliance and its potential impact on cognitive and metacognitive development warrant further exploration.

Given these findings, this study emphasizes the forward-thinking approach of teacher-student-GenAI collaboration as a promising direction for future research. Rather than positioning AI as a standalone tool or a substitute for human instruction, a triadic interaction model—in which teachers, students, and AI interact dynamically—should be explored in this field. As higher education continues to integrate GenAI, this study provides foundational yet valuable insights for researchers, educators, and policymakers, guiding the innovative, pedagogically sound, and ethical application of AI in education.

Acknowledgements The author expresses sincere appreciation to the editor and anonymous reviewers for their incisive critiques and meticulous review. Their constructive feedback and insightful suggestions significantly strengthened the rigor and clarity of this work.

Funding The author did not receive funding to conduct this research.

Data Availability The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics Approval N/A.

Informed Consent This is a systematic literature review study. No ethical approval or informed consent is needed.

Competing interests The author declares no competing interests.

References

Allen, M., Naeem, U., & Gill, S. S. (2024). Q-Module-Bot: A generative AI-based question and answer bot for module teaching support. *IEEE Transactions on Education, 67*(5), 793–802. https://doi.org/10.1109/TE.2024.3435427

Bedington, A., Halcomb, E. F., McKee, H. A., Sargent, T., & Smith, A. (2024). Writing with generative AI and human-machine teaming: Insights and recommendations from faculty and students. *Computers and Composition, 71*, 102833. https://doi.org/10.1016/j.compcom.2024.102833

Bender, E. M. (2021). On the dangers of Stochastic Parrots: Can language models be too big? FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (pp. 610 – 623). https://doi.org/10.1145/3442188.3445922

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education, 21*(4), 4. https://doi.org/10.1186/s41239-023-00436-z

Carbonell, J. R. (1970). AI in CAI: An artificial-intelligence approach to computer-assisted instruction. *IEEE Transactions on Man-Machine Systems, 11*(4), 190–202. https://doi.org/10.1109/TMMS.1970.299942

Chheang, V., Sharmin, S., Marquez-Hern, R., Patel., M., Rajasekaran, D, & Caulfield., G. (2024). Towards anatomy education with generative AI-based virtual assistants in immersive virtual reality environments. Retrieved February 1, 2025, from https://arxiv.org/pdf/2306.17278

Collins, B. R., Black, E. W., & Rarey, K. E. (2024). Introducing AnatomyGPT: A customized artificial intelligence application for anatomical sciences education. *Clinical Anatomy, 37*(6), 661–669. https://doi.org/10.1002/ca.24178

Corbin, J., & Strauss, A. (2014). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Sage.

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education, 20*, 22. https://doi.org/10.1186/s41239-023-00392-8

Crompton, H., & Burke, D. (2024). The educational affordances and challenges of ChatGPT: State of the field. *TechTrends, 68*, 380–392. https://doi.org/10.1007/s11528-024-00939-0

Cummings, C., Monroe, S., & Watkins, M. (2024). Generative AI in first-year writing: An early analysis of affordances, limitations, and a framework for the future. *Computers and Composition, 71*, 102827. https://doi.org/10.1016/j.compcom.2024.102827

Du, H., Jia, Q., Gehringer, E., & Wang, X. (2024). Harnessing large language models to auto-evaluate the student project reports. *Computers and Education: Artificial Intelligence, 7*, 100268. https://doi.org/10.1016/j.caeai.2024.100268

EDUCAUSE. (2023). EDUCAUSE quick poll results: Adopting and adapting to generative AI in higher educational technology. Retrieved January 5, 2025, from https://er.educause.edu/articles/2023/4/educause-quickpoll-results-adopting-and-adapting-to-generative-ai-in-higher-ed-tech

Escalante, J., Pack, A., & Barrett, A. (2024). AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education, 20*, 57. https://doi.org/10.1186/s41239-023-00425-2

Essel, H. B., Vlachopoulos, D., Essuman, A. B., & Amankwa, J. O. (2024). ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models (LLMs). *Computers and Education: Artificial Intelligence, 6*, 100198. https://doi.org/10.1016/j.caeai.2023.100198

Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K., & McGrath, C. (2024). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices. *Assessment & Evaluation in Higher Education, 49*(3), 363–375. https://doi.org/10.1080/02602938.2023.2241676

Francis, N. J., Jones, S., & Smith, D. P. (2025). Generative AI in higher education: Balancing innovation and integrity. *British Journal of Biomedical Science, 81*, 14048. https://doi.org/10.3389/bjbs.2024.14048

Gao, Z. Y., Cheah, J. H., Lim, X. J., & Luo, X. (2024). Enhancing the academic performance of business students using generative AI: An interactive-constructive-active-passive (ICAP) self-determination perspective. *The International Journal of Management Education, 22*(2), 100958. https://doi.org/10.1016/j.ijme.2024.100958

Google Notebook LM. (n.d.). Think smarter, not harder. Retrieved December 25, 2024, from https://notebooklm.google/

Groothuijsen, S., van den Beemt, A., Remmers, J. C., & van Meeuwen, L. W. (2024). AI chatbots in programming education: Students' use in a scientific computing course and consequences for learning. *Computers and Education: Artificial Intelligence, 7*, 100290. https://doi.org/10.1016/j.caeai.2024.100290

Guo, K., & Li, D. (2024). Understanding EFL students' use of self-made AI chatbots as personalized writing assistance tools: A mixed methods study. *System, 124*, 103362. https://doi.org/10.1016/j.system.2024.103362

Guo, K., Pan, M., Li, Y., & Lai, C. (2024). Effects of an AI-supported approach to peer feedback on university EFL students' feedback quality and writing ability. *The Internet and Higher Education, 63*, 100962. https://doi.org/10.1016/j.iheduc.2024.100962

Habib, S., Vogel, T., Xiao, A., & Thorne, A. (2024). How does generative artificial intelligence impact student creativity? *Journal of Creativity, 34*(1), 100072. https://doi.org/10.1016/j.yjoc.2023.100072

Hu, Y., Hsieh, C., & Salac, E. (2024). Advancing freshman skills in information literacy and self-regulation: The role of AI learning companions and Mandala Chart in academic libraries. *The Journal of Academic Librarianship, 50*(3), 102885. https://doi.org/10.1016/j.acalib.2024.102885

Huang, H., Liu, Y., Dong, M., & Lu, C. (2024). Integrating AIGC into product design ideation teaching: An empirical study on self-efficacy and learning outcomes. *Learning and Instruction, 92*, 101929. https://doi.org/10.1016/j.learninstruc.2024.101929

Hudon, A., Kiepura, B., Pelletier, M., & Phan, V. (2024). Using ChatGPT in psychiatry to design script concordance tests in undergraduate medical education: Mixed methods study. *JMIR Medical Education, 10*(1), e54067. https://doi.org/10.2196/54067

Hwang, G. J., & Chang, C. Y. (2023). ChatGPT for language learning: Opportunities and challenges. *Computers & Education, 201*, 104829. https://doi.org/10.1016/j.compedu.2023.104829

IEEE Spectrum. (2024). What Is Generative AI? Spectrum explains large language models, the transformer architecture, and how it all works. Retrieved February 1, 2025, from https://spectrum.ieee.org/what-is-generative-ai

John, D. (2023). Multi-Modal generative AI systems: Bridging text, vision and speech with advanced LLM Architectures. *International Journal of Science and Research Archive, 9*(2), 1044–1058. https://doi.org/10.30574/ijsra.2023.9.2.0619

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., … Kasneci, G. (2023). *ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education*. https://doi.org/10.35542/osf.io/5er8f

Knoth, N., Tolzin, A., Janson, A., & Leimeister, J. M. (2024). AI literacy and its implications for prompt engineering strategies.

Computers and Education: Artificial Intelligence, 6, 100225. https://doi.org/10.1016/j.caeai.2024.100225
Knox, J. (2023). *AI and education in China: Imagining the future, excavating the past*. Routledge.
Koh, E., Zhang, L., Lee, A. V. Y., & Wang, H. (2024). Revolutionizing word clouds for teaching and learning with generative artificial intelligence: Cases from China and Singapore. *IEEE Transactions on Learning Technologies, 17*, 1416–1427. https://doi.org/10.1109/TLT.2024.3385009
Lee, G., Shi, L., Latif, E., Gao, Y., Bewersdorff, A., Nyaaba, M., Guo, S., Wu, Z., Liu, Z., Wang, H., Mai, G., Liu, T., & Zhai, X. (2023). Multimodality of AI for education: Towards artificial general intelligence. Retrieved January 5, 2025, from https://arxiv.org/abs/2312.06037
Li, T., Ji, Y., & Zhan, Z. (2024). Expert or machine? Comparing the effect of pairing student teacher with in-service teacher and ChatGPT on their critical thinking, learning performance, and cognitive load in an integrated-STEM course. *Asia Pacific Journal of Education, 44*(1), 45–60. https://doi.org/10.1080/02188791.2024.2305163
Liu, J., Li, S., & Dong, Q. (2024). Collaboration with generative artificial intelligence: An exploratory study based on learning analytics. *Journal of Educational Computing Research*. https://doi.org/10.1177/07356331241242441
Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). Cambridge University Press.
McCormack, M. (2023). EDUCAUSE QuickPoll results: Adopting and adapting to generative AI in higher ed tech. Retrieved December 31, 2024, from https://er.educause.edu/articles/2023/4/educause-quickpoll-results-adopting-and-adapting-to-generative-ai-in-higher-ed-tech
McKinsey & Company. (2024). What is generative AI? Retrieved January 5, 2025, https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai
Medium. (2024). How Generative AI is revolutionizing film and music. Retrieved January 5, 2025, from https://medium.com/@researchgraph/how-generative-ai-is-revolutionising-film-and-music-2a622bd14fad
Meta. (2023a). Multimodal generative AI systems. Retrieved January 10, 2025, from https://ai.meta.com/tools/system-cards/multimodal-generative-ai-systems/
Meta. (2023b). Meta movie gen. Retrieved January 5, 2025, from https://ai.meta.com/research/movie-gen/
Milić, T., Tomić, B., Marinković, S., & Jeremić, V. (2024). ESPRIT adventure: Assessing hybrid fuzzy-crisp rule-based AI method effectiveness in teaching key performance indicators. *The International Journal of Management Education, 22*(3), 101022. https://doi.org/10.1016/j.ijme.2024.101022
MIT Management. (2024). AI image generation tools. Retrieved January 5, 2025, from https://mitsloanedtech.mit.edu/ai/tools/images/
MIT News. (2023). Explained: Generative AI. How do powerful generative AI systems like ChatGPT work, and what makes them different from other types of artificial intelligence? Retrieved January 10, 2025, from https://news.mit.edu/2023/explained-generative-ai-1109
Mohebi, L. (2024). Empowering learners with ChatGPT: Insights from a systematic literature exploration. *Discover Education, 3*, 36. https://doi.org/10.1007/s44217-024-00120
Mollick, E., & Mollick, L. (2023). *Co-intelligence: Living and working with AI*. Harvard Business Review Press.
Moorhouse, B. L., Wan, Y., Ho, T. Y., & Lin, A. M. (2024). Generative AI-assisted, evidence-informed use of L1 in L2 classrooms. *ELT Journal, 78*(4), 453–465.
Muengsan, T. S., & Chatwattana, P. (2024). The game-based learning (GbL) platform with Generative AI to enhance digital and technology literacy skills. *Higher Education Studies, 14*(1), 46–53. https://ccsenet.org/journal/index.php/hes/article/view/0/49704
Natale, S. (2019). If software is narrative: Joseph Weizenbaum, artificial intelligence and the biographies of ELIZA. *New Media & Society, 21*(3), 712–728.
Nayak, B. (2025). The evolution and architecture of multimodal AI systems. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 11*(1), 1007–1017. https://doi.org/10.32628/cseit251112108
Netland, T., von Dzengelevski, O., Tesch, K., & Kwasnitschka, D. (2025). Comparing human-made and AI-generated teaching videos: An experimental study on learning effects. *Computers & Education, 224*, Article 105074. https://doi.org/10.1016/j.compedu.2024.105074
Nguyen, A., Hong, Y., Dang, B., & Huang, X. (2024). Human-AI collaboration patterns in AI-assisted academic writing. *Studies in Higher Education, 49*(5), 847–864. https://doi.org/10.1080/03075079.2024.2323593
Ogunleye, B., Zakariyyah, K. I., Ajao, O., Olayinka, O., & Sharma, H. (2024). A systematic review of generative AI for teaching and learning practice. *Education Sciences, 14*(6), 636. https://doi.org/10.3390/educsci14060636
Owan, V. J., Abang, K. B., Idika, D. O., Etta, E. O., & Bassey, B. A. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *EURASIA Journal of Mathematics, Science and Technology Education, 19*(8), em2307. https://doi.org/10.29333/ejmste/13428
Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *Systematic Reviews, 10,* Article 89.https://doi.org/10.1186/s13643-021-01626-4
Partnership for 21st Century Learning. (2019). Framework for 21st century learning. Retrieved December 5, 2024, from http://www.p21.org
Pinochet, L. H. C., Moreira, M. A. L., Fávero, L. P., dos Santos, M., & Pardim, V. I. (2023). Collaborative work alternatives with ChatGPT based on evaluation criteria for its use in higher education: Application of the PROMETHEE-SAPEVO-M1 method. *Procedia Computer Science, 221*, 107970. https://doi.org/10.1016/j.procs.2023.07.025
Preiksaitis, C. M., & Rose, C. (2023). Opportunities and limitations of generative artificial intelligence in medical education: Scoping review. *JMIR Medical Education, 9*(1), e48785. https://doi.org/10.2196/48785
Prensky, M. (2012). From *digital natives to digital wisdom: Hopeful essays for 21st century learning*. Corwin Press.
Robert, J. (2024). 2024 EDUCAUSE AI landscape study. Retrieved January 5, 2025, from https://www.educause.edu/ecar/research-publications/2024/2024-educause-ai-landscape-study/policies-and-procedures
Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.
Selwyn, N. (2022). *Education and technology: Key issues and debates* (3rd ed.). Bloomsbury Academic.
Shortliffe, E. H., Davis, R., Axline, S. G., Buchanan, B. G., Green, C. C., & Cohen, S. N. (1975). Computer-based consultations in clinical therapeutics: Explanation and rule acquisition capabilities of the MYCIN system. *Computers and Biomedical Research, 8*(4), 303–320.
Sillitoe, J. (2017). Scaling educational technologies: A framework for integrating scalable infrastructure into higher education

systems. *Educational Technology Research and Development, 65*(4), 913–928.

Smerdon, D. (2024). AI in essay-based assessment: Student adoption, usage, and performance. *Computers and Education: Artificial Intelligence, 7*, 100288. https://doi.org/10.1016/j.caeai.2024.100288

Song, Y., Wu, K., & Ding, J. (2024). Developing an immersive game-based learning platform with generative artificial intelligence and virtual reality technologies – "LearningverseVR." *Computers and Education: X Reality, 4*, 100069. https://doi.org/10.1016/j.cexr.2024.100069

State Council of China. (2017). Next generation artificial intelligence development plan. Retrieved January 30, 2025, from https://www.beijing.gov.cn/fuwu/lqfw/gggs/202106/t20210609_2409514.html

Tang, T. Q., Sha, J. R., Zhao, Y. N., Wang, S. D., Wang, Z. B., & Shen, S. (2024). Unveiling the efficacy of ChatGPT in evaluating critical thinking skills through peer feedback analysis: Leveraging existing classification criteria. *Thinking Skills and Creativity, 53*, 101607. https://doi.org/10.1016/j.tsc.2024.101607

Tegmark, M. (2017). *Life 3.0: Being human in the age of artificial intelligence.* Alfred A. Knopf.

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology, 8*, 45. https://doi.org/10.1186/1471-2288-8-45

Tsao, J., & Nogues, C. (2024). Beyond the author: Artificial intelligence, creative writing and intellectual emancipation. *Poetics, 102*, 101865. https://doi.org/10.1016/j.poetic.2024.101865

Tzirides, A. O., Zapata, G., Kastania, N. P., Saini, A. K., Castro, V., Ismael, S. A., You, Y. L., dos Santos, T. A., Searsmith, D., O'Brien, C., Cope, B., & Kalantzis, M. (2024). Combining human and artificial intelligence for enhanced AI literacy in higher education. *Computers and Education Open, 6*, 100184. https://doi.org/10.1016/j.caeo.2024.100184

U.S. Department of Education. (2023). *Artificial intelligence and the future of teaching and learning: Insights and recommendations*. Retrieved January 5, 2025, from https://www2.ed.gov/documents/aireport/ai-report.pdf

Van den Berg, G., & du Plessis, E. (2023). ChatGPT and Generative AI: Possibilities for its contribution to lesson planning, critical thinking and openness in teacher education. *Education Science, 13*, 998. https://doi.org/10.3390/educsci13100998

Waluyo, B., & Kusumastuti, S. (2024). Generative AI in student English learning in Thai higher education: More engagement, better outcomes? *Social Sciences & Humanities Open, 10*, 101146. https://doi.org/10.1016/j.ssaho.2024.101146

Wang, N., Wang, X., & Su, Y. S. (2024). Critical analysis of the technological affordances, challenges and future directions of Generative AI in education: A systematic review. *Asia Pacific Journal of Education, 44*(1), 139–155. https://doi.org/10.1080/02188791.2024.2305156

Weizenbaum, J. (1966). ELIZA: A computer program for the study of natural language communication between man and machine. *Communications of the ACM, 9*(1), 36–45. https://doi.org/10.1145/365153.36516

Wiboolyasarin, W., Wiboolyasarin, K., Suwanwihok, K., Jinowat, N., & Muenjanchoey, R. (2024). Synergizing collaborative writing and AI feedback: An investigation into enhancing L2 writing proficiency in wiki-based environments. *Computers and Education: Artificial Intelligence, 6*, 100228. https://doi.org/10.1016/j.caeai.2024.100228

Xie, Z. H., Wu, X. Z., & Xie, Y. X. (2024). Can interaction with generative artificial intelligence enhance learning autonomy? A longitudinal study from comparative perspectives of virtual companionship and knowledge acquisition preferences. *Journal of Computer Assisted Learning, 40*(1), 49–64. https://doi.org/10.1111/jcal.12863

Yan, W. J., Hu, B., Liu, Y. L., Li, C. Y., & Song, C. L. (2024). Does usage scenario matter? Investigating user perceptions, attitude and support for policies towards ChatGPT. *Information Processing & Management, 61*(6), 103867. https://doi.org/10.1016/j.ipm.2024.103867

Yang, Y., Luo, J., Yang, M., Yang, R., & Chen, J. (2024). From surface to deep learning approaches with Generative AI in higher education: An analytical framework of student agency. *Studies in Higher Education, 49*(5), 817–830. https://doi.org/10.1080/03075079.2024.2327003

Yilmaz, R., & Yilmaz, F. (2023). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Computers and Education: Artificial Intelligence, 4*, 100147. https://doi.org/10.1016/j.caeai.2023.100147

Zawacki-Richter, O., Marín, V. I., & Bond, M. (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*, 39. https://doi.org/10.1186/s41239-019-0171-0

Zhong, C., & Kim, J. B. (2024). Teaching case: Teaching business students logistic regression in R with the aid of ChatGPT. *Journal of Information Systems Education, 35*(2), 138–143. https://doi.org/10.62273/DYLI2468