IJTE International Journal of Technology in Education
www.ijte.net

The Impact of GenAI Chatbots on Student Learning in Higher Education: A Literature Review

Merete Badger [1 *], Nuria López [2], Camilla Falk Rønne Nissen [3]

[1] Copenhagen Business School, Solbjerg Plads 3, 2000 Frederiksberg, Denmark, 0000-0002-2072-8619
[2] Copenhagen Business School, Solbjerg Plads 3, 2000 Frederiksberg, Denmark, 0009-0005-7123-2882
[3] Copenhagen Business School, Solbjerg Plads 3, 2000 Frederiksberg, Denmark, 0009-0007-6092-8731
Corresponding author: Merete Badger (mba.edq@cbs.dk)

Article Info

Article History

Received:
19 June 2025

Revised:
11 October 2025

Accepted:
17 November 2025

Published:
1 January 2026

Keywords

Artificial Intelligence
Generative AI
Chatbots
Student learning
Higher education

Abstract

The integration of generative AI (GenAI) chatbots in Higher Education raises questions about the impact of the new technology on student learning. This systematic literature review synthesizes findings from 49 empirical studies published between 2022 and 2024, focusing on how GenAI chatbots influence key dimensions of student learning, including motivation, engagement, self-efficacy, self-regulation, comprehension, critical thinking, problem-solving, and learning performance. Results indicate that GenAI chatbots can enhance learning by providing personalized support, immediate feedback, and opportunities for self-directed learning. However, concerns persist regarding over-reliance on AI, reduced critical thinking, and academic integrity. The review highlights that guided and pedagogically sound integration of GenAI tools is essential to maximize benefits and mitigate risks. These findings underscore the importance of developing AI literacy and ethical usage guidelines to support meaningful and equitable learning experiences in Higher Education.

Citation: Badger, M., López, N., & Nissen, C. F. R. (2026). The impact of GenAI chatbots on student learning in higher education: A literature review. *International Journal of Technology in Education (IJTE), 9*(1), 43-69. https://doi.org/10.46328/ijte.5111

ISSN: 2689-2758 / © International Journal of Technology in Education (IJTE).

Introduction

The future of Higher Education is in many ways linked with the developments of emerging new technologies and for most students, generative artificial intelligence (GenAI) will be an integrated part of their studies and their lives in general. In this review, we focus on how the access to an AI assistant may affect student learning in Higher Education. Tutoring has previously proven to be an extremely effective strategy for achieving high assessment scores in university and college education. Bloom (1984) found that the average student who received one-to-one tutoring performed two standard deviations better than the average student in a traditional classroom setting. Through the so-called "2 Sigma Problem", Bloom also acknowledged that one-to-one tutoring is resource demanding and therefore not always practical or scalable at educational institutions. In response to this challenge, the use of web-based intelligent tutoring systems and AI assistants has been explored in Higher Education to uncover the potential for catalysing student learning.

Intelligent tutoring systems are designed specifically to enhance learning experiences through personalized and adaptive instruction, whereas AI assistants are general-purpose tools providing responses based on user queries (Crompton & Burke, 2023). The launch of ChatGPT by OpenAI in the autumn of 2022 made AI assistants, based on large language models (LLM), available to anyone with access to a web browser. GenAI technology represented a leap forward in the AI assistant's capabilities of generating human-like text, code, images, and other types of output and a range of other GenAI chatbots have emerged since then (e.g. Copilot by Microsoft, Claude by Anthropic, and Gemini by Google). The impact of GenAI on Higher Education is being debated with a focus on both the challenges and opportunities it presents for teaching and learning. On the positive side, students can customize their learning experience according to their needs and preferences by engaging with AI-based learning tools. Concerns have also been expressed, particularly regarding issues of transparency and academic integrity, as GenAI is integrated in educational settings.

The amount of scientific literature on the use of GenAI chatbots in Higher Education is increasing dramatically over time. A vast number of conceptual publications about the anticipated implications of GenAI for Higher Education institutions and for society exist. Numerous studies apply the Technology Acceptance Model (Davis, 1985) to explore student's and educator's perceived usefulness and perceived ease of use of GenAI tools, and other studies test the performance of GenAI tools when it comes to solving assignments such as quizzes, coding exercises, or essay writing. Fewer publications are available on the effect of GenAI on student's learning outcomes as it takes time to collect and present such evidence. The purpose of this review is to establish an overview of early findings on the effect of student's use of GenAI chatbots on their learning outcomes in Higher Education. This is formulated through the following research question:

RQ: How does the use of GenAI chatbots impact student learning in Higher Education?

Previous Reviews on AI Chatbots and Student Learning

Zirar (2023) examined the impact of language models, particularly ChatGPT, on student learning and assessment. The review of the 25 selected articles yielded two overarching conclusions. First, student learning may be

negatively affected when there is exclusive reliance on language models without critical evaluation. Second, the role played by language models in assisting educators to produce teaching and assessment resources should be specific and controlled.

Bhullar et al. (2024) analysed the role of ChatGPT in teaching and learning within Higher Education institutions. Their review of 47 selected articles identified four key thematic clusters: academic integrity, learning environment, student engagement, and scholarly research. The findings indicated that research primarily focused on assessment practices, academic integrity, student learning and discipline-specific applications. Although the integration of AI in education could potentially enhance academic performance and student motivation, the authors emphasized that students' use of these technologies should be monitored to ensure responsible use.

Wu & Yu (2024) conducted a meta-analysis of 24 studies on the effect of AI chatbots on students' learning outcomes. They concluded that the use of AI chatbots had a positive effect on overall learning outcomes. Improvements were observed in relation to learning performance, motivation, self-efficacy, interest and perceived value of learning. These positive results could be attributed to the chatbots' capacity to deliver personalized support, provide continuous opportunities for practice, and create a relaxed learning environment – particularly beneficial for learners who may feel anxious about making mistakes in front of teachers or peers.

As part of their comprehensive SWOT (strengths, weaknesses, opportunities, threads) analysis on the use of ChatGPT in teaching and learning, Mai et al. (2024) evaluated the effect of ChatGPT integration in 51 selected studies. The authors identified 13 strengths (e.g., ChatGPT's capacity to act as a teaching assistant), 10 weaknesses (e.g., beginners' limitations to use the chatbot critically), 5 opportunities (e.g., personalised learning experience) and 4 threats (e.g., overreliance on AI-generated outputs) of using ChatGPT. Similar conclusions were drawn by Faisal (2024) in a review of 52 studies on the potential benefits of integrating ChatGPT in Higher Education in the Saudi Arabian context. The possibility of using ChatGPT to respond to individuals' learning needs was seen as particularly useful to support language proficiency in the case of Saudi students enrolled in courses taught in English. The risk of ChatGPT perpetuating biases from the training data, as well as the potential diminishment of students' critical thinking skills, were seen as key weaknesses of this technology.

Khurma et al. (2024) explored student engagement within ChatGPT interactions through their systematic review of 16 studies. Four dimensions of student engagement were investigated: academic, behavioural, cognitive and psychological engagement. The review identified several ways in which interacting with ChatGPT could enhance student engagement, including the provision of immediate feedback. A key finding was that student's academic engagement – such as engagement with the course material – is strengthened the most when using ChatGPT, compared to any of the other dimensions of engagement. However, the review also highlighted some potential negative effects of ChatGPT, such as the risk of students becoming over dependent on the tool.

Deng et al. (2025) conducted a meta-analysis of 69 experimental studies investigating the impact of ChatGPT on student learning. The data revealed that ChatGPT use could enhance academic performance and affective-motivational states, promote higher-order cognitive processes and reduce mental effort. However, it was found to

have no significant influence on students' self-efficacy. The authors caution that many studies did not clearly report whether the use of ChatGPT was permitted during post-intervention assessments of academic performance. As a result, it remains uncertain whether the observed gains reflect actual improvements in student learning or are primarily attributable to the higher quality of output generated by the tool itself.

Bruun et al. (2024) presented a research overview based on a literature review of 141 studies published between 2018 and 2024 on the use of AI chatbots in Higher Education, specifically within the humanities and social science. In relation to student learning, the authors noted a lack of theoretical frameworks explicitly analysing how AI-supported learning occurs, as well as a tendency in the research to focus on the role of AI to promote individual students' learning, without sufficient consideration to the use of this technology in the larger social context. The capacity of AI-chatbots to support self-regulated learning was highlighted in many of the reviewed studies. Moreover, the interactive nature of the technology, which can adopt various roles to support learning, makes it possible to personalise the instruction by identifying students' challenges and suggesting solutions. However, some studies caution that students' perceptions of the effectiveness of AI chatbots may be more variable than generally assumed.

Dimensions of Learning

To approach our research question in a systematic manner, this review will focus on eight of dimensions that refer to key aspects of meaningful student learning (see Figure 1). These dimensions emerged, among others, through the thematic analysis of the articles during the review process.

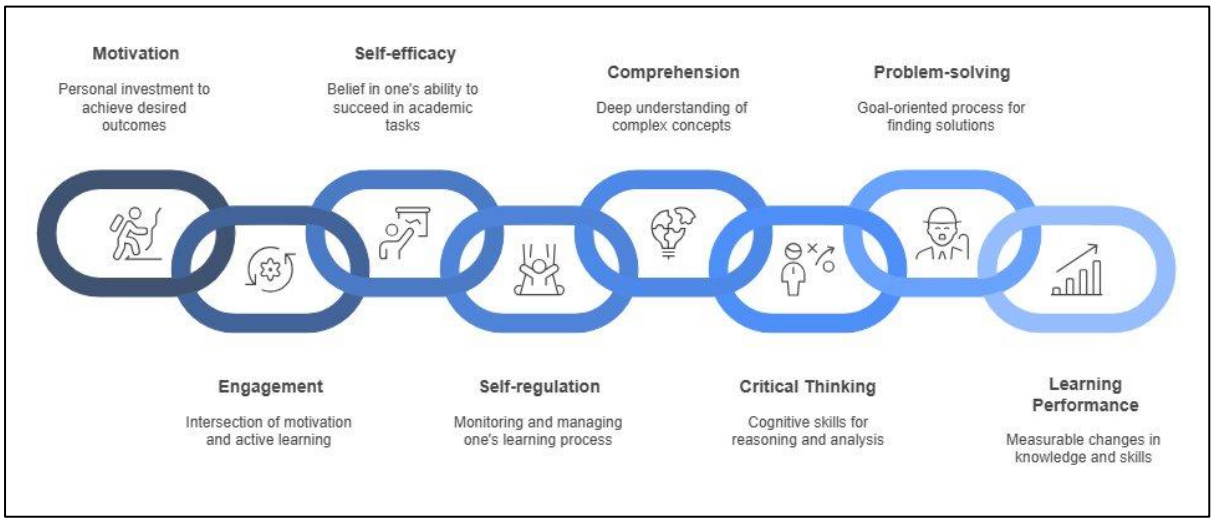

Figure 1. Dimensions of Student Learning Analysed in This Literature Review

The final selection of eight dimensions – encompassing affective, metacognitive, and cognitive aspects – was made to enable a multidimensional analysis of GenAI chatbots' impact on student learning. The selection ensures comprehensive coverage of key facets of learning. For instance, motivation is a central factor in the affective domain; self-regulation captures the key role of metacognitive skills in learning; and problem-solving represents one of the core cognitive skills. Moreover, the selected dimensions reflect the inherently interconnected nature of

learning domains, as affective, metacognitive, and cognitive processes intersect in multiple ways. As a result, clear overlaps emerge among the selected dimensions. For instance, while the comprehension of complex concepts may be defined as a cognitive process, it is nevertheless influenced by affective factors such as motivation. Similarly, self-efficacy impacts the effective application of problem-solving skills, and self-regulation supports and strengthens critical thinking. These eight dimensions form the analytical framework of this review and are described below.

Motivation

Motivation plays a key role in determining the behaviours that students adopt when learning, affecting the direction, intensity and quality of their personal investment to achieve a desired outcome (Lovett, 2023). For decades now, motivation has been explained through the lens of the Expectancy-Value Theory (EVT), which identifies the subjectively perceived value of a goal (value) and the expectations to successfully achieve it (expectancies) as the two main aspects that influence motivation (Rosenzweig et al., 2019). More recent research has suggested that the perceived negative consequences of completing a task (cost) can also have an impact on this aspect of learning (Flake et al., 2015). Students will be more motivated to pursue goals that have high value to them (e.g., sense of accomplishment from solving a complex task), imply a low cost (e.g. effort and time required seem reasonable), and that they believe they are capable of achieving (e.g., prior knowledge and experience seem relevant to attain the goal).

Engagement

Student engagement can be defined as the mental state where motivation (the feeling aspect of engagement) and active learning (the thinking aspect of engagement) intersect (Barkley & Major, 2020). Motivation and active learning interact synergistically to create engagement, leading to a series of learning behaviours that determine how students approach tasks, courses, or educational experiences in general. There is general agreement in educational research that more engagement leads to better academic achievement, manifested in improved acquisition of course content and retention (Casuso-Holgado et al., 2013; Lee, 2014). Learning behaviours that are associated with a high level of engagement are for example active participation and contribution to class activities and discussions, perseverance when dealing with challenges, and adequate time and effort spent on the tasks (Bowden et al., 2021). These behaviours are generally shaped by affective and social aspects such as students' well-being, self-efficacy, and sense of belonging.

Self-Efficacy

Self-efficacy, the degree to which students believe they are capable of succeeding in specific academic tasks, is a strong predictor of their performance and achievements. Self-efficacy beliefs affect the effort invested in learning as well as the level of resilience shown in difficult situations. They also influence students' perceptions of how difficult tasks really are. Those with a weak sense of self-efficacy are prone to approach tasks with a sense of helplessness, thinking that they are too complex for them to complete successfully. A strong sense of self-efficacy,

however, will help students approach even challenging tasks willingly and confidently (Van Dinther et al., 2011) Self-efficacy is bolstered when students experience a challenging situation that they have been able to overcome successfully (mastery experiences), when they observe peers succeeding in tasks like the ones they must carry out (vicarious experiences), and when they receive realistic verbal encouragement (Beatson et al., 2018).

Self-Regulation

Self-regulation involves monitoring one's learning, being aware of the emotions, behaviors, motivations and other factors that influence the learning process (Barkley & Major, 2020). Self-regulated learners are able to plan the most efficient way of approaching a task, evaluate their own progress and choose appropriate learning strategies. Self-regulated thoughts and behaviours are self-generated, guiding students to personally set goals, consider strengths and limitations, and decide task-related strategies that will help them become more effective learners. These self-regulatory abilities increase motivation and empower students to work in their optimal challenge zone, which subsequently makes them more likely to succeed academically (Zimmerman, 2002). Although self-regulation stems largely from self-efficacy beliefs and intrinsic motivation, research has also indicated that self-regulation can be taught through systematic interventions, such as providing guidance for students to practice self-regulation strategies like goal setting, self-assessment, and self-reflection (Bol & Garner, 2011).

Comprehension

Comprehension is a cognitive process by which students develop a deep understanding of complex concepts and ideas. This process involves connecting new information to prior knowledge, constructing meaning, and applying knowledge in a variety of situations. Acquiring deep comprehension of complex content or attaining mastery of complex tasks requires having sufficient and adequate prior knowledge and being competent in foundational sub-skills. Furthermore, it is essential to engage in targeted practice and be able to transfer knowledge and skills to novel contexts (Lovett, 2023). The cognitive load imposed by the tasks to be mastered is a critical factor to consider in supporting students' comprehension. In this regard, the Cognitive Load Theory (CLT) offers valuable insights into how to best design instruction for acquiring complex cognitive skills and competencies (e.g., addressing deficient prior knowledge, scaffolding practice with an appropriate gradual increase in task difficulty) (Kirschner, 2002).

Critical Thinking

Critical thinking includes a broad range of cognitive skills, such as reasoning, inference-making, analysis, evaluation, interpretation, and argumentation. Critical thinkers are characterized by being inquisitive, aware of their own biases, and committed to inquiry and the identification of relevant information (Davies, 2015). Since critical thinking entails examining issues from multiple perspectives and considering alternative viewpoints, it is a fundamental component of effective problem-solving and decision-making. The development of critical thinking skills is intrinsically linked to self-regulation, as abilities such as being able to self-monitor one's progress and making decisions about appropriate learning strategies are positively correlated with critical thinking (Akcaoğlu

et al., 2023). Some instructional strategies that foster students' critical thinking include the use of Socratic questioning to explore complex ideas and the implementation of peer and self-assessment techniques to evaluate the strengths and weaknesses of one's reasoning (Fahim & Masouleh, 2012).

Problem-Solving Skills

Problem solving is a goal-oriented process, which involves skills such as identifying the problem, organizing knowledge, thinking creatively about multiple solution paths, developing solution strategies, and monitoring one's progress toward the goal (Harlim & Belski, 2013). What specific skills are necessary to apply in problem-solving tasks will depend on the problem's structure and complexity. Students can encounter well-structured problems and ill-structured problems. The former type requires the implementation of a limited number of principles in an organized manner, whereas the latter tends to be more complex, calling for divergent solutions and the integration of knowledge from several domains (Jonassen, 2000). Metacognition and critical thinking are positively related to problem solving. The development of metacognitive abilities such as self-assessing the actions taken to achieve the goal and critical thinking competences such as making reasonable inferences based on evidence can enhance students' problem-solving performance (Prakong, 2024).

Learning Performance

Learning performance refers to the observable and measurable short-term changes in knowledge, skills and attitudes that occur during the teaching and learning process. It is commonly used to evaluate attainment of learning outcomes within courses or activities, although it may not necessarily reflect long-term retention or mastery (Soderstrom & Bjork, 2015). Certain factors are critical to ensure that learning performance is evaluated adequately. Assessment tools must be specifically designed for the learning goals that are to be evaluated, in order to yield useful evidence that can further inform meaningful changes in teaching and learning (Suskie, 2018). Furthermore, educational research also recognizes the importance of employing diverse assessment approaches to measure learning performance, as they offer complementary perspectives and illuminate different aspects of learning. For example, although self-reporting primarily reflects students' subjective experiences and attitudes rather than providing objective indicators of performance, it is widely acknowledged as a valid tool for capturing affective and metacognitive dimensions of learning, such as motivation and self-regulation (Caspersen et al., 2017; Pekruna, 2020).

Methods

Literature Search

This literature review follows the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines (Page et al., 2021). We first composed a set of keywords combining generative AI terms with the two terms 'student learning' or 'learning outcome'. Based on the keywords, we built search strings compatible with three academic research databases: Web of Science, Scopus, and the Education Resource Information Center (ERIC) under EBSCOhost. The search string for Scopus is shown in Table 1 as an example.

Table 1. The Search String Used to Search in Scopus

Data base	Search string
Scopus	((TITLE-ABS-KEY ((ai OR gai OR genai OR "generative ai" OR "generative artificial intelligence" OR llm* OR "large language model*" OR chatgpt OR "bing chat" OR copilot OR gemini OR chatbot*)) AND TITLE-ABS-KEY (student W/3 learning OR "learning outcome"))) AND TITLE-ABS-KEY ("higher education" OR HE OR universit* OR college) AND PUBYEAR > 2021 AND PUBYEAR < 2025

Additional filtering was applied to restrict the search to Higher Education and colleges and to set the period of interest to 2022-24 i.e. the era after ChatGPT was released. In ERIC, this filtering was performed via the web interface. All three data bases were searched on June 6, 2024 and this led to 1431 hits in total (813, 550, and 68, respectively) as shown in Figure 2.

Screening of Publications

The collection of references was imported in the tool Covidence, which we used for screening of abstracts and full texts and for extraction of relevant content. A total of 265 references were identified as duplicates and subtracted from the collection. The abstracts of the remaining 1165 references were then screened by one reviewer according to a set of inclusion/exclusion criteria (Table 2). Publications were included if they contained evidence of student's learning outcome, learning experience, strategies for learning, or perception of learning when using GenAI chatbots in connection with their Higher Education studies. Publications were excluded if they were not research publications, if they were conceptual or theoretical rather than evidence based, or if they were not addressing higher or college education. Further, publications were excluded if GenAI chatbots did not feature or if the publications were addressing other aspects than student learning, such as the design of chatbots, chatbot's ability to solve problems, student's acceptance of technology, or teacher's perspectives on using AI chatbots in the classroom.

The title and abstract screening reduced the collection to 210 publications. Full-text screening was then carried out by three reviewers where each publication was read by at least two reviewers. After the full-text screening, the collection consisted of 49 publications, which form the basis of our review. Of these publications, 41 were published in international journals, whereas the remaining eight publications were in conference proceedings.

Table 2. Inclusion and Exclusion Criteria Applied in the Screening

Inclusion criteria	Exclusion criteria
The publication is a research publication *(e.g. a journal or conference article, a book chapter*)	The publication is not a research publication *(e.g. a cover for conference proceedings)*
The study is about GenAI chatbots	The study is about chatbots that do not take

Inclusion criteria	Exclusion criteria
(e.g. ChatGPT, Copilot, Claude, Gemini)	advantage of GenAI technology
	The study is about AI-tools other than chatbots *(e.g. tools for writing, grammar checking, pronunciation, transcription, coding, music, visualisation of learning trajectories*)
The study takes place in a Higher Education setting *(e.g. university, college*)	The study does not take place in a Higher Education setting *(e.g. primary school, K-12, nursing school*)
The study is based on evidence *(e.g. based on survey responses, assessment outcomes)*	The study is conceptual or theoretical or addresses the wider implications of AI on Higher Education or society
The study shows evidence of student's learning outcome, learning experience, strategies for learning, or perception of learning	The study takes a teacher's perspective *(e.g. of student learning, course design)*
	The study is about the design of chatbots or the performance of LLMs *(e.g. solving assignments, predicting student success)*
	The study is about technology acceptance *(e.g. the ease of use or the perceived usefulness of chatbots)*
	The effect of GenAI chatbots on student learning cannot be separated from other variables *(e.g. the course design or learning platform*)

Extraction of Information

Based on an extraction template, information relevant for our review was extracted from the 49 selected publications. The extraction template held information about each study such as the location, duration, academic level, number of students, the type of intervention and control experiment, and the chosen analysis method. The template was then populated with key findings about the different dimensions of learning to be investigated: motivation, engagement, self-efficacy, self-regulation, comprehension, critical thinking, problem-solving, and learning performance. The filled extraction table was used to structure the outcomes of this review.

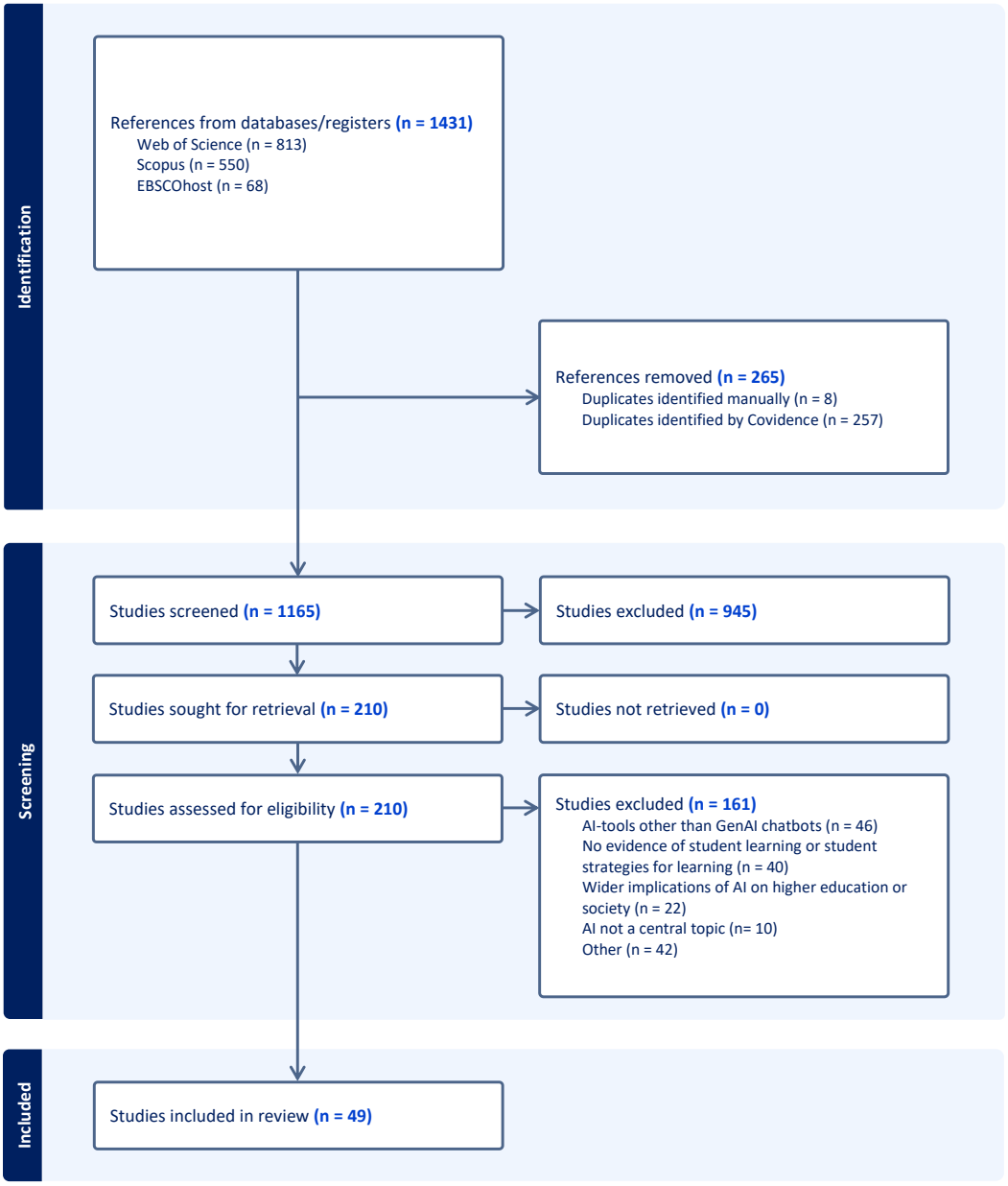

Figure 2. Selection Process Applied in This Literature Review (source: Covidence).

Analytics

Table 3 shows the distribution of the 49 studies considered in this review according to geographical location, discipline, data collection method, and the type of GenAI tool used. Almost half of the studies (23) were conducted in Asia whereas the other half took place in Europe (9), the Americas (11), Africa (2) and Australia (1). A total of 12 studies were conducted within the discipline computer science and seven were conducted in the context of language; typically in English as a Foreign Language (EFL) classes. The rest of the studies were distributed across a wide range of disciplines embracing the humanities, natural, and social sciences. A vast majority of the studies were conducted on the basis of student's self-reporting, either through surveys (34) or interviews (9), and 12 studies were based on experiments. A total of 39 publications reported explicitly that ChatGPT was used, whereas the remaining 10 publications described the use of GenAI tools in more general terms, or they did not specify

which tool was applied in the works.

Table 3. Analytics of the Publications Selected for This Review

Location of study		Discipline	
Asia	23	Computer science	12
Europe	9	Language	7
North America	8	Business and management	6
South America	3	Engineering	6
Africa	2	Medicine	4
Australia	1	Science and math	4
Other/not specified	3	Music	1
		Social science	1
		Other/not specified	8
Method		**GenAI tool**	
Survey amongst students	34	ChatGPT	39
Test results (experimental)	12	Other/not specified	10
Interviews with students	9		
Other/not specified	10		

Results

This review considers results derived from the variety of methodological frameworks and assessment methods used across the selected studies. Both qualitative (e.g., interviews with students and open-ended survey questions) and quantitative data (e.g., closed-end survey questions and test results) are considered. Moreover, a substantial portion of the data is based on students' self-reported perceptions of their learning experience and performance. In the following, we examine how the selected studies address the dimensions of learning described above and how students' use of GenAI affects each of the dimensions.

Motivation

Out of the 49 articles considered in this review, a total of 20 mentioned student motivation explicitly in connection with the use of GenAI chatbots. The findings were largely positive and highlighted that the incorporation of GenAI chatbots into the learning process – either in class or at home – can boost students' motivation through enhanced enjoyment and satisfaction (Dahri et al., 2024; Hamid et al., 2023; Shoufan, 2023; Song & Song, 2023). Other studies found that motivation was sparked by the chatbot's versatility and accessibility in various learning activities (Karataş et al., 2024; Ou et al., 2024) and Zhang et al. (2023) suggested that individuals' association with a community that supports artificial intelligence can in itself increase motivation and lead to improved learning outcomes.

Hmoud et al. (2024) investigated student motivation when learning with ChatGPT and identified five main categories of motivation: task enjoyment, reported effort, result assessment, perceived relevance, and interaction. Their findings led to the conclusion that ChatGPT can be a potent tool for enhancing task motivation among Higher Education students. Silitonga et al. (2023) compared measures of learners' intrinsic and extrinsic motivation for student populations with and without access to ChatGPT. Students who used the chatbot for English writing tasks scored higher on their intrinsic as well as extrinsic motivation, which was attributed to the instant feedback received on their writing.

Hsu (2023) recorded the time dedicated by students to learning medical terminology and found that students who had access to a chatbot practised for 7-10 hours per week whereas the control group practised only two hours. This led to the conclusion that incorporation of the tool could significantly enhance learning motivation. A survey amongst students in computer science revealed that motivation was ranked the third most important benefit of using ChatGPT, after it was used to generate codes and provide explanations (Singh et al., 2023). Gao et al. (2024) found that business students exhibited heightened motivation in AI-enhanced learning environments, particularly those emphasizing interactive and constructive learning methodologies. In the context of language learning, students expressed that they felt more motivated to practise and learn a foreign language when the learning was augmented by an AI chatbot (Muthmainnah et al., 2022).

On the negative side, Hmoud et al. (2024) stressed that learning with ChatGPT also comes with an element of anxiety associated with incorrect information provided by the chatbot. In a survey by Zhao et al. (2023), some students claimed that the use of a GenAI chatbot had a demotivating effect with respect to learning, as it was possible to generate answers with little effort. Holland & Ciachir (2024) found that ChatGPT has the potential to dent student motivation in the absence of perceived fair attainment and differentiation.

Engagement

A total of 13 studies in this review shed light on the impact of GenAI chatbots on student engagement. Overall, the integration of these chatbots appears to lead to increased student engagement by facilitating more active and effective participation in the learning process. One key factor contributing to this effect is the chatbots' capacity to deliver personalized learning experiences. The tools can achieve this personalisation in different ways, for example by offering guidance adapted to the individual's knowledge level (Hsu, 2023), providing immediate feedback (Song & Song, 2023), or supporting some metacognitive activities such as setting up learning goals and monitoring progress (Zhao et al., 2023).

Karataş et al. (2024) investigated how the implementation of ChatGPT in foreign language learning affected students' learning experiences, including engagement. In semi-structured interviews, students indicated that interacting with the chatbot had enhanced their engagement in various language-related activities, including writing and grammar exercises. The chatbots' instant feedback and ease of use were identified as key factors contributing to this enhanced engagement. Similarly, Hamid et al. (2023), through questionnaires and follow-up interviews with undergraduate pharmacy students, found that integrating ChatGPT into the problem-based

learning process introduced a positive emotional dimension that further promoted student engagement. Notably, introverted learners – who typically hesitate to ask questions in classroom settings – benefited significantly, as interacting with the chatbot gave them the opportunity to seek clarification and solve doubts that they would not want to voice in class.

Four out of the 13 articles were experimental studies. In all four, the chatbots were carefully integrated in the instruction and students in the experimental groups received guidance on how to use the technology. Learners were instructed about what specific tasks to complete with AI assistance, how to interact effectively with the tool, and how to utilize it as a learning facilitator (Hsu, 2023; Jost et al., 2024; Song & Song, 2023; Zhou & Kim, 2024). In one case, students also had the support of two experienced assistants who were available while they were using the chatbots to complete their assignments (Jost et al., 2024). Similarly, students taking part in some of the non-experimental studies conducted in specific courses received instructions on how to maximize the use of the chatbots (Araujo & Cruz-Correia, 2024; Karataş et al., 2024; Yang et al., 2024). This type of guidance for students to use GenAI effectively leads to productive interactions with the chatbots and is therefore linked to their positive impact on student engagement. On the contrary, the absence of guidelines and instructions reduces engagement, as students struggle to make optimized use of the tools, for example due to difficulties in creating suitable prompts (Tossell et al., 2024).

Self-Efficacy

Seven of the reviewed articles yielded results about the connection between the use of GenAI chatbots and self-efficacy, all of which indicated a positive impact of GenAI chatbots on students' perceived self-efficacy. The findings suggest that the interactive nature of the tools, the possibility of asking for step-by-step explanations, and the provision of personalised assistance and immediate feedback contribute to boosting learners' confidence in their abilities, consequently increasing self-efficacy (Guo & Lee, 2023; Hamid et al., 2023; Inoferio et al., 2024; Liang et al., 2023). This effect has been observed across a variety of disciplines, such as mathematics, business and languages, as well as in connection with different types of activities, such as writing, problem-solving and analysis of information.

Among these studies, Liang et al. (2023) included the largest number of participants. The results from a questionnaire survey completed by 389 undergraduate students from different Chinese universities and a variety of disciplines revealed that, when used as learning assistants, GenAI tools can lead to increased self-efficacy. GenAI's adaptability to students' level of understanding enables these tools to deliver tailored and varied explanations of complex concepts, thereby personalizing the learning process. Consequently, interaction with this technology can facilitate learners' comprehension of complex content and increase their ability to solve more challenging tasks, which in turn increases their perceived ability to achieve learning goals.

Despite the reported positive correlation between the use of GenAI chatbots and increased self-efficacy, it should be noted that two of the studies also indicated that the interaction with GenAI chatbots can reduce self-efficacy for some learners. Whereas some students perceive GenAI's assistance as a way of jump-starting their thinking,

others feel that their creativity is being outsourced to a tool, which results in a loss of confidence in their own ability to think by themselves, particularly if they struggle to come up with ideas beyond the ones suggested by the chatbot (Habib et al., 2024). Furthermore, although the immediate assistance obtained by interacting with GenAI chatbots can accelerate learning, it can also prevent students from making the cognitive effort that is needed to gain confidence in their own capabilities (Gao et al., 2024).

Self-Regulation

Self-regulation is addressed in four of the reviewed articles – two experimental studies and two non-experimental studies – which reported that interaction with GenAI chatbots has the potential to enhance students' self-regulation skills. Improvements were observed on students' ability to self-assess, re-examine their reasoning, evaluate the reliability of sources, set goals, and apply appropriate learning strategies (Elkhodr et al., 2023; Hmoud et al., 2024; Muthmainnah et al., 2022; Song & Song, 2023).

In the two experimental studies, students received substantial guidance on how to use the technology as an assistive tool. In the experiment conducted by Elkhodr et al. (2023), which involved undergraduate and graduate ICT students, instructors facilitated and promoted the use of ChatGPT as a self-directed learning aid. Students in the experimental groups used ChatGPT to analyse a case study. The chatbot was recommended for supporting content comprehension and providing examples of concepts, whereas the mere generation of answers was discouraged. In Song & Song (2023) study, where undergraduate EFL students used ChatGPT for writing tasks, participants in the experimental group were guided to use the tool ethically, ensuring that their unique style and ideas were preserved in their writing. While the integration of ChatGPT-generated feedback was promoted, considerable emphasis was placed on the preservation of academic integrity. Both studies concluded that collaborating with ChatGPT as a supplementary aid to learning can enhance students' self-regulation capacities, for example by helping them identify gaps in knowledge and become more reflective and self-directed learners.

Similar results were obtained in the two non-experimental studies where the development of self-regulation skills was observed. In a survey completed by 453 EFL undergraduate students, they indicated that they had self-assessed their judgements, re-examined their interpretations and considered carrying out additional research while interacting with the tools (Muthmainnah et al., 2022). These self-reports resemble closely the information provided by the 15 participants interviewed in Hmoud et al. (2024) study, where students stated that they had improved their self-evaluation skills as well as their ability to verify sources and the chatbot-provided information.

Comprehension

Comprehension featured in 25 of the articles in our collection. The interaction with GenAI chatbots was reported to enhance students' comprehension of complex subject matter in a variety of disciplines, including chemistry (Guo & Lee, 2023), computer science (Marquardson, 2024), Information and Communication Technology (Elkhodr et al., 2023), entrepreneurship (Hammoda, 2024), math (Gouia-Zarrad & Gunn, 2024; Inoferio et al.,

2024), music (Zhou & Kim, 2024), and pharmacy (Hamid et al., 2023). A statistical analysis by Dahri et al. (2024) supported these findings and concluded that effective student interaction with AI-tools - primarily chatbots - can enhance comprehension, engagement, and knowledge retention, resulting in improved academic performance. Komba (2024) highlighted that the straightforward manner and simple language of ChatGPT helped students grasp complex concepts more easily.

Specifically for language learning, Zhang et al. (2023) found that a heightened comprehension of AI chatbots among students correlates with an increased likelihood of their proficient utilisation in the context of language acquisition. Xiao & Zhi (2023) found that the extent to which ChatGPT could possibly contribute to language development was determined by whether it was used as an assistant or as a content creator, where the latter had little effect on student's language learning. Karatas et al. (2024) suggested that ChatGPT-assisted language learning has a positive impact on writing and grammar skills and new vocabulary acquisition. In contrast, they claimed that speaking practise could be negatively affected and over-reliance on the chatbot would also have a negative impact on student's language acquisition.

In connection with programming, Haindl & Weinberger (2024) and Singh et al. (2023) reported that students found ChatGPT more suitable for learning programming concepts than for the actual implementation of programming. Some students were afraid of losing their proficiency in programming if they relied too heavily on code generated with GenAI. This concern was supported by Jost et al. (2024) who found that students' use of a GenAI chatbot for seeking additional explanations in connection with programming exercises did not have a statistically significant impact on students' final grades when they were examined without any aids. Rajala et al. (2023) concluded that the positive implications of using ChatGPT for programming outweighed the negative ones, which were in their study mostly associated with unreliable outputs.

Critical Thinking

A total of 19 articles reported findings related to the impact of GenAI chatbots on critical thinking skills. These studies indicated that while the use of this technology has potential to enhance students' critical thinking abilities, it can also impede their development under certain circumstances. The interaction with GenAI chatbots, when integrated in teaching and learning activities, could improve learners' critical thinking in different ways. Aware of the occasional inaccuracies and erroneous information in GenAI-created output, students engaged in verifying the facts contained in the output and evaluating the validity of the provided information. They also made decisions about to what extent they could rely on the tool, balancing the benefits of using the technology with the active use of their human intervention. This type of interaction with GenAI chatbots, which involved analysing, evaluating, and inferencing, contributed to the development of students' critical thinking abilities (Hmoud et al., 2024; Sánchez-Guerrero et al., 2024; Zhao et al., 2023).

Furthermore, collaborating with the tools exposed learners to new perspectives on the topics being studied, challenging them to question their existing arguments, explore new ideas and pose further questions. Guo & Lee´s study (2023), with 29 undergraduate chemistry students as participants, explored ChatGPT's role in fostering

students' critical thinking skills. Students interacted with the chatbot in three different stages of an essay-writing activity, the third stage fully devoted to analysing and critiquing ChatGPT's output. Through the completion of a survey, students reported an improvement in their critical thinking competences, mentioning for example the ability to pose incisive questions and draw logical conclusions. Similar results, also based on students' self-reporting, were obtained in Muthmainnah et al. (2022) research. A survey completed by 453 EFL students revealed that interacting with the GenAI tools had helped them consider opposing arguments and justify their own conclusions, thus enhancing their logical reasoning abilities.

Despite these positive findings, the reviewed articles have also reported the potential risks that the use of GenAI chatbots can pose for the development of students' critical thinking. Students could become overly reliant on the tools instead of engaging in the sort of collaborative interaction mentioned above, they become dependent on the quick responses provided by the chatbots and do less independent thinking (Hamid et al., 2023; Song & Song, 2023). Some of the consequences of this passive approach would be the production of responses that, being too influenced by the GenAI-generated content, are lacking in depth, originality and personal style (Araujo & Cruz-Correia, 2024; Hyde et al., 2024). These challenges were reported by students themselves in several of the studies. For example, in the questionnaire completed by 430 international students at a British university, 39% of the participants identified reduced critical thinking and originality as the largest threat associated with the use of ChatGPT in academia (Singh et al., 2023).

Problem-Solving Skills

In this review, 14 articles provided results regarding the impact of GenAI chatbots on problem-solving skills. The findings in these studies indicated that the use of GenAI chatbots can both enhance and hinder the development of students' problem-solving abilities. Students can benefit from using GenAI chatbots by accelerating some tasks such as accessing information and obtaining clarification for doubts, which subsequently leaves them with more time to focus on the more complicated steps and makes the problem-solving process more time-efficient (Hmoud et al., 2024; Singh et al., 2023). Brainstorming, aiding with the comprehension of the underlying problem to solve, and organizing information were also mentioned as tasks that this technology can assist with, having a positive impact on students' problem-solving skills (Bravo & Cruz-Bohorquez, 2024); particularly when utilised to support tasks that involve creativity and user-centric design principles (Elkhodr et al., 2023).

The use of GenAI chatbots was also shown to enhance learners' perception of self-efficacy and preparedness when engaging with problem-solving tasks. The possibility of obtaining individual support aligned with their needs and receiving immediate feedback increased students' confidence levels, motivation and perseverance to overcome challenges. Hamid et al. (2023) investigated the perceptions of 18 undergraduate pharmacy students towards the use of ChatGPT in process-driven problem-based learning (PDPBL). Participants reported that interacting with the chatbot made the problem-solving process more engaging and increased their motivation to take an active role in it. Similar results were obtained in Inoferio et al. (2024) study, which involved 20 undergraduate students from different BSs (e.g., computer science, engineering, etc.) and enrolled in math courses. This research revealed that when used in a supporting role, AI models helped learners cope with math anxiety and

low confidence and empowered them to become more autonomous and proactive in problem-solving activities.

While integrating GenAI chatbots can have a positive influence on students' development of problem-solving skills, it can also make a negative impact if the technology is not used appropriately. One aspect that is emphasized in the reviewed articles is the risk of students over-relying on the tools, by using them as generators of solutions rather than as learning aids. This can prevent learners from engaging in the cognitive processes that help them succeed in problem-solving tasks independently (Bravo & Cruz-Bohorquez, 2024; Elkhodr et al., 2023; Jost et al., 2024). AI models can also impede the improvement of problem-solving abilities if students are not able to evaluate the accuracy and relevance of the chatbot's output, possibly failing to identify incorrect information and using it in the problem-solving tasks (Hamid et al., 2023; Swapna & Jin SHIM, 2023).

Learning Performance

Student's learning performance was addressed in 16 publications in our collection. Numerous experiments across disciplines allowed students to use a GenAI chatbot for practising before taking a test in a controlled environment without access to the chatbot. Most of these studies revealed that students who had practised with a chatbot performed better than the control groups who were not using chatbots. Zhou & Kim (2024) found that music students scored higher in a test of their cognitive grasp and comprehension of fundamental concepts, genres, scales, chords, harmony, and composition techniques pertaining to music when they used ChatGPT for preparation. Likewise, Kavadella et al. (2024) showed that dental students performed better in a multiple-choice quiz when they had used ChatGPT for preparation and Hsu (2023) found that students performed better in a test of medical terminology when they were allowed to use a chatbot for practising. ChatGPT and a dedicated 'Termbot' were compared, and the highest test scores were obtained by the students using ChatGPT. Song & Song (2023) compared test results for students who were trained in English as a Foreign Language (EFL) where the experimental group, who had practised with ChatGPT, demonstrated an enhanced proficiency in various aspects of writing, including organization, coherence, grammar, and vocabulary. Habib et al. (2024) carried out an experiment to test student's creative thinking skills with and without access to ChatGPT. They concluded that the use of GenAI enhanced divergent thinking, increasing the diversity of ideas but with a potential for getting answers that are too generic and have a negative impact on student's creativity.

In contrast to the studies mentioned above, Jost et al. (2024) identified a significant negative correlation between the average use of LLMs and students' final grades in a programming course; especially when the chatbot was used for code generation or debugging. In their experiment, students were allowed to use a chatbot for practising but not for the final test. Based on the experiment, Jost et al. (2024) concluded that reliance on LLMs correlates with diminished academic performance in programming assignments. Another coding experiment was conducted by Qureshi (2023). Two groups of ICT students worked on the same coding exercises where the experimental group utilized ChatGPT and the control group did not. The experimental group performed better than the control group, achieving higher scores in less time. However, the experimental group was not able to achieve perfect scores due to inaccuracies or inconsistencies in the code generated by ChatGPT.

Three studies compared assignment outcomes for student cohorts pre-ChatGPT with outcomes of students who had access to ChatGPT and this led to contradicting results. Hyde et al. (2024) found that students performed significantly better in a standardised management exam after ChatGPT was introduced. In a study by Kim et al. (2024), lab reports by engineering students were assessed before and after the reports had been through a revision aided by ChatGPT. The report quality improved due to students' enhanced lab report genre understanding but the use of ChatGPT also lead students to provide false claims, incorrect lab procedures, or extremely broad statements. Tossell et al. (2024) found through an essay assignment that students who had used ChatGPT to develop their hand-ins did not perform better than the previous year's cohorts who were taught by the same teacher but without access to a chatbot.

Students' self-reported performance increased when a chatbot assisted in diverse tasks and assignments including coding-related tasks, project proposals, assignments and report writing (Komba, 2024). A survey amongst students in Malaysia and Pakistan indicated that the rising use of AI tools has improved students' satisfaction levels and significantly impacted students learning outcomes (Dahri et al., 2024). Ou et al. (2024) concluded from their large-scale survey of Swedish students that GenAI tools can enhance students' academic communication performance and facilitate personal language development. The study emphasised how AI-powered language tools are also conducive to transforming the academic writing process into an additional learning space.

Discussion

Our purpose was to establish an overview of early findings regarding the impact of GenAI chatbots on student learning in Higher Education. Based on the literature outlined above, our impression is that students' use of GenAI has high potential to impact positively on the learning dimensions explored in this review – motivation, engagement, self-regulation, self-efficacy, comprehension, critical thinking, problem-solving skills, and learning performance. However, across all these dimensions, some negative effects are also evident, and although these are in many cases outweighed by the benefits, they remain significant enough to be taken into consideration in drawing conclusions. The coexistence of predominantly positive but also some negative effects of GenAI-chatbots' use on student learning underscores not only the complexity of examining the intersection between GenAI technologies and human learning, but also the relatively short period of time we have had to use, explore and research this technology.

Our findings align well with conclusions drawn in previous review articles where the use of GenAI chatbots improved academic performance, motivation, interest, and perceived value of learning (Wu & Yu, 2024; Deng et al., 2025; Mai et al., 2024). In line with Abu Khurma et al. (2024), our findings show that students' use of GenAI chatbots can boost academic engagement, particularly when the interaction is guided through completion of structured tasks. Our review indicates that the use of GenAI can increase students' confusion or anxiety about inaccurate or misleading outputs and make them question the fairness of assessment. This finding is contradictory to statements by Wu & Yu (2024), who conclude that GenAI chatbots can create a relaxed learning environment, especially helpful for anxious learners. These different conclusions may be explained by the fact that anxiety is a highly individualised emotion and we also speculate that it may vary geographically depending on the way Higher

Education is structured and financed and the pressure this can put on students. Evidence from our review and from previous reviews (Bruun et al., 2024; Faisal, 2024) shows unambiguously that students' use of GenAI chatbots supports self-regulation and independent learning. This can be a benefit in some contexts whereas in other contexts, the benefit is counteracted by a risk that more social aspects of learning are obstructed (Bruun et al., 2024).

From the publications in our collection, we observe that GenAI was used in various ways, ranging from carefully structured and guided classroom exercises to students' independent collaboration with GenAI chatbots, both in class and at home. Many examples of guided use incorporated a reflective aspect where students were asked to critically reflect about their learning experience using this technology, which included evaluating the information provided by the chatbots and assessing the reliability of the tools (Hmoud et al., 2024; Zhao et al., 2023). We note that these studies were generally positive about the effects of GenAI chatbots on student learning and the implementation of GenAI in Higher Education. This can be interpreted as an indication of how critical it is that learners understand the limitations of the technology, apply their own judgement when using it, and recognise the value of their human intervention.

Another example of studies which yielded overall positive results were those which provided students with clear guidelines about how to use the chatbots to assist their learning and how to preserve academic integrity. Students received for example instructions about what specific tasks they could complete with GenAI support and were reminded of the importance of preserving their own ideas and writing style (Hsu, 2023; Song & Song, 2023). Conversely, in studies where learners used GenAI chatbots without guidance, they would most typically used them to explain complex concepts or to practise their knowledge in advance of a test. In this case, students would be left on their own to interpret and critically assess the chatbot outcomes, and the absence of guidelines and instructions could reduce engagement and lead to a lack of depth in students' responses (Araujo & Cruz-Correia, 2024; Tossell et al., 2024). This seems to indicate that a carefully guided integration of GenAI chatbots in learning environments is paramount to ensure that students make an ethical and effective use of this technology to support their learning.

In connection with the above, the risk of over-reliance on AI tools was mentioned in numerous studies, noting that it can hinder the development of independent thinking, creativity, and problem-solving skills. This is well aligned with previous reviews (Faisal, 2024; Mai et al., 2024). Cognitive shortcuts taken through AI use can reduce the mental effort needed for meaningful learning and long-term retention and students may also lose confidence in their own abilities when they depend too much on AI-generated ideas. In practical areas like programming, when students rely on AI for implementation rather than understanding, there is a risk of skill degradation. This is particularly evident from coding experiments allowing students to use a GenAI chatbot for practising but not for the final test (Jost et al., 2024). Similar effects were observed in ChatGPT- assisted language learning, where over-reliance on the technology impacted language acquisition negatively (Karataş et al., 2024).

Most of the studies considered in this review base their conclusions on consensus by student majorities in the populations investigated, or on mean values of test scores. It is important to keep the diversity of students in mind

when interpreting such results. Students have different learning attitudes (Tu & Hwang, 2023) and growth mindsets (Tu, 2024), which determine their benefits of GenAI chatbots. Tu (2024) argued that students with a so called 'low learning attitude' would mostly use ChatGPT to achieve lower-order skills such as comprehension of concepts whereas students with a 'high learning attitude' could more effectively make use of ChatGPT to stimulate their creativity and achieve higher-order thinking skills (Bloom et al., 1956) (Krathwohl, 2002). In other words, there is a risk that access to GenAI tools might enhance inequalities amongst students (Holland & Ciachir, 2024).

Conclusions

This literature review reveals that GenAI chatbots are most powerful for supporting student learning when they take the role of tutors or assistants guiding students through dialogue to establish their own findings or solutions. This contrasts with using chatbots to generate complete solutions, thereby bypassing problem-solving, critical thinking, and creativity, which are all important steps and skills for achieving deep learning. A human tutor would typically not provide complete answers to students' assignments; rather the good tutor would provide guiding questions to support the student's progress towards a given learning objective.

The human tutor might take proactive steps e.g. by asking a student challenging questions. This capability is less pronounced for tutoring by means of GenAI chatbots, where it takes some initiative and insights from the student's side to get meaningful answers from the chatbot. In the not so distant future, a combination of intelligent tutoring systems, designed for learning and built into learning management systems, in combination with underlying GPT-technology might offer the optimal tutoring that promotes deep learning and represents a viable solution to the Bloom's "2 Sigma Problem" mentioned in the Introduction (Bloom, 1984).

Recommendations for Integrating Genai Chatbots in Learning Environments

The findings in this review have several important implications that can offer valuable guidance for educators planning the integration of GenAI chatbots to support student learning. The identified potential benefits of these tools should encourage teachers to explore their use with confidence and curiosity, seeking applications that can allow them to support their students even better than they already do. Also based on the review's results, it is recommended that the integration of GenAI chatbots be carefully planned. This includes clarifying which GenAI uses are encouraged or discouraged, ensuring they are appropriate for students' existing knowledge and level of study, and addressing potential interferences with academic integrity and ethical regulations.

Providing students with clear guidelines, fostering their AI literacy skills, and incorporating the evaluation of GenAI-generated output as part of the activities are some of the measures that can ensure a pedagogical sound integration of this technology. Finally, it is also important to acknowledge and be prepared for differences in how students experience their interaction with GenAI chatbots. While many learners may gain confidence and become more independent in their learning by being assisted by the tools, others may doubt their ability to think critically and create original content.

Limitations

This review is based on the earliest publications about GenAI and student learning, which were all published within 1.5 years from the launch of ChatGPT in the autumn of 2022. Most of these early studies are based on small-scale analyses with relatively few participants or on larger-scale studies in less controlled environments where students have used different GenAI tools for different purposes. Larger systematic studies including control experiments are needed to establish more firm conclusions. Another limitation of this review is that outcomes of the different studies might be heavily influenced by the way experiments were setup. Most importantly, the conclusions drawn from experiments where students' use of GenAI was permitted for solving an assignment are bound to be different from those where GenAI was allowed for practising but not for the final assignment.

Future Research Directions

The findings of this review also highlight certain gaps in the existing literature that can be addressed by future studies. Further research is still needed to gain a more nuanced understanding of how students can effectively use GenAI chatbots as learning partners to enhance, rather than substitute, their original ideas and critical thinking skills. For instance, future research could explore whether this positive outcome is more easily achievable in certain disciplines, at specific levels of expertise, or in relation to particular types of tasks. Additionally, as both educators and learners gain more experience with GenAI chatbots in educational settings, future studies should try to shed more light on whether or how cognitive offloading affects learning outcomes and performance. Another area where future research could focus concerns the emotional dimension of GenAI use, especially in the case of those learners for whom these tools seem to have a negative effect in terms of undermining their trust in their human abilities or prompting unfavorable comparisons between their original contributions and the ones generated by the chatbots.

Acknowledgements

The authors would like to thank Mette Bechmann from CBS Library for her guidance and support throughout the PRISMA review process. They also want to express their gratitude to Sidsel-Marie Winther Prag and Mathias Lund for reviewing the manuscript and their insightful comments.

References

Abu Khurma, O., Albahti, F., Ali, N., & Bustanji, A. (2024). AI ChatGPT and student engagement: Unraveling dimensions through PRISMA analysis for enhanced learning experiences. *Contemporary Educational Technology*, *16*(2). https://doi.org/10.30935/cedtech/14334

Akcaoğlu, M. Ö., Mor, E., & Külekçi, E. (2023). The mediating role of metacognitive awareness in the relationship between critical thinking and self-regulation. *Thinking Skills and Creativity*, *47*. https://doi.org/10.1016/j.tsc.2022.101187

Araujo, S. M., & Cruz-Correia, R. (2024). Incorporating ChatGPT in Medical Informatics Education: Mixed

Methods Study on Student Perceptions and Experiential Integration Proposals. *JMIR Medical Education*, *10*. https://doi.org/10.2196/51151

Barkley, E. F., & Major, C. H. (2020). *Student engagement techniques: a handbook for college faculty* (2nd ed.). Jossey-Bass.

Beatson, N. J., Berg, D. A. G., & Smith, J. K. (2018). The impact of mastery feedback on undergraduate students' self-efficacy beliefs. *Studies in Educational Evaluation*, *59*, 58–66. https://doi.org/10.1016/j.stueduc.2018.03.002

Bhullar, P. S., Joshi, M., & Chugh, R. (2024). ChatGPT in higher education - a synthesis of the literature and a future research agenda. *Education and Information Technologies*. https://doi.org/10.1007/s10639-024-12723-x

Bloom, B. S. (1984). The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*, 4–16.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook 1: Cognitive domain* (B. S. Bloom, Ed.). David McKay.

Bol, L., & Garner, J. K. (2011). Challenges in supporting self-regulation in distance education environments. *Journal of Computing in Higher Education*, *23*(2–3), 104–123. https://doi.org/10.1007/s12528-011-9046-7

Bowden, J. L. H., Tickle, L., & Naumann, K. (2021). The four pillars of tertiary student engagement and success: a holistic measurement approach. *Studies in Higher Education*, *46*(6), 1207–1224. https://doi.org/10.1080/03075079.2019.1672647

Bravo, F. A., & Cruz-Bohorquez, J. M. (2024). Engineering Education in the Age of AI: Analysis of the Impact of Chatbots on Learning in Engineering. *Education Sciences*, *14*(5). https://doi.org/10.3390/educsci14050484

Bruun, M. H., Krause-Jensen, J., & Hasse, C. (2024). Store sprogmodeller og AI-chatbots på videregående uddannelser. In *Pædagogisk Indblik* (Vol. 26). https://dpu.au.dk/fileadmin/edu/Paedagogisk_Indblik/AI_paa_videregaaende_uddannelser/26_Store_Sprogmodeller_og_AI-chatbots_paa_videregaaende_uddannelser_-11-12-2024.pdf

Caspersen, J., Smeby, J. C., & Olaf Aamodt, P. (2017). Measuring learning outcomes. *European Journal of Education*, *52*(1), 20–30. https://doi.org/10.1111/ejed.12205

Casuso-Holgado, M. J., Cuesta-Vargas, A. I., Moreno-Morales, N., Labajos-Manzanares, M. T., Barón-López, F. J., & Vega-Cuesta, M. (2013). The association between academic engagement and achievement in health sciences students. *BMC Medical Education*, *13*(33). http://www.biomedcentral.com/1472-6920/13/33

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: the state of the field. *International Journal of Educational Technology in Higher Education*, *20*(1). https://doi.org/10.1186/s41239-023-00392-8

Dahri, N. A., Yahaya, N., Al-Rahmi, W. M., Vighio, M. S., Alblehai, F., Soomro, R. B., & Shutaleva, A. (2024). Investigating AI-based academic support acceptance and its impact on students' performance in Malaysian and Pakistani higher education institutions. *Education and Information Technologies*. https://doi.org/10.1007/s10639-024-12599-x

Davies, M. (2015). A Model of Critical Thinking in Higher Education. In M. B. Paulsen (Ed.), *Higher Education: Handbook of Theory and Research* (pp. 41–92). Springer International Publishing Switzerland. https://doi.org/10.1007/978-3-319-12835-1_2

Davis, F. (1985). *A technology acceptance model for empirically testing new end-user information systems : theory and results* [Massachusetts Institute of Technology]. http://hdl.handle.net/1721.1/15192

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, *227*, 105224. https://doi.org/10.1016/j.compedu.2024.105224

Elkhodr, M., Gide, E., Wu, R., & Darwish, O. (2023). ICT students' perceptions towards ChatGPT: An experimental reflective lab analysis. *STEM Education*, *3*(2), 70–88. https://doi.org/10.3934/steme.2023006

Fahim, M., & Masouleh, N. S. (2012). Critical thinking in higher education: A pedagogical look. *Theory and Practice in Language Studies*, *2*(7), 1370–1375. https://doi.org/10.4304/tpls.2.7.1370-1375

Faisal, E. (2024). Unlock the potential for Saudi Arabian higher education: a systematic review of the benefits of ChatGPT. *Frontiers in Education*, *9*. https://doi.org/10.3389/feduc.2024.1325601

Flake, J. K., Barron, K. E., Hulleman, C., McCoach, B. D., & Welsh, M. E. (2015). Measuring cost: The forgotten component of expectancy-value theory. *Contemporary Educational Psychology*, *41*, 232–244. https://doi.org/10.1016/j.cedpsych.2015.03.002

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

Gouia-Zarrad, R., & Gunn, C. (2024). Enhancing students' learning experience in mathematics class through ChatGPT. *International Electronic Journal of Mathematics Education*, *19*(3). https://doi.org/10.29333/iejme/14614

Guo, Y., & Lee, D. (2023). Leveraging ChatGPT for Enhancing Critical Thinking Skills. *Journal of Chemical Education*, *100*(12), 4876–4883. https://doi.org/10.1021/acs.jchemed.3c00505

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

Haindl, P., & Weinberger, G. (2024). Students' Experiences of Using ChatGPT in an Undergraduate Programming Course. *IEEE ACCESS*, *12*, 43519–43529. https://doi.org/10.1109/ACCESS.2024.3380909

Hamid, H., Zulkifli, K., Naimat, F., Yaacob, N. L. C., & Ng, K. W. (2023). Exploratory study on student perception on the use of chat AI in process-driven problem-based learning. *Currents in Pharmacy Teaching and Learning*, *15*(12), 1017–1025. https://doi.org/10.1016/j.cptl.2023.10.001

Hammoda, B. (2024). ChatGPT for Founding Teams: An Entrepreneurial Pedagogical Innovation. *International Journal of Technology in Education*, *7*(1), 154–173. https://doi.org/https://doi.org/10.46328/ijte.530

Harlim, J., & Belski, I. (2013). Long-term Innovative Problem Solving Skills: Redefining Problem Solving. *International Journal of Engineering Education*, *29*(2), 280–290.

Hmoud, M., Swaity, H., Hamad, N., Karram, O., & Daher, W. (2024). Higher Education Students' Task Motivation in the Generative Artificial Intelligence Context: The Case of ChatGPT. *Information*, *15*(1). https://doi.org/10.3390/info15010033

Holland, A., & Ciachir, C. (2024). A qualitative study of students' lived experience and perceptions of using ChatGPT: immediacy, equity and integrity. *Interactive Learning Environments*, *33*(1), 483–494. https://doi.org/10.1080/10494820.2024.2350655

Hsu, M. H. (2023). Mastering medical terminology with ChatGPT and Termbot. *Health Education Journal*. https://doi.org/10.1177/00178969231197371

Hyde, S. J., Busby, A., & Bonner, R. L. (2024). Tools or Fools: Are We Educating Managers or Creating Tool-Dependent Robots? *Journal of Management Education*. https://doi.org/10.1177/10525629241230357

Inoferio, H. V., Espartero, M. M., Asiri, M. S., Damin, M. D., & Chavez, J. V. (2024). Coping with math anxiety and lack of confidence through AI-assisted Learning. *Environment and Social Psychology*, *9*(5). https://doi.org/10.54517/esp.v9i5.2228

Jonassen, D. H. (2000). Toward a Design Theory of Problem Solving. *Educational Technology Research and Development*, *48*, 63–85. https://doi.org/https://doi.org/10.1007/BF02300500

Jost, G., Taneski, V., & Karakatic, S. (2024). The Impact of Large Language Models on Programming Education and Student Learning Outcomes. *Applied Sciences*, *14*(10), 4115. https://doi.org/10.3390/app14104115

Karataş, F., Abedi, F. Y., Ozek Gunyel, F., Karadeniz, D., & Kuzgun, Y. (2024). Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners. *Education and Information Technologies*, *29*, 19343–19366. https://doi.org/10.1007/s10639-024-12574-6

Kavadella, A., Da Silva, M. A. D., Kaklamanos, E. G., Stamatopoulos, V., & Giannakopoulos, K. (2024). Evaluation of ChatGPT's Real-Life Implementation in Undergraduate Dental Education: Mixed Methods Study. *JMIR Medical Education*, *10*(1). https://doi.org/10.2196/51344

Kim, D., Majdara, A., & Olson, W. (2024). A Pilot Study Inquiring into the Impact of ChatGPT on Lab Report Writing in Introductory Engineering Labs. *International Journal of Technology in Education*, *7*(2), 259–289. https://doi.org/10.46328/ijte.691

Kirschner, P. A. (2002). Cognitive load theory: implications of cognitive load theory on the design of learning. *Learning and Instruction*, *12*(1), 1–10. https://doi.org/https://doi.org/10.1016/S0959-4752(01)00014-7

Komba, M. M. (2024). The influence of ChatGPT on digital learning: experience among university students. *Global Knowledge, Memory and Communication*. https://doi.org/10.1108/GKMC-10-2023-0390

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, *41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Lee, J. S. (2014). The relationship between student engagement and academic performance: Is it a myth or reality? *Journal of Educational Research*, *107*(3), 177–185. https://doi.org/10.1080/00220671.2013.807491

Liang, J., Wang, L. L., Luo, J., Yan, Y. F., & Fan, C. (2023). The relationship between student interaction with generative artificial intelligence and learning achievement: serial mediating roles of self-efficacy and cognitive engagement. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1285392

Lovett, M. C. ; B. M. W. ; D. M. (2023). *How learning works: eight research-based principles for smart teaching* (2nd ed.). Jossey-Bass.

Mai, D. T. T., Da, C. Van, & Hanh, N. Van. (2024). The use of ChatGPT in teaching and learning: a systematic review through SWOT analysis approach. *Frontiers in Education*, *9*. https://doi.org/10.3389/feduc.2024.1328769

Marquardson, J. (2024). Embracing Artificial Intelligence to Improve Self-Directed Learning: A Cybersecurity

Classroom Study. *Information Systems Education Journal*, *22*(1), 4–13. https://doi.org/https://doi.org/10.62273/WZBY3952

Muthmainnah, Seraj, P. M. I., & Oteir, I. (2022). Playing with AI to Investigate Human-Computer Interaction Technology and Improving Critical Thinking Skills to Pursue 21st Century Age. *Education Research International*, *2022*. https://doi.org/10.1155/2022/6468995

Ou, A. W., Stöhr, C., & Malmström, H. (2024). Academic communication with AI-powered language tools in higher education: From a post-humanist perspective. *System*, *121*. https://doi.org/10.1016/j.system.2024.103225

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. *BMJ*, *372*(n71). https://doi.org/10.1136/bmj.n71

Pekruna, R. (2020). Commentary: Self-report is indispensable to assess students' learning. *Frontline Learning Research*, *8*(3), 185–193. https://doi.org/10.14786/flr.v8i3.637

Prakong, S. (2024). The Role of Critical Thinking in Enhancing Students' Problem-Solving Abilities in Higher Education. *Journal of Education, Humanities, and Social Research*, *1*(1), 10–16. https://doi.org/10.70088/scx8x622

Qureshi, B. (2023). ChatGPT in Computer Science Curriculum Assessment: An analysis of Its Successes and Shortcomings. *ACM International Conference Proceeding Series*, 7–13. https://doi.org/10.1145/3613944.3613946

Rajala, J., Hukkanen, J., Hartikainen, M., & Niemelä, P. (2023). "Call me Kiran" ChatGPT as a Tutoring Chatbot in a Computer Science Course. *ACM International Conference Proceeding Series*, 83–94. https://doi.org/10.1145/3616961.3616974

Rosenzweig, E. Q., Wigfield, A., & Eccles, J. S. (2019). Expectancy-value theory and its relevance for student motivation and learning. In *The Cambridge Handbook of Motivation and Learning* (pp. 617–644). Cambridge University Press. https://doi.org/10.1017/9781316823279.026

Sánchez-Guerrero, J., Miranda-López, X., Buenaño, H., & Lozada-Miranda, D. (2024). ChatGPT to Motivate Critical Thinking in the Teaching–Learning Process of First Semester Students at the Technical University of Ambato. In A. Rocha, J. M. R. Rodríguez, & C. H. Fajardo-Toro (Eds.), *Smart Innovation, Systems and Technologies* (Vol. 380, pp. 169–179). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-981-99-8894-5_15

Shoufan, A. (2023). Exploring Students' Perceptions of ChatGPT: Thematic Analysis and Follow-Up Survey. *IEEE ACCESS*, *11*, 38805–38818. https://doi.org/10.1109/ACCESS.2023.3268224

Silitonga, L. M., Hawanti, S., Aziez, F., Furqon, M., Zain, D. S. M., Anjarani, S., & Wu, T.-T. (2023). The Impact of AI Chatbot-Based Learning on Students' Motivation in English Writing Classroom. In H. Y.-M. & T. Rocha (Eds.), *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 14099 LNCS* (pp. 542–549). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/978-3-031-40113-8_53

Singh, H., Tayarani-Najaran, M.-H., & Yaqoob, M. (2023). Exploring Computer Science Students' Perception of

ChatGPT in Higher Education: A Descriptive and Correlation Study. *Education Sciences*, *13*(9). https://doi.org/10.3390/educsci13090924

Soderstrom, N. C., & Bjork, R. A. (2015). Learning Versus Performance: An Integrative Review. *Perspectives on Psychological Science*, *10*(2), 176–199. https://doi.org/10.1177/1745691615569000

Song, C. P., & Song, Y. P. (2023). Enhancing academic writing skills and motivation: assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1260843

Suskie, L. (2018). *Assessing student learning: a common sense guide* (3rd ed.). Jossey-Bass.

Swapna, G., & Jin SHIM, K. (2023). Exploring students' adoption of ChatGPT as a mentor for undergraduate computing projects: PLS-SEM analysis. *Proceedings of the 31st International Conference on Computers in Education Conference, Matsue, Shimane, Japan, 2023 December 4-8.*, 40–45. https://ink.library.smu.edu.sg/sis_research/8503/

Tossell, C. C., Tenhundfeld, N. L., Momen, A., Cooley, K., & Visser, E. J. de. (2024). Student Perceptions of ChatGPT Use in a College Essay Assignment: Implications for Learning, Grading, and Trust in Artificial Intelligence. *IEEE Transactions on Learning Technologies*, *17*, 1069–1081. https://doi.org/10.1109/TLT.2024.3355015

Tu, Y.-F. (2024). Roles and Functionalities of ChatGPT for Students with Different Growth Mindsets: Findings of Drawing Analysis. *Educational Technology & Society*, *27*(1), 198–214. https://doi.org/10.30191/ETS.202401_27(1).TP01

Tu, Y.-F., & Hwang, G.-J. (2023). University students' conceptions of ChatGPT-supported learning: a drawing and epistemic network analysis. *Interactive Learning Environments*, *32*(10), 6790–6814. https://doi.org/10.1080/10494820.2023.2286370

Van Dinther, M., Dochy, F., & Segers, M. (2011). Factors affecting students' self-efficacy in higher education. *Educational Research Review*, *6*(2), 95–108. https://doi.org/10.1016/j.edurev.2010.10.003

Wu, R., & Yu, Z. (2024). Do AI chatbots improve students learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology*, *55*(1), 10–33. https://doi.org/10.1111/bjet.13334

Xiao, Y., & Zhi, Y. (2023). An Exploratory Study of EFL Learners' Use of ChatGPT for Language Learning Tasks: Experience and Perceptions. *Languages*, *8*(3). https://doi.org/10.3390/languages8030212

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

Zhang, Y. Y., Viriyavejakul, C., & Sumettikoon, P. (2023). Integrating Chatbots in Educational Administration for Improved Language Learning Outcomes. *Eurasian Journal of Educational Research*, *104*, 142–163. https://doi.org/10.14689/ejer.2023.104.009

Zhao, X., Aydeniz, M., & Yuan, F. (2023). Exploring Opportunities and Challenges of AI-incorporated Biomedical Informatics Education: A Qualitative Study. *Proceedings 2023 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*. https://doi.org/10.1109/BHI58575.2023.10313394

Zhou, W., & Kim, Y. (2024). Innovative music education: An empirical assessment of ChatGPT-4's impact on student learning experiences. *Education and Information Technologies*, *29*, 20855–20881.

https://doi.org/10.1007/s10639-024-12705-z

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, *41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Zirar, A. (2023). Exploring the Impact of Language Models, Such as ChatGPT, on Student Learning and Assessment. *Review of Education*, *11*(3). https://doi.org/10.1002/rev3.3433