The influence of ChatGPT on student engagement: A systematic review and future research agenda

Chung Kwan Lo [a,*], Khe Foon Hew [b], Morris Siu-yung Jong [c]

[a] *Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong SAR, China*
[b] *Faculty of Education, The University of Hong Kong, Hong Kong SAR, China*
[c] *Department of Curriculum and Instruction & Centre for Learning Sciences and Technologies, The Chinese University of Hong Kong, Hong Kong SAR, China*

A B S T R A C T

ChatGPT, a state-of-the-art artificial intelligence (AI) chatbot, has gained considerable attention as a transformative yet controversial tool for enhancing teaching and learning experiences. Several reviews and numerous articles have been written about harnessing ChatGPT in education since its release on November 30, 2022. Besides summarising its strengths, weaknesses, opportunities, and threats (SWOT) as identified in previous systematic reviews of ChatGPT research, this systematic review aims to develop a new understanding of its influence on student engagement by synthesising the existing related research using a three-dimensional framework comprising behavioural, emotional, and cognitive aspects. We searched relevant databases and included 72 empirical studies published within one year of ChatGPT's initial release. The findings reveal robust but narrowly focused evidence related to behavioural engagement (i.e., work with ChatGPT) and disengagement (i.e., academic dishonesty). The evidence related to the emotional aspect is mixed, with instances of both engagement (e.g., satisfaction and interest/fun) and disengagement (e.g., disappointment and worry/anxiety). There is broad but weak evidence regarding cognitive engagement (e.g., increased understanding and positive self-perception) and disengagement (e.g., reduced critical thinking and overreliance). Our review uncovers several under-explored indicators of student engagement, pointing to the need for further research. Specifically, future studies could focus on students' study habits and attendance (behavioural engagement), social interaction (emotional engagement), and self-regulation and critical thinking (cognitive engagement) in ChatGPT-supported learning environments.

1. Introduction

There has been significant growth in the use of artificial intelligence (AI) chatbots in education (Hwang & Chang, 2023). The capabilities of these chatbots, which are computer programmes designed to mimic human-like conversations, have remarkably improved due to advancements in AI and natural language processing (NLP). A notable example in the field is ChatGPT, a pre-trained language model developed by OpenAI (Flanagin et al., 2023; Roumeliotis & Tselikas, 2023). Language models are statistical models designed to depict the probability distribution of natural language (Wu et al., 2023). They are often referred to as large language models as they are trained on large amounts of text data (Roumeliotis & Tselikas, 2023). The pre-training process allows them to learn

---

* Corresponding author.
  E-mail address: chungkwanlo@eduhk.hk (C.K. Lo).

the patterns of a language and thus generate human-like responses (Roumeliotis & Tselikas, 2023; Wu et al., 2023). ChatGPT was developed in two stages (Roumeliotis & Tselikas, 2023; Shen et al., 2023). The first stage involved unsupervised pre-training, in which the model was trained on an extensive text corpus using language modelling techniques. This stage aimed to equip the model with a thorough understanding of natural language structure and the relationships between words and sentences. In the second stage, the model underwent a supervised fine-tuning process whereby ChatGPT was trained on labeled datasets for tasks such as text completion and dialogue generation. This process involved iteratively adjusting the model's parameters to minimise the inconsistencies between the model's predictions and the correct task-specific labels. This training process enables ChatGPT to engage in realistic interactions with users, offering responses to inquiries, acknowledging errors, and refusing inappropriate requests (OpenAI, 2022). Remarkably, within just two months of its launch, it had attracted an estimated 100 million active users, making it the fastest-growing consumer application in history (Hu, 2023).

The rapid and widespread adoption of ChatGPT by both students and teachers highlights the growing importance and influence of AI tools in education (Ahmed & Sab, 2023). ChatGPT has become an integral component of the teaching and learning process. With its capacity to simulate human-like dialogue, ChatGPT has the potential to enhance student engagement by providing immediate responses to enquiries, facilitating group discussions, and even aiding teachers in lesson planning (Imran & Almusharraf, 2023; Lo, 2023; Vargas-Murillo et al., 2023). However, the use of ChatGPT in education also raises concerns about incorrect information and academic dishonesty (Gödde et al., 2023; Lo, 2023; Zhang & Tur, 2023). Moreover, overreliance on ChatGPT could potentially impair students' engagement in learning by depriving them of the opportunity to engage in critical thinking and problem-solving (Sallam, 2023; Vargas-Murillo et al., 2023). Therefore, it is essential to understand the influence of ChatGPT on student engagement to optimise its usage in educational settings and guide the direction of future research agendas for ChatGPT-supported learning.

To address this need, we first consolidated our understanding by summarising the major findings of previous systematic reviews of ChatGPT research in terms of its strengths, weaknesses, opportunities, and threats (SWOT). As presented in Section 4.1.5, significant knowledge gaps remain unaddressed, particularly regarding the holistic influence of ChatGPT on student engagement. Therefore, we conducted a systematic review of empirical studies of ChatGPT in educational settings and examined their findings relating to student engagement. As in the approach adopted by Crompton and Burke (2023), we sought to understand the characteristics of the research conducted on ChatGPT. We then drew on the three-dimensional model of student engagement (i.e., behavioural, emotional, and cognitive) proposed by Fredricks et al. (2004). The following research questions (RQ1 to RQ5) were posed to guide the review.

- RQ1: What are the main findings of previous systematic reviews of ChatGPT research based on SWOT analysis?
- RQ2: What are the characteristics of the research conducted on ChatGPT (including the geographical locations of the first authors, the academic levels and subject domains considered, and the methodological approaches employed)?
- RQ3: How does ChatGPT influence students' behavioural engagement?
- RQ4: How does ChatGPT influence students' emotional engagement?
- RQ5: How does ChatGPT influence students' cognitive engagement?

The remainder of this article is organised as follows. Section 2 introduces our analytical framework for student engagement. Section 3 details our research methods, including our search strategies, criteria for including and excluding articles, and approach to data extraction and analysis. Section 4 first summarises the key findings from previous systematic reviews of ChatGPT research (RQ1), their limitations and the knowledge gaps that remain unaddressed. The section then presents a synthesis of 72 journal articles, in which we discuss findings pertaining to RQ2 to RQ5. Finally, Section 5 concludes the systematic review, acknowledges its limitations, and provides recommendations for future studies.

2. Framework for student engagement

Fredricks et al. (2004) proposed that student engagement comprises three dimensions: behavioural, emotional, and cognitive. This three-dimensional model, derived through an extensive synthesis of relevant research, broadly covers all aspects of educational settings (see Fredricks et al., 2004 for a review). The model has been widely adopted for synthesising research in the context of technology-enhanced learning, with multiple indicators identified for each dimension (see Bond, 2020; Bond & Bedenlier, 2019; Bond et al., 2020; Lo & Hew, 2021). Therefore, this framework is appropriate for use in this review. To establish an initial analytical framework, we undertook a preliminary review of the literature using the Web of Science database, thereby facilitating our subsequent research synthesis.

2.1. Behavioural engagement (for RQ3)

Behavioural engagement encompasses actions that are indicative of students' participation, effort, study habits, task completion, and time spent on tasks (Bond, 2020; Bond & Bedenlier, 2019; Bond et al., 2020; Lo & Hew, 2021). According to Escalante et al. (2023), ChatGPT can enhance students' participation and involvement in learning due to its round-the-clock availability. They can work with ChatGPT as a collaborative partner (Jo, 2023a) to help proofread and refine their writing (Liu & Ma, 2024). Without time usage constraints, ChatGPT potentially improves study habits, such as aiding exam preparation and home learning (Lo, 2023; Zhang & Tur, 2023). In the study by Chan and Hu (2023), one student shared that ChatGPT facilitated task completion by handling repetitive tasks, thereby allowing greater focus on studying and advanced learning tasks.

Bond (2020) and Bond et al. (2020) identified absence, inattentiveness, lack of preparation, poor conduct, and task incompletion as

common indicators of behavioural disengagement. In terms of poor conduct, in particular, Sallam (2023) highlighted the potential misuse and academic dishonesty that may arise from ChatGPT's ability to pass reputable exams. This includes the submission of ChatGPT-generated materials as students' own work, thereby raising plagiarism concerns (Lo, 2023; Zhang & Tur, 2023). As one student shared, "all you need to do is just copying and pasting a few keywords and a few clicks for final production. From the perspective of ideal plagiarism, it seems perfect" (Yan, 2023, p. 13957).

2.2. Emotional engagement (for RQ4)

Bond and Bedenlier (2019) and Lo and Hew (2021) identified various indicators of emotional engagement, including excitement, interest/fun, enjoyment, satisfaction, positive social interaction, and reduced anxiety. Uddin et al. (2023) incorporated ChatGPT into an engineering class, where students found its functions to be exciting, fun, and enjoyable. The students reported high satisfaction with the learning experience provided by ChatGPT. In terms of positive social interactions, ChatGPT has the potential to foster knowledge sharing among students and create a collaborative and participatory learning environment (Duong, Vu, & Ngo, 2023). Furthermore, Wu et al. (2024) found that the use of ChatGPT reduced students' anxiety by eliminating fears associated with unresolved problems.

According to Bond (2020) and Bond et al. (2020), worry and anxiety are key indicators of emotional disengagement. Furthermore, our preliminary review of the literature suggested that the use of ChatGPT might impair social interaction and that some students felt upset or disappointed with their learning experiences. In the study by Chan and Hu (2023), the students were concerned about using ChatGPT due to its lack of operational transparency. If students perceive ChatGPT to be highly capable, they might begin to question the role of teachers, potentially leading to a negative impact on the student-teacher relationship (Chan & Hu, 2023). However, ChatGPT's responses may be biased or incorrect due to its training data (Lo, 2023). Therefore, as Cai et al. (2023) observed, some students may be disappointed with its accuracy and credibility.

2.3. Cognitive engagement (for RQ5)

Bond and Bedenlier (2019) identified various indicators of cognitive engagement, including learning from peers, critical thinking, self-regulation, positive self-perceptions, and understanding. In the study by Escalante et al. (2023), some students noted the benefits of receiving feedback from both their teachers and ChatGPT, expressing that this combination enabled them "to learn more from both" (p. 12). Thus, in the context of ChatGPT-supported education, learning could be an indicator of cognitive engagement. The findings of Zhang and Tur (2023) suggested that incorporating ChatGPT could stimulate students' critical thinking, probably because they had to assess the correctness of ChatGPT's answers. Regarding self-regulation, Cai et al. (2023) noted that ChatGPT allowed students to easily regulate their learning pace and style. Wu et al. (2024) found that ChatGPT-supported learning not only promoted students' self-regulation but also increased their self-efficacy. Finally, some students asserted that the use of ChatGPT enhanced the depth of their thinking and understanding (Chan & Hu, 2023). Students' achievement can serve as a proxy indicator of cognitive engagement, as it reflects their understanding (Huang et al., 2019).

Indicators of cognitive disengagement include an unwillingness to use and rejection of the technology (Bond, 2020; Bond et al., 2020). Our preliminary review of the literature suggested that the use of ChatGPT might reduce students' critical thinking and lead to an overreliance on the tool (Vargas-Murillo et al., 2023). As Cai et al. (2023) discovered, overdependence and diminished intellectual engagement were the major weaknesses of ChatGPT-supported learning. This could potentially result in a decrease in critical thinking, with students making decisions solely based on the information provided by ChatGPT (Chan & Hu, 2023).

3. Methods

3.1. Narrative review on previous systematic reviews of ChatGPT research

To avoid repeating the work of previous systematic reviews, we began with a narrative review of their major findings. In our exploration of the literature, we found 10 systematic reviews of ChatGPT research at the time of writing. In line with the approach of Roumeliotis and Tselikas (2023), we addressed RQ1 and summarised these reviews in terms of their numbers of included articles, objectives, and statements reflecting their overall conclusions. A content analysis of their findings was also conducted using the SWOT framework. This framework, utilised in the reviews by Gödde et al. (2023) and Zhang and Tur (2023), served as a useful lens for organising our current understanding of the topic.

3.2. Systematic review on empirical studies of ChatGPT

3.2.1. Search strategies

We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement (Moher et al., 2009) when selecting relevant articles. The final search was conducted on December 11, 2023. Five electronic databases were used: (1) Academic Search Ultimate; (2) Education Research Complete; (3) ERIC; (4) Scopus; and (5) Web of Science. The search string with Boolean operators was as follows: ChatGPT OR "GPT 4" OR GPT-4. We did not include other terms (e.g., "student" and "engagement") in the string to avoid filtering out relevant studies that did not use those terms. We applied this string to each database to search for relevant articles that contained either of the search terms in their title, abstract, or keywords. The publication period was specified as December 2022 to November 2023.

3.2.2. Inclusion and exclusion criteria

Journal articles published between December 1, 2022 and November 30, 2023 (i.e., within one year after the release of ChatGPT) were reviewed. In line with Korpershoek et al. (2016), we only selected articles from journals and excluded other sources (e.g., preprints and conference papers). The rigorous review processes associated with journals better assure the methodologically robust of the studies they publish (Korpershoek et al., 2016). Therefore, despite its potential to exclude some relevant studies, this criterion was necessary to facilitate the initial identification of high-quality articles. To be included in this review, the articles had to focus on the use of ChatGPT and/or its subsequent release (i.e., GPT-4) related to education. No constraints were imposed regarding educational contexts. However, the articles had to report on empirical studies that collected and analysed data from student participants to examine indicators of their engagement (i.e., the focus of this review). Although we excluded articles that were not empirical studies (e.g., commentaries, position papers, and editorials without student data), we referred to some of them as background references. Only English-language articles were included in the review. Table 1 summarises the inclusion and exclusion criteria applied when selecting the articles.

3.2.3. Data extraction and analysis

To address RQ2, we extracted data from each article on the author(s), year of publication, location, disciplines, academic levels, study types (i.e., observational or interventional), research approaches (i.e., quantitative, qualitative, and mixed methods), and data sources (e.g., surveys and interviews). We classified studies as either observational or interventional, following Thiese (2014). A study was classed as observational if it collected information to measure participants' characteristics and perspectives with no attempt (i.e., an intervention) to interfere with or change any of the measured attributes. A study was classed as interventional if it involved some type of intervention, such as incorporating ChatGPT in team-based learning (Kirpalani et al., 2023) or using ChatGPT-generated materials in teaching (Jauhiainen & Guerra, 2023). Following Creswell (2012), the interventional studies were further classified as follows.

- A pre-experimental design entails assessing a single group of participants without the inclusion of a comparison group.
- A true experimental design involves the random assignment of participants to two or more groups, thereby enabling an examination of interventional effects and a comparison across groups.
- A quasi-experimental design resembles a true experimental design, but without the random assignment of participants to different groups.

To address the research questions regarding student engagement (RQ3 to RQ5), the findings/results and discussions (or equivalent sections) of the included studies were analysed thematically using a deductive thematic approach. As described in Section 3, the general framework for the thematic analysis followed the three-dimensional model of student engagement (Fredricks et al., 2004), consisting of behavioural engagement (RQ3), emotional engagement (RQ4), and cognitive engagement (RQ5). The initial framework of the thematic analysis was established based on the seminal work of Bond and her colleagues (Bond, 2020; Bond & Bedenlier, 2019; Bond et al., 2020) and Lo and Hew (2021), as shown in Table 2. Although the framework provided a basis for analysis, we were open to refining or adding to it when new sub-themes emerged.

Among the included studies, we identified a set of observational studies (Boubker, 2024; Cai et al., 2023; Foroughi et al., 2023; Jo, 2023a; Liu & Ma, 2024; Saxena & Doleck, 2023; Strzelecki, 2023; Tiwari et al., 2023; Zou & Huang, 2023) that used the technology acceptance model (Davis, 1989) and its subsequent versions, such as the extended unified theory of acceptance and use of technology (Venkatesh et al., 2012), to inform their survey design. The researchers developed and tested their hypotheses using structural equation modelling or partial least squares structural equation modelling approaches (Hair et al., 2014) and path analysis. Hence, the direct effect of a predictor on student engagement was represented by a path coefficient. This set of studies enriched our understanding of the causes (i.e., predictors) and effects related to student engagement. As shown in Table 3, we identified three effects – actual use, satisfaction, and learning effectiveness – that were relevant to this review. Due to word limitations, the definitions of their associated predictors are provided in the appendix (Table A.1). However, the variability in the hypotheses examined across the studies and the limited number of studies ($N \leq 3$) per hypothesis made it infeasible to conduct a quantitative synthesis of their findings. Therefore, given the current state of the literature, we used descriptive statistics as a suitable means to summarise the results (Fan & Chen, 2001).

All of the articles were coded by the first author. To establish coding reliability, 25% of the studies were randomly selected and independently coded by another research team member. Interrater reliability exceeded 80%, calculated by the percent-agreement method for calculating a consensus estimate (Stemler, 2004). Any discrepancies between the data extracted and coded were

Table 1
Inclusion and exclusion criteria for article selection.

Criterion	Inclusion	Exclusion
Topic	Focused on the use of ChatGPT and/or GPT-4; education-related	Not focused on the use of ChatGPT and/or GPT-4; not education-related
Study type	Empirical studies that collected and analysed data from student participants; some indicators of student engagement reported	Non-empirical studies; no indicators of student engagement reported
Source	Journals	Sources other than journals
Period	December 1, 2022 to November 30, 2023	Articles outside the sample period
Language	English	Not English

Table 2
Sub-themes of the initial analytical framework for student engagement.

Dimension	Engagement	Disengagement
Behavioural	Participation/involvement	Absence
	Work/effort	Inattentiveness
	Study habits	Lack of preparation
	Task completion	Poor conduct
	Time on tasks	Task incompletion
Emotional	Excitement	Worry/anxiety
	Interest/fun	Disappointment
	Enjoyment	Reduced social interaction
	Satisfaction	
	Positive social interaction	
	Reduced anxiety	
Cognitive	Learning from ChatGPT	Unwillingness
	Critical thinking	Rejection
	Self-regulation	Overreliance
	Positive self-perceptions	Reduced critical thinking
	Understanding/achievement	

Table 3
Effects related to student engagement identified in the path analysis.

Dimension	Effects	Description
Behavioural	Actual use	The frequency with which students regularly work with ChatGPT as part of their learning process (Jo, 2023a; Strzelecki, 2023).
Emotional	Satisfaction	Students' affective responses derived from their learning experiences with ChatGPT (Boubker, 2024; Cai et al., 2023).
Cognitive	Learning effectiveness	The ways that technology enhances students' understanding and mastery of subjects, thereby improving their learning process and outcomes (Boubker, 2024; Cai et al., 2023).

reviewed, discussed, and resolved. Multiple reviews were conducted to ensure the consistency of coding.

4. Findings and discussion

4.1. RQ1: what are the main findings of previous systematic reviews of ChatGPT research based on SWOT analysis?

Table 4 summarises 10 systematic reviews of ChatGPT research in terms of their numbers of included articles, objectives, and statements reflecting their overall conclusions. Their objectives can be categorised into (1) medical and healthcare applications of ChatGPT (Garg et al., 2023; Gödde et al., 2023; Levin et al., 2024; Sallam, 2023), (2) roles of ChatGPT in education (Imran & Almusharraf, 2023; Lo, 2023; Vargas-Murillo et al., 2023; Zhang & Tur, 2023; İpek et al., 2023), and (3) applications of ChatGPT in different areas (Zamfiroiu et al., 2023). To provide a more structured presentation of the findings from these reviews, we compiled a list of their statements referring to the strengths, weaknesses, opportunities, and threats associated with ChatGPT (Table A.2 in appendix). The following subsections present the findings of the reviews corresponding to each aspect of this framework. After that, their limitations and unaddressed gaps are highlighted.

4.1.1. Strengths

ChatGPT exhibits multiple strengths in the areas of education and health professions. It can generate cohesive and informative human-like responses to user input (Lo, 2023), making it capable of providing students with immediate personalised feedback and enhancing their achievement (Zhang & Tur, 2023). ChatGPT is also proficient in communication and writing. It articulates ideas clearly (Gödde et al., 2023; Vargas-Murillo et al., 2023) and improves scientific writing (Sallam, 2023). Furthermore, it can translate text from one language to another, aiding students in writing papers in a non-native language to ensure accuracy and grammatical correctness (Imran & Almusharraf, 2023). Its ability to generate high-quality responses was also highlighted (Gödde et al., 2023). In their meta-analysis of ChatGPT's performance in medical examinations, Levin et al. (2024) found that it correctly answered most multiple-choice questions, demonstrating a performance level close to a passing grade.

4.1.2. Weaknesses

ChatGPT exhibits several weaknesses. The accuracy of ChatGPT's responses was occasionally questionable (Gödde et al., 2023; Imran & Almusharraf, 2023; İpek et al., 2023). Lo (2023) provided a summary of the results reported in 21 studies across subject disciplines. These studies used tests, exams, or prompts to evaluate ChatGPT's capabilities. Their results indicated that its performance varied across knowledge domains. Although its performance was outstanding in terms of critical thinking and the field of economics, it was considered barely satisfactory to unsatisfactory in the fields of law, medical education, mathematics, software testing, sports science, and psychology. The issues related to ChatGPT's inaccurate information stem from its training data, which may be biased or

Table 4
Summary of the identified systematic reviews.

Reference	No. of included articles	Major objectives	Statements reflecting overall conclusions
Garg et al. (2023)	118[NE,PP]	Explore the potential of ChatGPT in patient care and its role in medical research.	"Although it can help with patient treatment and research, there are issues with accuracy, authorship, and bias. ChatGPT can serve as a 'clinical assistant' and be a help in research and scholarly writing" (p. 183).
Gödde et al. (2023)	160[NE,PP]	Analyse the role of ChatGPT in the medical literature in terms of its strengths, weaknesses, opportunities, and threats.	"Over 400 quotes with information on strengths, weaknesses, opportunities, and threats were detected. By far, most ($n = 142$, 34.8%) were related to weaknesses" (p. 1).
Imran and Almusharraf (2023)	30	Analyse the role of ChatGPT as a writing assistant at higher education levels.	"The analysed literature identifies different opinions and scenarios associated with using ChatGPT as a writing assistant and how to interact with it" (p. 1).
İpek et al. (2023)	40	Reveal the potential effects of ChatGPT on education.	"In conclusion, this review revealed the critical applications of ChatGPT for educational settings and the potential negative impact of its application" (p. 26).
Levin et al. (2024)	19	Conduct a meta-analysis of studies reporting ChatGPT's performance in medical examinations.	"ChatGPT correctly answered the majority of multiple-choice questions in medical examinations and demonstrated a performance of approximately a passing grade" (p. 379).
Lo (2023)	50[NE,PP]	Understand ChatGPT's capabilities across subject domains, how it can be used in education, and potential issues regarding its use.	"ChatGPT's performance varied across subject domains … ChatGPT has the potential to serve as an assistant for instructors … and a virtual tutor for students … there were challenges associated with its use" (p. 1).
Sallam (2023)	60[NE,PP]	Investigate the utility of ChatGPT in healthcare education, research, and practice and highlight its potential limitations.	"The promising applications of ChatGPT can induce paradigm shifts in health care education, research, and practice. However, the embrace of this AI chatbot should be conducted with extreme caution considering its potential limitations" (p. 1).
Vargas-Murillo et al. (2023)	16	Analyse the use of ChatGPT in higher education.	"ChatGPT can potentially enhance both academic- and librarian-related processes, although it is important to reconsider the ethics of using technology such as this" (p. 122).
Zamfiroiu et al. (2023)	59[NE] + 21[NE]	Identify scenarios where ChatGPT could be used and understand how to use it appropriately.	"In the future ChatGPT could definitely improve or even replace the consulting field … However, it must be taken into account that we cannot rely 100% on the information received from ChatGPT" (p. 13).
Zhang and Tur (2023)	13	Identify the strengths, weaknesses, opportunities, and threats associated with the use of ChatGPT in K-12 education.	"ChatGPT could empower educators through curriculum, lesson planning, materials generation, differentiation, and optimising student learning experience through personalised learning. However, concerns regarding academic integrity and output quality must be addressed" (p. 1).

NE: includes non-empirical publications; PP: includes preprints.

outdated (Gödde et al., 2023; Lo, 2023; İpek et al., 2023). In addition, the free version of ChatGPT, which cannot handle images and lacks access to the Internet to verify information, has limited utility (Lo, 2023; Zhang & Tur, 2023). The generation of fake references (i.e., non-existent sources) is also a widely reported issue (Imran & Almusharraf, 2023; Lo, 2023; Sallam, 2023).

4.1.3. Opportunities

ChatGPT offers numerous opportunities, particularly in education and health professions. According to Vargas-Murillo et al. (2023), it can revolutionise how students create and process texts and assignments. Zhang and Tur (2023) further stated that it can revolutionise traditional teaching approaches. For example, ChatGPT can function as a virtual tutor, assisting students in online independent study by responding to their questions; it can also enhance group dynamics by proposing a discussion framework and providing real-time feedback (Lo, 2023; Zhang & Tur, 2023). In the field of language education, ChatGPT has the potential to create an engaging and adaptive learning environment (Imran & Almusharraf, 2023). It can assist teachers in generating assessments and learning tasks (Lo, 2023; Vargas-Murillo et al., 2023; Zhang & Tur, 2023), reducing their workload and enhancing teaching practices. Lo (2023) further stated that it has the potential to assist teachers in grading students' assessments. In health professions, ChatGPT could potentially serve as a clinical assistant and aid in the research process and medical writing (Garg et al., 2023; Zamfiroiu et al., 2023). As İpek et al. (2023) concluded, ChatGPT can play a growing role in various fields, offering significant benefits that improve human life and the world.

4.1.4. Threats

Despite the various opportunities created by ChatGPT, it also presents certain challenges that warrant attention. A primary concern is that ChatGPT may generate harmful content (Imran & Almusharraf, 2023), potentially leading to undesirable consequences (İpek et al., 2023). It may also facilitate misconduct, including plagiarism and cheating (Lo, 2023; Sallam, 2023; Vargas-Murillo et al., 2023). Given its ability to produce seemingly high-quality text, there is a risk of students misusing it for dishonest purposes, such as generating essays or coursework (Lo, 2023). Vargas-Murillo et al. (2023) and Zhang and Tur (2023) further noted that students who over-rely on

ChatGPT to complete their learning tasks may develop a dependency that hampers their productivity and intellectual growth. Unfortunately, plagiarism detection software and even experts in the field may not be able to identify whether a piece of content has been generated by ChatGPT (Gödde et al., 2023; Zamfiroiu et al., 2023). Furthermore, ethical concerns surround the use of ChatGPT, including biased data, user privacy, and data security (Garg et al., 2023; Sallam, 2023; Zhang & Tur, 2023). Therefore, researchers have called for the establishment of robust guidelines and preventative measures to ensure the responsible and ethical use of ChatGPT (Lo, 2023; Sallam, 2023; Zhang & Tur, 2023).

4.1.5. Limitations of and gaps unaddressed by previous systematic reviews

As discussed above, previous systematic reviews have enriched our understanding of the application of ChatGPT in health professions and education. However, they also exhibit certain limitations and have left knowledge gaps that need to be addressed. First, as shown in Table 4, half of the 10 reviews included non-empirical publications such as editorials, notes, or brief commentaries that were largely based on author opinions and lacked sufficient empirical data to support their claims. Second, four of the reviews included preprint articles that had not undergone rigorous peer review, casting doubt on the quality of their evidence. As acknowledged by some of the authors (e.g., Garg et al., 2023; Lo, 2023; Sallam, 2023), including those non-empirical publications and pre-prints might have affected the quality of the evidence presented in their systematic reviews. In contrast, Vargas-Murillo et al. (2023) and Zhang and Tur (2023) focused on peer-reviewed journal articles in their reviews. Despite their adoption of this rigorous approach, however, the short timeframe and limited scope of their reviews resulted in the inclusion of only a small number of articles ($N \leq 20$). This constrains the generalisability of their findings. Given the increasing number of journal articles being published on the topic, our current systematic review focused specifically on evidence-based empirical journal articles investigating the use of ChatGPT in education.

In addition to their limitations, the previous systematic reviews left significant knowledge gaps unaddressed. The reviews covered the SWOT associated with the use of ChatGPT in diverse roles such as clinical assistant, writing assistant, and virtual tutor (see Table A.2 in the appendix). However, none of these reviews examined the influence of ChatGPT on student engagement from a holistic perspective encompassing its behavioural, emotional, and cognitive aspects (Fredricks et al., 2004). This shortcoming indicates a critical knowledge gap that needs to be addressed. Student engagement can be broadly defined as "the student's active participation in academic and co-curricular or school-related activities, and commitment to educational goals and learning" (Christenson et al., 2012, p. 816). Sinatra et al. (2015) described it as "the holy grail of learning" (p. 1). Therefore, gaining insights into the influence of ChatGPT on student engagement is critical for optimising its usage in educational settings and guiding future research initiatives in ChatGPT-supported learning.

4.2. Study selection to address RQ2 to RQ5

A total of 2583 results were retrieved through the database search on December 11, 2023. Duplicate articles were removed, yielding 1888 unique records for screening. Many of the records retrieved did not involve student participants and thus were outside the scope of this review, including articles about ChatGPT performance or function ($n = 792$) and opinion papers ($n = 632$). After reviewing the titles and abstracts, we assessed 87 full-text articles for eligibility. Of these, 15 articles were excluded based on the inclusion and exclusion criteria: no data on student engagement ($n = 8$) and no student participants involved ($n = 7$). Finally, 72
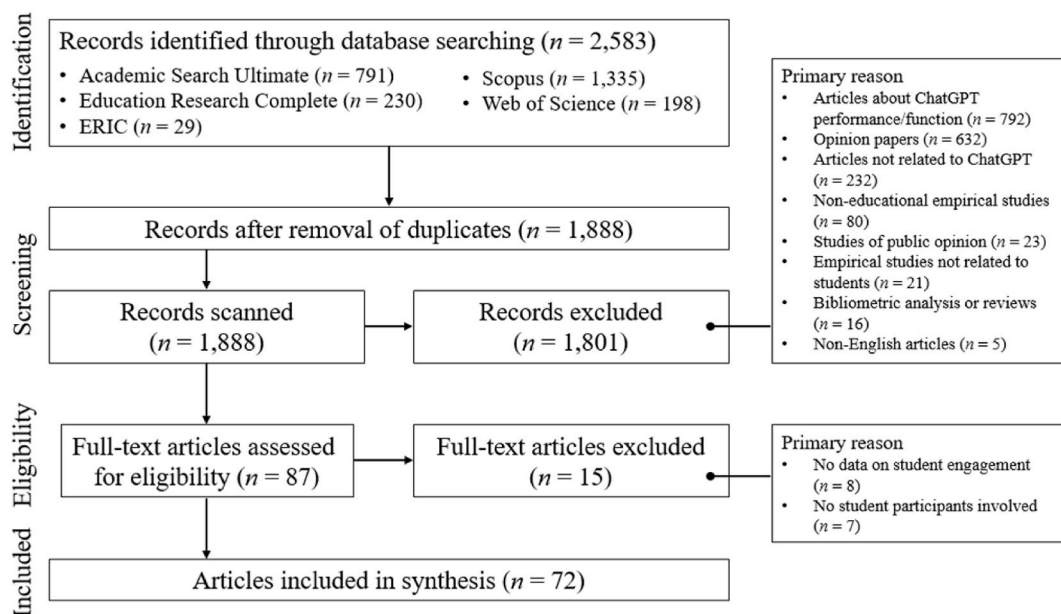

Fig. 1. PRISMA flow diagram of article selection.

articles were included in this review. Fig. 1 provides an overview of the article selection process.

4.3. RQ2: what are the characteristics of the research conducted on ChatGPT?

To address RQ2, we summarised the geographical locations of the first authors, the academic levels and subject domains considered, and the methodological approaches employed. As Fig. 2 shows, the 72 studies originated from six continents: Asia ($n = 42$; 58.3%), Europe ($n = 15$; 20.8%), North America ($n = 10$; 13.9%), Africa ($n = 2$; 2.8%), South America ($n = 2$; 2.8%), and Oceania ($n = 1$; 1.4%). There were disparities in the regions represented in the research. Notably, only a few of the studies originated from Africa, South America, and Oceania. The majority of the included studies involved students across subject disciplines ($n = 19$; 26.4%) followed by subject-specific studies, including language education ($n = 14$; 19.4%) and health professions ($n = 8$; 11.1%), among others (Fig. 3). The other included studies spanned a diverse range of subject areas, such as computer science ($n = 4$), business ($n = 2$), and mathematics ($n = 2$). The application of ChatGPT across these wide-ranging areas suggests that researchers from various fields recognise the potential of the technology to enhance student learning and engagement. However, the small number of studies published in these subject areas also reveals the need for further research to examine its impact in these under-researched disciplines. In terms of academic levels, the majority of the included studies focused on university students ($n = 57$; 79.2%). A few studies were conducted in K-12 contexts, including primary school ($n = 2$; 2.8%) and secondary school ($n = 6$; 8.3%) students. Some studies were cross-grade ($n = 4$; 5.6%), while the academic level was not reported in a few instances ($n = 3$; 4.2%). There was a substantial emphasis on university students and a lack of research involving K-12 students. This finding highlights the need for future studies to explore the application and impact of ChatGPT at these educational levels.

Table A.3 in the appendix presents a breakdown of the included studies in terms of their research approaches. The table shows that there were more observational studies ($n = 44$; 61.1%) than interventional studies ($n = 28$; 38.9%). In the observational studies, the focus was generally on investigating students' behaviours and perspectives on using ChatGPT. The number of student participants ranged from 5 (Alqasham, 2023; Xiao & Zhi, 2023) to 6311 (von Garrel & Mayer, 2023), $M = 432.52$, $SD = 983.12$. Most of the interventional studies examined the use of ChatGPT in educational settings. The majority used pre-experimental designs ($n = 18$; e.g., Chan et al., 2023; Kirpalani et al., 2023), followed by quasi-experimental designs ($n = 5$; e.g., Escalante et al., 2023; Wu et al., 2024) and true experimental designs ($n = 5$; e.g., Hsu, 2024; Yilmaz & Yilmaz, 2023). The number of student participants in the interventional studies ranged from 8 (Al-Obaydi et al., 2023) to 140 (Waltzer et al., 2023), $M = 50.43$, $SD = 35.08$. The duration of these studies varied, encompassing a few lessons or learning tasks ($n = 21$; e.g., Aktay et al., 2023; Amaro et al., 2023; Bitzenbauer, 2023), one to two months ($n = 5$; e.g., Ahmed, 2023; Javier & Moorhouse, 2023; Yilmaz & Yilmaz, 2023), and three to four months ($n = 2$; Alneyadi & Wardat, 2023; Chen et al., 2023).

In terms of research approaches, Table A.3 shows that the majority of the 72 studies were quantitative ($n = 35$; 48.6%), followed by qualitative ($n = 21$; 29.2%) and mixed methods ($n = 16$; 22.2%). Surveys ($n = 49$; 68.1%) and interviews ($n = 20$; 27.8%) were the most common data sources (Fig. 4). Few studies used more objective measures, such as achievement tests ($n = 5$; 6.9%), student work ($n = 5$; 6.9%), and ChatGPT records ($n = 2$; 2.8%). As the numbers indicate, relatively few of the studies utilised a mixed-methods approach. We recommend that this approach be used more frequently in future research because it facilitates a deeper understanding of quantitative results. For example, in their analysis of students' essays, Bašić et al. (2023) discovered that some students were likely to have incorporated text produced by ChatGPT. However, the investigation could have been enhanced if the students had been interviewed. This additional step would have afforded a deeper understanding of their experiences and shed light on the reasons behind the identified issue.
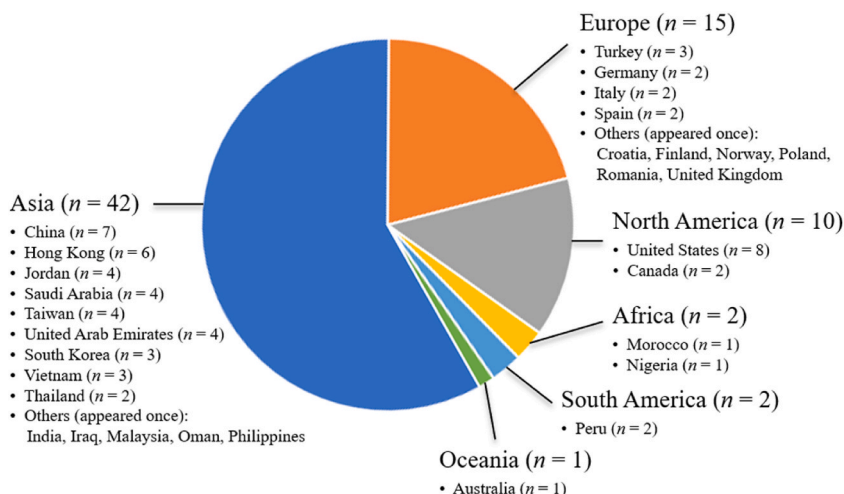

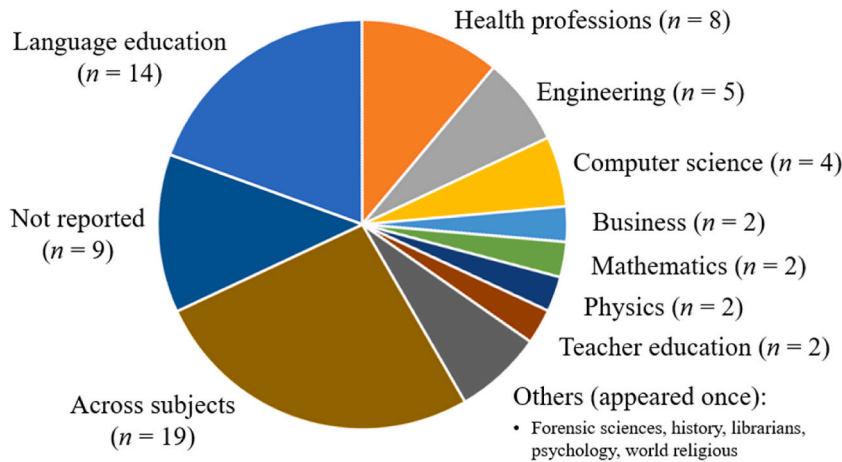**Fig. 2.** Locations of first authors of the included studies ($n = 72$).

Fig. 3. Subject disciplines of the included studies ($n = 72$).
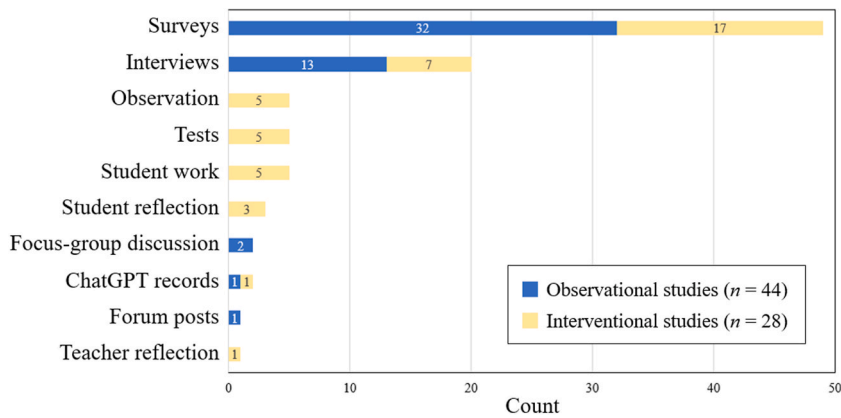

Fig. 4. Data sources used in the included studies ($n = 72$). Note: The sum of the number of studies across data sources is greater than 72 because some studies used multiple data sources.

4.4. Answers to RQs related to student engagement and future research agenda

Indicators of engagement and disengagement were identified across behavioural, emotional, and cognitive aspects (Fig. 5). Relatively few indicators were associated with behavioural engagement ($n = 3$) and disengagement ($n = 1$). Six indicators were identified for emotional engagement and three indicators for disengagement. Cognitive engagement and disengagement were reflected by five and two indicators, respectively.

4.4.1. RQ3: how does ChatGPT influence students' behavioural engagement? Robust but narrow evidence

Table 5 summarises the major indicators of students' behavioural engagement and disengagement in the 72 studies. The evidence extracted from the studies consistently highlighted students' active involvement in working with ChatGPT ($n = 50$; the highest frequency among all indicators; e.g., Adetayo, 2023; Duong, Bui, et al., 2023; Limo et al., 2023; Romero-Rodríguez et al., 2023; Wandelt et al., 2023). For example, students used ChatGPT as a tool to enhance their assignment drafts (Barrett and Pack, 2023; Boubker, 2024; Dai et al., 2023) and generate ideas (Alqasham, 2023; Barrett and Pack, 2023; Chan & Hu, 2023). Table 6 shows that students' ChatGPT usage was influenced by personal factors: behavioural intention (Jo, 2023a; Liu & Ma, 2024; Strzelecki, 2023) and habit (Strzelecki, 2023); factors related to ChatGPT: perceived usefulness and output quality (Boubker, 2024); and other external factors: facilitating conditions (Strzelecki, 2023) and social influence (Boubker, 2024). Furthermore, using ChatGPT could increase the likelihood of students completing their learning tasks ($n = 17$; e.g., Ahmed, 2023; Bitzenbauer, 2023; Chen et al., 2023). As some teachers observed, ChatGPT enhanced students' punctuality in the submission of their learning tasks (Hasanein & Sobaih, 2023).

The negative consequences associated with the use of ChatGPT, particularly concerning plagiarism and cheating ($n = 18$; e.g., Alshurafat et al., 2024; Singh et al., 2023), were well documented in the included studies. These consequences were directly evident from their empirical results and were reinforced by the perspectives of both teachers and students. For example, Bašić et al. (2023) used an AI detector to check whether students' writing assignments were human- or AI-generated. Their results indicated that text
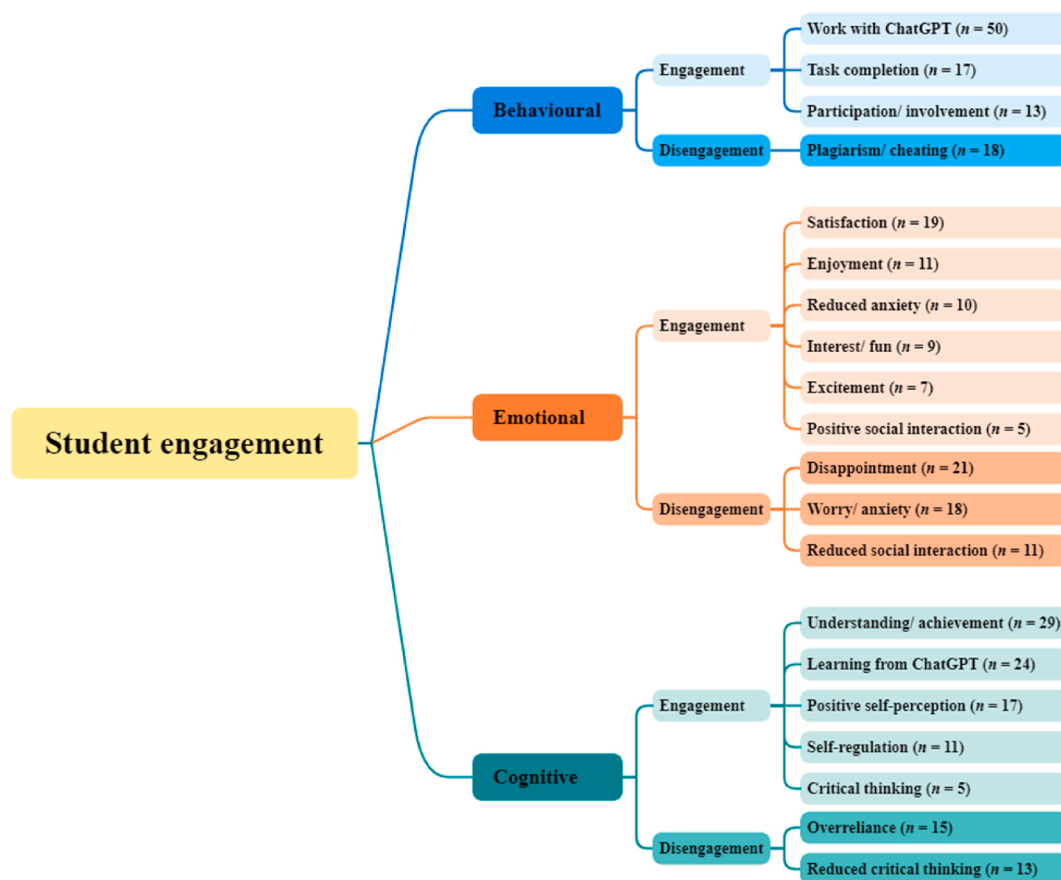
Fig. 5. Engagement and disengagement indicators identified across the 72 studies.

Table 5
Major indicators of behavioural engagement across the 72 studies.

Themes and sub-themes	Count	Representative quotes
Engagement		
Work with ChatGPT	50	Alqasham (2023): "[A student] Emphasizes ChatGPT's role in kickstarting the writing process, acting as a guide during the brainstorming phase" (p. 1256).
Task completion	17	Lozano and Blanco Fontao, 2023: "The most beneficial aspects of using ChatGPT (question 5) identified by students are the speed in completing tasks …" (p. 5).
Participation/ involvement	13	Hamid et al. (2023): "the use of ChatGPT was found to enhance students' willingness to actively participate in PBL activities" (p. 1021).
Disengagement		
Plagiarism/cheating	18	Bašić et al. (2023): "The ChatGPT group had three possible and five cases [out of nine students' essays] likely produced by AI, while one case was labeled as unclear" (p. 3).

potentially generated by AI was more prevalent in the assignments of students using ChatGPT than in the assignments of those not using it. In the study by Fuchs and Aguilos (2023), some students confessed to copying and pasting text generated by ChatGPT into their essays without providing appropriate citations. This behaviour could be attributed to their inclination towards laziness. In the words of one student, "When I feel lazy, I just do it online (with the chatbot) to save time" (p. 1368). Echoing some previous reviews (e. g., Lo, 2023; Sallam, 2023; Zhang & Tur, 2023), several of the included studies (e.g., Bernabei et al., 2023; Chan, 2023; Gorichanaz, 2023; Sallam et al., 2023; Tlili et al., 2023) stressed the need to establish guidelines and policy regulating the ethical use of AI tools in educational settings. Furthermore, considering that students' ChatGPT usage is influenced by facilitatory conditions and social influence (Table 6), it is essential for educators to (1) inform students about the use of plagiarism detection software to monitor their work, as regular checks could deter them from misusing AI tools; and (2) foster peer influence through open discussions, thus leveraging the power of social influence by encouraging students to hold each other accountable for maintaining academic integrity.

Despite the robust evidence demonstrating the influence of ChatGPT on students' behavioural engagement, several crucial indicators were absent from the included studies. First, "study habits," a major indicator of behavioural engagement in technology-

Table 6
Effects associated with behavioural engagement and their causes across studies using path analysis ($n = 9$).

Effect ← Cause	No. of studies	Cumulative sample size	Average sample size[a]	Range of path coefficients		No. with sig. effect	No. with non-sig. effect	% sig.
				Lower[b]	Upper[a]			
Actual use								
← Behavioural intention	3	1171	390	0.42	0.77	3	0	100
← Habit	1	534		0.25		1	0	100
← Facilitating conditions	1	534		0.18		1	0	100
← Social influence	1	319		0.34		1	0	100
← Perceived usefulness	1	319		0.26		1	0	100
← Output quality	1	319		0.13		1	0	100

Note: The sum of the numbers of studies did not add up to nine because some studies examined multiple causes, and some studies did not focus on the effects associated with behavioural engagement.
   [a] For multiple studies only.
   [b] Value of path coefficient for single study.

enhanced learning (Bond, 2020; Bond et al., 2020), remain inadequately addressed. Although a few studies (e.g., Boubker, 2024; Dai et al., 2023; Hasanein & Sobaih, 2023) noted ChatGPT's potential assistance in exam preparation and practice scenarios, the extent to which it shapes students' study habits remains unclear within the current literature. Second, the "time on task" indicator lacks comprehensive exploration. Among the included studies, only Hsu (2024) used students' weekly study time as an objective measure of their learning behaviour. In her undergraduate nursing course, she found that students who used ChatGPT generally spent more time ($n = 20$; $M = 10$ h) studying medical terminology than those who did not use it ($n = 20$; $M = 2$ h). She explained that incorporating ChatGPT as a learning tool could create a more engaging learning environment and enhance student motivation. However, further research is needed to investigate the influence of ChatGPT on shaping students' study habits and their time allocation on learning tasks, thereby broadening our understanding of its influence on behavioural engagement.

Lastly, there is a lack of studies concerning students' attendance and attention in class. Although some students expressed views suggesting ChatGPT as a potential substitute for traditional teaching methods, it is important to consider the broader implications. For example, one student in the study by Xiao and Zhi (2023) visited the school's language centre less frequently after using ChatGPT. Chan and Lee (2023) further reported that "some participants did believe that GenAI usage would hinder or even defeat the purpose of teaching and learning in certain areas" (p. 14). The use of ChatGPT may reduce the need for students to be physically present in school. Nevertheless, this reduction in physical presence may not necessarily be negative. If ChatGPT or other AI tools can optimise learning and enable students to spend less time in school without compromising on their learning or development, this shift could be seen as positive. Furthermore, in online learning environments, AI tools could serve as virtual tutors and provide valuable support (Chan & Hu, 2023). Therefore, it is important to further investigate the influence of ChatGPT on students' class attention and attendance to fully understand its potential benefits and drawbacks in various learning environments.

4.4.2. RQ4: how does ChatGPT influence students' emotional engagement? Mixed and inconclusive findings
The findings from the included studies revealed a mixed landscape regarding students' emotional engagement (Table 7). While

Table 7
Major indicators of emotional engagement across the 72 studies.

Themes and sub-themes	Count	Representative quotes
Engagement		
Satisfaction	19	Shaikh et al. (2023): "The satisfaction score from the participants showed a positive experience with ChatGPT" (p. 1952).
Enjoyment	11	Elkhodr et al. (2023): "ICT students in Case Studies 1 and 2 generally reported enjoying the experience of using ChatGPT" (p. 75).
Reduced anxiety	10	Rad and Rad (2023): "others [students] saw chatbots as a helpful tool for improving communication and reducing social anxiety" (p. 50).
Interest/fun	9	A student of Uddin et al. (2023): "Thank you, Prof. [Name] for introducing us to this tool [ChatGPT] … This was a fun and unique class" (p. 14).
Excitement	7	A student of Shoufan (2023): "It is absolutely amazing what this technology [ChatGPT] is capable of" (p. 38810).
Positive social interaction	5	Hamid et al. (2023): "Students appreciated how ChatGPT stimulated meaningful discussions and allowed for the exploration of diverse perspectives of information" (p. 1021).
Disengagement		
Disappointment	21	Ahmed (2023): "43 of the 64 participants reported not being satisfied with ChatGPT in their EFL writing class" (p. 2730).
Worry/anxiety	18	Chan and Lee (2023): "Many [students and teachers] also expressed worries about the larger impact of GenAI on the job market and society" (p. 14).
Reduced social interaction	11	Hasanein and Sobaih (2023): "The use of ChatGPT has a negative impact on educational support and may reduce social interaction between students and faculty" (p. 2608).

students across 19 studies expressed satisfaction (e.g., Hu et al., 2023; Jo, 2023b; Sahari et al., 2023) with their experiences using ChatGPT, disappointment was reported in 21 studies. As shown in Table 8, the factors that predicted student satisfaction included perceptions of ChatGPT's usefulness, ease of use, and the quality of the output and information system (Boubker, 2024; Cai et al., 2023). The inconsistent performance of ChatGPT (Lo, 2023) thus explains the mixed findings. For example, students in the study by Cai et al. (2023) described ChatGPT as providing "very good language learning experiences" alongside "disappointing accuracy and credibility" issues (p. 8). Furthermore, several studies provided evidence of a reduction in students' anxiety ($n = 5$; e.g., Firat, 2023) because "ChatGPT provides a low-pressure environment where they can seek clarification and assistance without fear" (Hasanein & Sobaih, 2023, p. 2607). However, in other instances ($n = 18$), anxiety was caused by concerns about ChatGPT's privacy and confidentiality (Chan et al., 2023; Menon & Shilpa, 2023) and the potential job displacement by AI (Chan & Hu, 2023; Shoufan, 2023). These diverse findings highlight the need for continual effort to enhance ChatGPT's capabilities and reinforce privacy safeguards. Such improvements have the potential to foster students' emotional engagement and trust in AI-based educational tools.

The examination of social interaction, another indicator of students' emotional engagement, presented contrasting results regarding ChatGPT usage within the included studies. Table 7 shows that reduced social interaction emerged as a concern in 11 studies (e.g., Kevin & Veronica, 2023; Wardat et al., 2023), highlighting a potential drawback associated with ChatGPT usage. The insights gathered from these studies indicate the loss of personal interactions among students and teachers as the reason for this concern (Fuchs & Aguilos, 2023; Hasanein & Sobaih, 2023; Limna et al., 2023). Nevertheless, five studies reported a positive effect on social interaction because ChatGPT stimulates collaboration and meaningful discussion among students (Hamid et al., 2023; Kirpalani et al., 2023). Furthermore, ChatGPT helped students solve basic learning problems and thus created space for more intellectual exchanges between teachers and students (Dai et al., 2023). These mixed findings emphasise the need for further research to develop a comprehensive approach that leverages ChatGPT as an in-class facilitator of student discussions and an out-of-class virtual tutor.

Notably, the findings from the included studies only enable a preliminary understanding of ChatGPT's influence on students' emotional engagement. Although they provide evidence indicating that the use of ChatGPT can elicit positive affective responses (i.e., enjoyment, $n = 11$, e.g., Muñoz et al., 2023; interest or fun, $n = 9$, e.g., Muñoz et al., 2023; excitement, $n = 7$, e.g., Uddin et al., 2023, Table 7), most of the studies were of short duration. Hence, students' responses might have been influenced by a novelty effect. In the study by Uddin et al. (2023), for example, many students expressed enthusiastic feedback, such as "ChatGPT is awesome" and "The future is exciting" (p. 14), probably because they were first-time users of ChatGPT. These findings suggest a potential short-term boost in engagement due to the introduction of new technology (Clark, 1983). Therefore, further studies with a longer duration (e.g., one semester) are necessary to explore the influence of ChatGPT on students' emotional engagement over time.

4.4.3. RQ5: how does ChatGPT influence students' cognitive engagement? Broad but weak evidence

Table 9 summarises the major indicators of students' cognitive engagement and disengagement identified in the 72 studies. However, the evidence regarding some indicators of cognitive engagement was not robust because there were few interventions over a sufficiently long period that objectively measured those indicators. Specifically, the findings related to understanding or achievement ($n = 29$; e.g., Ho & Lee, 2023; Ibrahim et al., 2023; Vecchiarini & Somià, 2023) through ChatGPT usage largely relied on short-term interventions or self-reported data (e.g., interviews and surveys). For example, Kirpalani et al. (2023) incorporated ChatGPT to facilitate students' team-based learning in one medicine lesson. A subsequent survey of the students suggested that ChatGPT fostered a deeper understanding of clinical cases. Similarly, Chan et al. (2023) introduced the use of ChatGPT to nursing students in a tutorial session. In the words of one student, "It provided an excellent guide to enhance my knowledge" (p. 201). However, the durability and reliability of these observations over longer periods remain uncertain and follow-up studies are required. Furthermore, Elkhodr et al. (2023) highlighted the potential limitations of relying solely on subjective measures because they are vulnerable to response bias, social desirability bias, and other factors that may affect students' responses. Further research should use more objective measures,

Table 8
Effects associated with emotional engagement and their causes across studies using path analysis ($n = 9$).

Effect ← Cause	No. of studies	Cumulative sample size	Average sample size[a]	Range of path coefficients		No. with sig. effect	No. with non-sig. effect	% sig.
				Lower[b]	Upper[a]			
Satisfaction								
← Perceived usefulness/ performance expectancy	2	777	389	0.32	0.43	2	0	100
← Output/information system quality	2	777	389	0.18	0.26	2	0	100
← Hedonic motivation	1	458		0.24		1	0	100
← Self-regulation	1	458		0.08		1	0	100
← Perceived ease of use	1	319		0.21		1	0	100
← Actual use	1	319		0.16		1	0	100
← Disconfirmation of expectations from prior use	1	106		0.79		1	0	100

Note: The sum of the numbers of studies did not add up to nine because some studies examined multiple causes, and some studies did not focus on the effects associated with emotional engagement.
[a] For multiple studies only.
[b] Value of path coefficient for single study.

Table 9

Major indicators of cognitive engagement across the 72 studies.

Themes and sub-themes	Count	Representative quotes
Engagement		
Understanding/ achievement	29	Alneyadi and Wardat (2023): "The results indicate a clear improvement in the post-test scores for the experimental group compared to the control group, demonstrating the effectiveness of ChatGPT in enhancing students' learning outcomes" (p. 9).
Learning from ChatGPT	24	Hosseini et al. (2023): "Another mentioned possibility was to use ChatGPT as a studying tool that (upon further improvements and approved accuracy) could describe medical concepts at a specific comprehension level" (pp. 6–7).
Positive self-perception	17	Foroughi et al. (2023): "This personalised approach [with ChatGPT] will likely improve students' performance expectancy" (p. 11).
Self-regulation	11	Cai et al. (2023): "ChatGPT allows students to easily self-regulate their learning pace and style" (p. 8).
Critical thinking	5	Dai et al. (2023): "The analyses of students' mock interactions and resource searching with ChatGPT show that ChatGPT, to some extent, was able to complement the supervisor's role in helping students develop critical thinking" (p. 81).
Disengagement		
Overreliance	15	Chen et al. (2023): "Some [students] fully relied on it to complete their homework" (p. 7).
Reduced critical thinking	13	A student of Chan and Hu (2023): "this may lead to a decrease in critical thinking and make decisions only based on the information that AI provides to them" (p. 11).

such as the quality of learning tasks and test performance, to serve as a more reliable proxy indicator of students' cognitive engagement (Huang et al., 2019).

In addition to incorporating objective measures and extending the duration of studies, future interventions of ChatGPT-supported learning should include student training on how to use ChatGPT. Table 10 shows that students' perceived usefulness of ChatGPT predicted their learning effectiveness (Boubker, 2024; Cai et al., 2023). However, students' actual use of ChatGPT did not necessarily predict their learning effectiveness (Boubker, 2024). One explanation for this discrepancy is that ChatGPT cannot effectively enhance students' learning if they do not use it skilfully. Their level of familiarity with ChatGPT, and therefore their skill in using it, may influence how their use of the technology enhances learning outcomes (Singh et al., 2023). Bašić et al. (2023) further noted that the quality and depth of students' prompts to ChatGPT significantly influence its output. In future interventions, students should be introduced to prompt strategies (Wang et al., 2023), such as defining a clear role for ChatGPT with a specific goal and additional information about response requirements. In the study by Dai et al. (2023), for example, doctoral students asked ChatGPT to act in a thesis supervisor role and consulted it on how to express a particular idea (goal) in a humble yet critical manner in academic writing (response requirements). Introducing students to these techniques enables them to use ChatGPT more effectively to support their learning.

Reports regarding the influence of ChatGPT on students' self-regulation ($n = 11$) were often too general due to a lack of specificity about which particular phases of self-regulation were enhanced. According to Zimmerman (2002), there are three phases in the self-regulated learning process: forethought, performance, and self-reflection. In the forethought phase, students should analyse the learning tasks, set specific learning goals, and devise strategies to achieve these goals. ChatGPT potentially offers guidance to assist students in attaining their goals (Ajlouni, Almahaireh, & Whaba, 2023a; Ajlouni, Almahaireh, & Whaba, 2023b). During the performance phase, ChatGPT can facilitate a progressive learning journey (Alqasham, 2023) and aid students in resolving learning problems. Finally, students can solicit ChatGPT's feedback on their learning outcomes (Chan & Lee, 2023; Fuchs & Aguilos, 2023), thereby promoting their self-reflection (Ajlouni, Almahaireh, & Whaba, 2023a,b). To strengthen the evidence in this area, further research should articulate the precise aspects or methods that enhance self-regulation. Such a detailed examination would provide more robust evidence regarding the influence of ChatGPT on the various phases of students' self-regulation.

5. Conclusion, limitations, and recommendations for future studies

In this systematic review, we synthesised the findings from 72 empirical studies related to ChatGPT within the first year of its release. The review offers initial yet comprehensive insights into the influence of ChatGPT on students' behavioural, emotional, and cognitive engagement. Our findings indicate both positive and negative influences in ChatGPT-supported learning environments.

● From the behavioural perspective, many students have actively worked with ChatGPT in their learning activities. However, the misuse of this tool, often stemming from insufficient literacy, knowledge, and understanding of its proper use, might lead to undesirable behaviours, such as plagiarism and cheating. Therefore, teachers and stakeholders should prioritise improving students' digital literacy to ensure they understand the ethical use of AI tools and develop clear guidelines on their acceptable use. The recommendations offered by European Network for Academic Integrity may be useful to assist academics and students in the ethical use of AI tools (Foltynek et al., 2023).

● The evidence regarding emotional engagement was mixed, with varied reports of student satisfaction and disappointment concerning the accuracy and credibility of ChatGPT's responses, as well as diverse impacts on social interaction. Teachers and stakeholders should be aware of the emotional impact of ChatGPT on students and provide emotional support, fostering an environment where students can express their concerns and seek advice. In addition, instructional designers may consider using Retrieval-Augmented Generation which integrates external sources of verifiable facts into the chatbot generative processes, thereby improving the relevance and accuracy of its responses (Chen et al., 2024).

Table 10
Effects associated with cognitive engagement and their causes across studies using path analysis ($n = 9$).

Effect ← Cause	No. of studies	Cumulative sample size	Average sample size[a]	Range of path coefficients		No. with sig. effect	No. with non-sig. effect	% sig.
				Lower[b]	Upper[a]			
Learning effectiveness								
← Perceived usefulness/ performance expectancy	2	777	389	0.33	0.35	2	0	100
← Satisfaction	2	777	389	0.07	0.50	1	1	50
← Behavioural intention	1	458		0.51		1	0	100
← Actual use	1	319		0.08		0	1	0

Note: The sum of the numbers of studies did not add up to nine because some studies examined multiple causes, and some studies did not focus on the effects associated with cognitive engagement.

[a] For multiple studies only.

[b] Value of path coefficient for single study.

● From the cognitive perspective, students were able to learn from ChatGPT, which increased their understanding and achievement. However, concerns were raised that the growing use of AI tools might lead to a decline in critical thinking among students. Teachers and stakeholders should continue to investigate pedagogical approaches that leverage ChatGPT to enhance students' understanding and critical thinking. For example, teachers can instruct students to fact-check and validate information produced by ChatGPT (Chan & Lee, 2023). This practice would help students develop a critical mindset by cross-referencing and evaluating the quality and relevance of ChatGPT's responses (Chen et al., 2023).

In addition to the review findings, we identified several areas that require further investigation. For example, the influence of ChatGPT on students' study habits (behavioural engagement), social interaction (emotional engagement), and self-regulation (cognitive engagement) are relatively unexplored. However, we recommend further studies be conducted over longer durations (e. g., one semester) to mitigate the effects of novelty and unfamiliarity with ChatGPT that students might initially experience. Future research would also benefit from the implementation of experimental designs and more objective measures. For example, assessing students' understanding through tests and tracking their study hours would provide quantifiable data to further enrich our understanding of ChatGPT-supported learning. Last but not least, in light of the rapid advancements in AI technology, the capabilities of ChatGPT are likely to have further improved since the time of writing. Future research should continue to evaluate these evolving capabilities in promoting student engagement and address any concerns associated with the use of ChatGPT in educational settings. In this regard, we recommend adopting the mixed-methods approach. This research approach enhances our understanding of quantitative results (e.g., students' assignment scores) via the incorporation of information obtained using qualitative methods (e.g., student interviews), and thereby uncovers the actual reasons behind the observations made in a study.

Finally, several limitations of this review must be acknowledged. We list them here with recommendations for future studies to address them.

● Aronson and Laughter (2016) noted that "no research synthesis is exhaustive" (p. 178) and acknowledged the potential for relevant research articles to be overlooked. In this review, our selection of articles was limited to peer-reviewed journal articles. Thus, our analysis may not have considered certain relevant studies published in conference proceedings. Future reviews could include research output from diverse sources (e.g., conference papers and book chapters) to widen their coverage of the research conducted on the topic.
● Our research synthesis could only focus on what the authors reported in their articles. The absence of an engagement indicator does not necessarily imply its absence in educational settings; rather, it merely indicates that the indicator was not explicitly reported in the articles. Furthermore, some of the studies lacked clear details on the precise context of the use of ChatGPT, such as whether it was used in formal or informal education settings and whether it was used within or outside the classroom. The absence of this information made it challenging to discern potential differences in student engagement across these contexts. We recommend that future research explicitly state the context in which ChatGPT is used in the study, thereby enhancing our understanding of its effects on student engagement in diverse learning environments.
● The majority of the reviewed studies were conducted in the language education and health professions and in higher education settings. The findings of this review may thus be biased towards these specific contexts. Further studies in other subject disciplines (e.g., computer science, business, and mathematics) and educational contexts (e.g., K-12 schools) are recommended.
● Our review only included studies published within one year of ChatGPT's initial release. The findings of these studies might have been influenced by (1) novelty effects, potentially resulting in a short-term boost in student engagement, or (2) students' unfamiliarity with ChatGPT usage, which could have hindered their engagement in learning. These factors may limit the generalisability of our findings. Future reviews could focus on studies with longer durations and examine their effect of using ChatGPT on student engagement.

CRediT authorship contribution statement

Chung Kwan Lo: Writing – review & editing, Writing – original draft, Visualization, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization, Funding acquisition. **Khe Foon Hew:** Writing – review & editing, Validation. **Morris Siu-yung Jong:** Writing – review & editing.

Declaration of competing interest

The authors declares that he has no conflict of interest.

Data availability

Data will be made available on request.

Acknowledgments

Appendix

Table A.1
Definitions of the predictors of the identified effects.

Predictor	Definition	Supporting studies
Actual use	The frequency with which students regularly work with ChatGPT as a part of their learning process.	Jo (2023a); Strzelecki (2023)
Behavioural intention	Students' plans and intentions to adopt and utilise ChatGPT as a learning tool in their studies.	Cai et al. (2023); Foroughi et al. (2023)
Disconfirmation of expectations from prior use	Instances where students' perceptions of ChatGPT's performance exceed their initial expectations.	Saxena and Doleck (2023)
Facilitating conditions	The extent to which students believe that the necessary resources and support are available for effective use of ChatGPT.	Strzelecki (2023)
Habit	The extent to which students have incorporated the regular and consistent use of ChatGPT into their academic routine.	Strzelecki (2023)
Hedonic motivation	The extent to which students find using ChatGPT for learning fun, entertaining, or enjoyable.	Cai et al. (2023); Foroughi et al. (2023); Strzelecki (2023)
Output/information system quality	The degree to which the output or information provided by ChatGPT aligns with the expected meaning.	Boubker (2024); Cai et al. (2023)
Perceived ease of use	The degree to which students believe that using ChatGPT will be free of effort.	Boubker (2024); Zou and Huang (2023)
Perceived usefulness/ performance expectancy	Students' perceptions of how ChatGPT contributes to achieving their learning goals and their evaluation of the benefits they receive from using ChatGPT. Note: According to Venkatesh et al. (2003), perceived usefulness is similar to performance expectancy.	Boubker (2024); Cai et al. (2023)
Satisfaction	Students' affective responses derived from their learning experience with ChatGPT.	Boubker (2024); Cai et al. (2023)
Self-regulation	The process through which students initiate and sustain behaviours, thoughts, and emotions related to the use of ChatGPT.	Cai et al. (2023)
Social influence	The extent to which students believe that their peers, teachers, or significant others in their social surroundings support or encourage them to use ChatGPT.	Boubker (2024); Foroughi et al. (2023); Strzelecki (2023)

Table A.2
Summary of previous systematic reviews in terms of the major strengths, weaknesses, opportunities, and threats associated with ChatGPT.

Reference	Strengths	Weaknesses	Opportunities	Threats
Garg et al. (2023)	● ChatGPT can assist with patient enquiries, note writing, decision-making, trial enrolment, data management, decision support, research support, and patient education.	● Solutions offered by ChatGPT are often insufficient and contradictory.	● ChatGPT can serve as a clinical assistant and help with research and scholarly writing. ● ChatGPT has the potential to serve as a personalised learning tool, encouraging critical thinking and problem-based learning among medical professionals.	● There are potential biases in ChatGPT's training data and also accuracy and reliability issues, privacy concerns, questions about authorship in academic papers, and ethical issues in its use.

Table A.2 (*continued*)

Reference	Strengths	Weaknesses	Opportunities	Threats
Gödde et al. (2023)	● ChatGPT excels at expressing ideas clearly and formulating general contexts comprehensibly. ● ChatGPT can assist in formulating medical issues for non-native English speakers.	● ChatGPT's responses are based on data from a limited time period and its responses may therefore require correction by experts.		● ChatGPT performs so well that even experts in the field have difficulty identifying abstracts generated by it.
Imran and Almusharraf (2023)	● ChatGPT can improve scientists' efficiency in writing review articles and enhance the quality of the articles. ● It has the ability to generate coherent and well-structured text on any topic, saving time and effort for students and educators. ● It can help students generate new ideas for their writing assignments by suggesting topics, themes, and perspectives they might not have considered otherwise. ● It can assist with various tasks such as writing and coding and can answer questions and guide discussions. While doing so, it can provide examples and follow up on being corrected. ● It can translate text from one language to another, aiding students writing papers in a non-native language to ensure accuracy and grammatical correctness.	● ChatGPT presents incorrect information and non-existent references. ● ChatGPT has limitations resulting from its lack of access to latest information.	● ChatGPT has the potential to promote engaging and adaptive language learning. ● It can revolutionise and challenge existing writing techniques and domains, offering opportunities for enhanced productivity and efficiency.	● ChatGPT may provide biased, incorrect, and harmful content.
İpek et al. (2023)	● ChatGPT can provide information, answer questions, and engage in conversations on various subjects.	● ChatGPT may produce incorrect answers and has limited knowledge of events beyond 2021.	● ChatGPT can play a growing role in education, medicine, and art as technology advances, offering significant benefits that improve human life and the world.	● ChatGPT has an algorithmic bias in its training data. ● It lacks independent ethical perception and relies on the ethical judgments of its programmers, which can lead to undesirable or objectionable responses. ● It may provide inappropriate responses, potentially leading to harmful consequences or causing fear.
Levin et al. (2024)	● ChatGPT demonstrates high accuracy in answering multiple-choice questions in medical examinations.	● The dataset used to train ChatGPT is not specifically focused on medical education.	● ChatGPT has the potential to become an important tool in medical education due to its ability for human-like conversation and access to a large quantity of data.	
Lo (2023)	● ChatGPT's performance is outstanding in economics. ● It is capable of generating cohesive and informative human-like responses to user input. ● It can aid in improving writing and research skills. ● Initial answers from ChatGPT can prompt further questioning and encourage students to apply their knowledge and reasoning skills.	● ChatGPT's performance is unsatisfactory in mathematics and software testing. ● It has limited up-to-date knowledge. ● It generates incorrect or fake information, including non-existent references and non-functional URLs.	● ChatGPT can serve as an assistant for instructors to generate course materials and provide suggestions. ● It can act as a virtual tutor for students, answering questions and facilitating collaboration.	● There are concerns about AI-assisted cheating, as students may use ChatGPT to complete written assignments and examinations on their behalf.
Sallam (2023)	● ChatGPT can improve scientific writing and enhance research equity and versatility.	● ChatGPT has limited knowledge.	● ChatGPT may contribute to a paradigm shift in healthcare	● There are concerns regarding bias, plagiarism, cybersecurity, ethical,

Table A.2 (*continued*)

Reference	Strengths	Weaknesses	Opportunities	Threats
	● It can efficiently analyse datasets, generate code, and conduct literature reviews in healthcare research. ● It can streamline workflows, save cost, and help documentation in healthcare practice. ● It can improve personalised learning, critical thinking, and problem-based learning in healthcare education.	● It may generate inaccurate content and incorrect citations.	education, research, and practice.	copyright, transparency, and legal issues.
Vargas-Murillo et al. (2023)	● ChatGPT is capable of creating text, processing data, and developing curricula. ● It complements students' learning processes by generating text responses based on knowledge gathered through machine learning. ● It can enhance academic and library management processes in higher education. ● It can generate ideas and arguments for further research, facilitate continuing education, and enhance teaching and learning processes.		● ChatGPT has revolutionised education by expanding the possibilities of what students and researchers can do to create and process texts and assignments.	● The use of ChatGPT may hinder the development of critical thinking skills in students, as it can provide answers to problems without encouraging independent thought. ● There are concerns about academic integrity and plagiarism resulting from potential overreliance on ChatGPT for completing assignments.
Zamfiroiu et al. (2023)	● ChatGPT can be used to predict diseases and explain medical concepts to users in a more accessible manner.		● ChatGPT has the potential to personalise learning experiences, assist in lesson planning, aid in language learning, support research and writing, facilitate professional development, and enable assessment and evaluation. ● It presents opportunities for improving medical writing and enhancing the research process, allowing researchers to allocate more time to the practical aspects of their work.	● ChatGPT generates text that cannot be identified by plagiarism software.
Zhang and Tur (2023)	● ChatGPT can be used across various subjects. ● It can enhance students' learning experience. ● It has a positive influence on student achievement. ● It can provide immediate and personalised feedback.	● ChatGPT has limited contextual understanding and reasoning abilities. ● It has technical and operational limitations, such as slow response times and lack of access to the Internet to verify facts.	● ChatGPT can empower educators through curriculum design, lesson planning, and materials generation. ● It can optimise students' learning experience by supporting personalised learning.	● There are concerns regarding academic integrity and output quality. ● Overreliance on ChatGPT may be detrimental to productivity.

Table A.3
Types of studies included (*n* = 72) and their research approaches.

Research approach	Observational studies		Interventional studies		Total
	Count	Representative study	Count	Representative study	
Quantitative	25	von Garrel and Mayer (2023) conducted a nationwide survey involving 6311 German university students to investigate the use of AI-based tools, particularly ChatGPT and GPT-4, in academic settings. Their findings revealed that almost two thirds of the participants had used such tools primarily for clarifying their understanding and explaining subject-specific concepts.	10	Hsu (2024) examined the impact of AI tools on nursing students' understanding of medical terminology. The study involved experimental groups using ChatGPT (*n* = 20) or Termbot (*n* = 20) and a control group using traditional textbooks (*n* = 20). The post-test results showed a significant improvement in learning outcomes for both experimental groups.	35

Table A.3 (*continued*)

Research approach	Observational studies		Interventional studies		Total
	Count	Representative study	Count	Representative study	
Qualitative	13	Xiao and Zhi (2023) explored students' perceptions of using ChatGPT for language learning tasks. Through semi-structured interviews with five students, they found that ChatGPT could serve as a learning partner, offering immediate feedback and personalised learning experiences. Nevertheless, students displayed critical judgment when evaluating ChatGPT's outputs.	8	In Yan's (2023) study, students participated in a one-week practicum applying ChatGPT's text generation feature to second language (L2) writing. The interview findings suggested that ChatGPT improved writing efficiency and was potentially applicable in L2 pedagogy. However, concerns about academic honesty and educational equity were also raised.	21
Mixed methods	6	Ngo (2023) investigated university students' perceptions of using ChatGPT for learning. Her survey of 200 students and interviews with 30 students revealed a generally favourable view of ChatGPT for its time-saving benefits, wide-ranging information, personalised tutoring, and idea generation. However, concerns were raised about source quality, citation accuracy, and word replacement.	10	Elkhodr et al. (2023) examined the performance and perceptions of 52 ICT students who used ChatGPT. Combining qualitative insights and quantitative rubric scores, their study indicated that most students found ChatGPT to be useful and enjoyable. It also improved their performance.	16
Total	44		28		72

References

* Adetayo, A. J. (2023). ChatGPT and librarians for reference consultations. *Internet Reference Services Quarterly, 27*(3), 131–147. https://doi.org/10.1080/10875301.2023.2203681.

* Ahmed, M. A. (2023). ChatGPT and the EFL classroom: Supplement or substitute in Saudi Arabia's Eastern region. *Information Sciences Letters, 12*(7), 2727–2734. https://doi.org/10.18576/isl/120704.

Ahmed, M., & Sab, C. (2023). Bibliometric analysis of ChatGPT's applications in medicine: A comprehensive assessment of its impact and potential. *International Journal of Medicine and Public Health, 13*(2), 74–83. https://www.ijmedph.org/index.php/ijmph/article/view/77.

* Ajlouni, A. O., Almahaireh, A. S., & Whaba, F. A. A. (2023a). Students' perception of using ChatGPT in counseling and mental health Education: The Benefits and Challenges. *International Journal of Emerging Technologies in Learning, 18*(20), 199–218. https://doi.org/10.3991/ijet.v18i20.42075.

* Ajlouni, A. O., Wahba, F. A. A., & Almahaireh, A. S. (2023b). Students' attitudes towards using ChatGPT as a learning tool: The case of the University of Jordan. *International Journal of Interactive Mobile Technologies, 17*(18), 99–117. https://doi.org/10.3991/ijim.v17i18.41753.

* Aktay, S., Gök, S., & Uzunoğlu, D. (2023). ChatGPT in education. *Türk Akademik Yayınlar Dergisi (TAY Journal), 7*(2), 378–406. https://doi.org/10.29329/tayjournal.2023.543.03.

* Al-Obaydi, L. H., Pikhart, M., & Klimova, B. (2023). ChatGPT and the general concepts of education: Can artificial intelligence-driven chatbots support the process of language learning?. *International Journal of Emerging Technologies in Learning, 18*(21), 39–50. https://doi.org/10.3991/ijet.v18i21.42593.

* Alneyadi, S., & Wardat, Y. (2023). ChatGPT: Revolutionizing student achievement in the electronic magnetism unit for eleventh-grade students in Emirates schools. *Contemporary Educational Technology, 15*(4), Article ep448. https://doi.org/10.30935/cedtech/13417.

* Alqasham, F. H. (2023). ChatGPT in the Saudi EFL classroom: A study of learner usage patterns and possibilities in learning optimization. *Migration Letters, 20*(7), 1251–1263 https://migrationletters.com/index.php/ml/article/view/4828.

* Alshurafat, H., Al Shbail, M. O., Hamdan, A., Al-Dmour, A., & Ensour, W. (2024). Factors affecting accounting students' misuse of ChatGPT: An application of the fraud triangle theory. *Journal of Financial Reporting & Accounting, 22*(2), 274–288. https://doi.org/10.1108/JFRA-04-2023-0182.

* Amaro, I., Barra, P., Della Greca, A., Francese, R., & Tucci, C. (2023). Believe in artificial intelligence? A user study on the ChatGPT's fake information impact. *IEEE Transactions on Computational Social Systems*. https://doi.org/10.1109/TCSS.2023.3291539. Advance online publication.

Aronson, B., & Laughter, J. (2016). The theory and practice of culturally relevant education: A synthesis of research across content areas. *Review of Educational Research, 86*(1), 163–206. https://doi.org/10.3102/0034654315582066

* Barrett, A., & Pack, A. (2023). Not quite eye to AI: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education, 20*, 59. https://doi.org/10.1186/s41239-023-00427-0.

* Bašić, Ž., Banovac, A., Kružić, I., & Jerković, I. (2023). ChatGPT-3.5 as writing assistance in students' essays. *Humanities and Social Sciences Communications, 10*, 750. https://doi.org/10.1057/s41599-023-02269-7.

* Bernabei, M., Colabianchi, S., Falegnami, A., & Costantino, F. (2023). Students' use of large language models in engineering education: A case study on technology acceptance, perceptions, efficacy, and detection chances. *Computers and Education: Artificial Intelligence, 5*, Article 100172. https://doi.org/10.1016/j.caeai.2023.100172.

* Bitzenbauer, P. (2023). ChatGPT in physics education: A pilot study on easy-to-implement activities. *Contemporary Educational Technology, 15*(3), ep430. https://doi.org/10.30935/cedtech/13176.

Bond, M. (2020). Facilitating student engagement through the flipped learning approach in K-12: A systematic review. *Computers & Education, 151*, Article 103819. https://doi.org/10.1016/j.compedu.2020.103819

Bond, M., & Bedenlier, S. (2019). Facilitating student engagement through educational technology: Towards a conceptual framework. *Journal of Interactive Media in Education, 2019*(1), 11. https://doi.org/10.5334/jime.528

Bond, M., Buntins, K., Bedenlier, S., Zawacki-Richter, O., & Kerres, M. (2020). Mapping research in student engagement and educational technology in higher education: A systematic evidence map. *International Journal of Educational Technology in Higher Education, 17*, 2. https://doi.org/10.1186/s41239-019-0176-8

* Boubker, O. (2024). From chatting to self-educating: Can AI tools boost student learning outcomes?. *Expert Systems with Applications, 238*, Article 121820. https://doi.org/10.1016/j.eswa.2023.121820.

* Cai, Q., Lin, Y., & Yu, Z. (2023). Factors influencing learner attitudes towards ChatGPT-Assisted language learning in higher education. *International Journal of Human-Computer Interaction*. https://doi.org/10.1080/10447318.2023.2261725. Advance online publication.

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

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

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

* Chan, M. M. K., Wong, I. S. F., Yau, S. Y., & Lam, V. S. F. (2023). Critical reflection on using ChatGPT in student learning: Benefits or potential risks?. *Nurse Educator, 48*(6), 200–201. https://doi.org/10.1097/NNE.0000000000001476.

* Chen, J., Lin, H., Han, X., & Sun, L. (2024). Benchmarking large language models in retrieval-augmented generation. *Proceedings of the AAAI Conference on Artificial Intelligence, 38*(16), 17754–17762. https://doi.org/10.1609/aaai.v38i16.29728.

Chen, B., Zhu, X., & del Castillo, H. F. D. (2023). Integrating generative AI in knowledge building. *Computers and Education: Artificial Intelligence, 5*, Article 100184. https://doi.org/10.1016/j.caeai.2023.100184

Christenson, S. L., Reschly, A. L., & Wylie, C. (2012). *Handbook of research on student engagement.* New York, NY: Springer. https://doi.org/10.1007/978-1-4614-2018-7

Clark, R. E. (1983). Reconsidering research on learning from media. *Review of Educational Research, 53*(4), 445–459. https://doi.org/10.3102/00346543053004445

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston: Pearson.

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

* Dai, Y., Lai, S., Lim, C. P., & Liu, A. (2023). ChatGPT and its impact on research supervision: Insights from Australian postgraduate research students. *Australasian Journal of Educational Technology, 39*(4), 74–88. https://doi.org/10.14742/ajet.8843.

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

* Duong, C. D., Bui, D. T., Pham, H. T., & Vu, A. T. (2023). How effort expectancy and performance expectancy interact to trigger higher education students' uses of ChatGPT for learning. *Interactive Technology and Smart Education.* https://doi.org/10.1108/ITSE-05-2023-0096. Advance online publication.

* Duong, C. D., Vu, T. N., & Ngo, T. V. N. (2023). Applying a modified technology acceptance model to explain higher education students' usage of ChatGPT: A serial multiple mediation model with knowledge sharing as a moderator. *International Journal of Management in Education, 21*(3), Article 100883. https://doi.org/10.1016/j.ijme.2023.100883.

* 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.

* Escalante, J., Pack, A., & Barrett, A. (2023). 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.

Fan, X., & Chen, M. (2001). Parental involvement and students' academic achievement: A meta-analysis. *Educational Psychology Review, 13*(1), 1–22. https://doi.org/10.1023/A:1009048817385

* Firat, M. (2023). What ChatGPT means for universities: Perceptions of scholars and students. *Journal of Applied Learning and Teaching, 6*(1), 57–63. https://doi.org/10.37074/jalt.2023.6.1.22.

Flanagin, A., Bibbins-Domingo, K., Berkwits, M., & Christiansen, S. L. (2023). Nonhuman "authors" and implications for the integrity of scientific publication and medical knowledge. *JAMA, 329*(8), 637–639. https://doi.org/10.1001/jama.2023.1344

Foltynek, T., Bjelobaba, S., Glendinning, I., Khan, Z. R., Santos, R., Pavletic, P., & Kravjar, J. (2023). ENAI recommendations on the ethical use of artificial intelligence in education. *International Journal for Educational Integrity, 19*, 12. https://doi.org/10.1007/s40979-023-00133-4

* Foroughi, B., Senali, M. G., Iranmanesh, M., Khanfar, A., Ghobakhloo, M., Annamalai, N., & Naghmeh-Abbaspour, B. (2023). Determinants of intention to use ChatGPT for educational purposes: Findings from PLS-SEM and fsQCA. *International Journal of Human-Computer Interaction.* https://doi.org/10.1080/10447318.2023.2226495. Advance online publication.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Garg, R. K., Urs, V. L., Agarwal, A. A., Chaudhary, S. K., Paliwal, V., & Kar, S. K. (2023). Exploring the role of ChatGPT in patient care (diagnosis and treatment) and medical research: A systematic review. *Health Promotion Perspectives, 13*(3), 183–191. https://doi.org/10.34172/hpp.2023.22

Gödde, D., Nöhl, S., Wolf, C., Rupert, Y., Rimkus, L., Ehlers, J., Breuckmann, F., & Sellmann, T. (2023). A SWOT (strengths, weaknesses, opportunities, and threats) analysis of ChatGPT in the medical literature: Concise review. *Journal of Medical Internet Research, 25*, Article e49368. https://doi.org/10.2196/49368

* Gorichanaz, T. (2023). Accused: How students respond to allegations of using ChatGPT on assessments. *Learning: Research and Practice, 9*(2), 183–196. https://doi.org/10.1080/23735082.2023.2254787.

Hair, J. F., Hult, G. T. M., & Ringle, C. M. (2014). *Partial least squares structural equation modeling (PLS-SEM).* Thousand Oaks, CA: SAGE Publications.

* 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.

* Hasanein, A. M., & Sobaih, A. E. E. (2023). Drivers and consequences of ChatGPT use in higher education: Key stakeholder perspectives. *European Journal of Investigation in Health, Psychology and Education, 13*(11), 2599–2614. https://doi.org/10.3390/ejihpe13110181.

* Ho, W., & Lee, D. (2023). Enhancing engineering education in the roblox metaverse: Utilizing ChatGPT for game development for electrical machine course. *International Journal of Advanced Science, Engineering and Information Technology, 13*(3), 1052–1058 https://insightsociety.org/ojaseit/index.php/ijaseit/article/view/18458.

* Hosseini, M., Gao, C. A., Liebovitz, D. M., Carvalho, A. M., Ahmad, F. S., Luo, Y., … Kho, A. (2023). An exploratory survey about using ChatGPT in education, healthcare, and research. *PLoS One, 18*(10), Article 0292216. https://doi.org/10.1101/2023.03.31.23287979.

* Hsu, M. H. (2024). Mastering medical terminology with ChatGPT and termbot. *Health Education Journal, 83*(4), 352–358. https://doi.org/10.1177/00178969231197371.

Hu, K. (2023). *ChatGPT sets record for fastest-growing user base – Analyst note.* Retrieved from https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/.

* Hu, J. M., Liu, F. C., Chu, C. M., & Chang, Y. T. (2023). Health care trainees' and professionals' perceptions of ChatGPT in improving medical knowledge training: Rapid survey study. *Journal of Medical Internet Research, 25*, Article e49385. https://doi.org/10.2196/49385.

Huang, B., Hew, K. F., & Lo, C. K. (2019). Investigating the effects of gamification-enhanced flipped learning on undergraduate students' behavioral and cognitive engagement. *Interactive Learning Environments, 27*(8), 1106–1126. https://doi.org/10.1080/10494820.2018.1495653

Hwang, G. J., & Chang, C. Y. (2023). A review of opportunities and challenges of chatbots in education. *Interactive Learning Environments, 31*(7), 4099–4112. https://doi.org/10.1080/10494820.2021.1952615

* Ibrahim, H., Liu, F., Asim, R., Battu, B., Benabderrahmane, S., Alhafni, B., … Zaki, Y. (2023). Perception, performance, and detectability of conversational artificial intelligence across 32 university courses. *Scientific Reports, 13*(1), Article 12187. https://doi.org/10.1038/s41598-023-38964-3.

Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology, 15*(4), ep464. https://doi.org/10.30935/cedtech/13605

İpek, Z. H., Gözüm, A. I. C., Papadakis, S., & Kallogiannakis, M. (2023). Educational applications of the ChatGPT AI system: A systematic review research. *Educational Process: International Journal, 12*(3), 26–55. https://doi.org/10.22521/edupij.2023.123.2

* Jauhiainen, J. S., & Guerra, A. G. (2023). Generative AI and ChatGPT in school children's education: Evidence from a school lesson. *Sustainability, 15*(18), Article 14025. https://doi.org/10.3390/su151814025.

* Javier, D. R. C., & Moorhouse, B. L. (2023). Developing secondary school English language learners' productive and critical use of ChatGPT. *TESOL Journal.* https://doi.org/10.1002/tesj.755. Advance online publication.

* Jo, H. (2023a). Decoding the ChatGPT mystery: A comprehensive exploration of factors driving AI language model adoption. *Information Development.* https://doi.org/10.1177/02666669231202764. Advance online publication.

* Jo, H. (2023b). Understanding AI tool engagement: A study of ChatGPT usage and word-of-mouth among university students and office workers. *Telematics and Informatics, 85*, Article 102067. https://doi.org/10.1016/j.tele.2023.102067.

* Kevin, F., & Veronica, A. (2023). Integrating artificial intelligence in higher education: Empirical insights from students about using ChatGPT. *International Journal of Information and Education Technology, 13*(9), 1365–1371. https://doi.org/10.18178/ijiet.2023.13.9.1939.

* Kirpalani, A., Grimmer, J., & Wang, P. Z. T. (2023). Med versus machine: Using ChatGPT in team-based learning. *Medical Education, 57*(11), 1159–1160 https://asmepublications.onlinelibrary.wiley.com/doi/10.1111/medu.15226.

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research, 86*(3), 643–680. https://doi.org/10.3102/0034654315626799

Levin, G., Horesh, N., Brezinov, Y., & Meyer, R. (2024). Performance of ChatGPT in medical examinations: A systematic review and a meta-analysis. *BJOG: An International Journal of Obstetrics and Gynaecology, 131*(3), 378–380. https://doi.org/10.1111/1471-0528.17641

* Limna, P., Kraiwanit, T., Jangjarat, K., Klayklung, P., & Chocksathaporn, P. (2023). The use of ChatGPT in the digital era: Perspectives on chatbot implementation. *Journal of Applied Learning and Teaching, 6*(1), 64–74. https://doi.org/10.37074/jalt.2023.6.1.32.

* Limo, F. A. F., Tiza, D. R. H., Roque, M. M., Herrera, E. E., Murillo, J. P. M., Huallpa, J. J., … Gonzáles, J. L. A. (2023). Personalized tutoring: ChatGPT as a virtual tutor for personalized learning experiences. *Social Space, 23*(1), 293–312 https://socialspacejournal.eu/menu-script/index.php/ssj/article/view/176/81.

* Liu, G., & Ma, C. (2024). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching, 18*(2), 125–138. https://doi.org/10.1080/17501229.2023.2240316.

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13*(4), 410. https://doi.org/10.3390/educsci13040410

Lo, C. K., & Hew, K. F. (2021). Student engagement in mathematics flipped classrooms: Implications of journal publications from 2011 to 2020. *Frontiers in Psychology, 12*, Article 672610. https://doi.org/10.3389/fpsyg.2021.672610

* Lozano, A., & Blanco Fontao, C. (2023). Is the education system prepared for the irruption of artificial intelligence? A study on the perceptions of students of primary education degree from a dual perspective: Current pupils and future teachers. *Education Sciences, 13*(7), 733. https://doi.org/10.3390/educsci13070733.

* Menon, D., & Shilpa, K. (2023). "Chatting with ChatGPT": Analyzing the factors influencing users' intention to use the open AI's ChatGPT using the UTAUT model. *Heliyon, 9*(11), Article e20962. https://doi.org/10.1016/j.heliyon.2023.e20962.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & the PRISMA Group. (2009). Reprint-preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Physical Therapy, 89*(9), 873–880. https://doi.org/10.1093/ptj/89.9.873

* Muñoz, S., Gayoso, G., Huambo, A., Tapia, R., Incaluque, J. L., Pon-go, O. E., … Arias, J. L. (2023). Examining the impacts of ChatGPT on student motivation and engagement. *Social Space, 23*(1), 1–27 https://www.researchgate.net/publication/370894242.

* Ngo, T. T. A. (2023). The perception by university students of the use of ChatGPT in education. *International Journal of Emerging Technologies in Learning (Online), 18* (17), 4–19. https://doi.org/10.3991/ijet.v18i17.39019.

OpenAI. (2022). *Introducing ChatGPT*. Retrieved from https://openai.com/blog/chatgpt.

* Rad, D., & Rad, G. (2023). Exploring the psychological implications of ChatGPT: A qualitative study. *Journal of Education, 32*(1), 43–55. https://doi.org/10.24250/jpe/Vol.321/2023/DR/GR.

* Romero-Rodríguez, J. M., Ramírez-Montoya, M. S., Buenestado-Fernández, M., & Lara-Lara, F. (2023). Use of ChatGPT at university as a tool for complex thinking: Students' perceived usefulness. *Journal of New Approaches in Educational Research, 12*(2), 323–339. https://doi.org/10.7821/naer.2023.7.1458.

Roumeliotis, K. I., & Tselikas, N. D. (2023). ChatGPT and open-AI models: A preliminary review. *Future Internet, 15*(6), 192. https://doi.org/10.3390/fi15060192

* Sahari, Y., Al-Kadi, A. M. T., & Ali, J. K. M. (2023). A cross sectional study of ChatGPT in translation: Magnitude of use, attitudes, and uncertainties. *Journal of Psycholinguistic Research, 52*, 2937–2954. https://doi.org/10.1007/s10936-023-10031-y.

Sallam, M. (2023). ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns. *Healthcare, 11*(6), 887. https://doi.org/10.3390/healthcare11060887

* Sallam, M., Salim, N. A., Barakat, M., Al-Mahzoum, K., Ai-Tammemi, A. B., Malaeb, D., Hallit, R., & Hallit, S. (2023). Assessing health students' attitudes and usage of ChatGPT in Jordan: Validation study. *JMIR Medical Education, 9*(1), Article e48254. https://doi.org/10.2196/48254.

* Saxena, A., & Doleck, T. (2023). A structural model of student continuance intentions in ChatGPT adoption. *Eurasia Journal of Mathematics, Science and Technology Education, 19*(12), Article e2366. https://doi.org/10.29333/ejmste/13839.

* Shaikh, S., Yayilgan, S. Y., Klimova, B., & Pikhart, M. (2023). Assessing the usability of ChatGPT for formal English language learning. *European Journal of Investigation in Health, Psychology and Education, 13*(9), 1937–1960. https://doi.org/10.3390/ejihpe13090140.

Shen, Y., Heacock, L., Elias, J., Hentel, K. D., Reig, B., Shih, G., & Moy, L. (2023). ChatGPT and other large language models are double-edged swords. *Radiology, 307* (2), Article e230163. https://doi.org/10.1148/radiol.230163

* 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.

Sinatra, G. M., Heddy, B. C., & Lombardi, D. (2015). The challenges of defining and measuring student engagement in science. *Educational Psychologist, 50*(1), 1–13. https://doi.org/10.1080/00461520.2014.1002924

* 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), 924. https://doi.org/10.3390/educsci13090924.

Stemler, S. E. (2004). A Comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research and Evaluation, 9*, 4. https://doi.org/10.7275/96jp-xz07

* Strzelecki, A. (2023). To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2023.2209881. Advance online publication.

Thiese, M. S. (2014). Observational and interventional study design types: An overview. *Biochemia Medica, 24*(2), 199–210. https://doi.org/10.11613/BM.2014.022

* Tiwari, C. K., Bhat, M. A., Khan, S. T., Subramaniam, R., & Khan, M. A. I. (2023). What drives students toward ChatGPT? An investigation of the factors influencing adoption and usage of ChatGPT. *Interactive Technology and Smart Education*. https://doi.org/10.1108/ITSE-04-2023-0061. Advance online publication.

* Tlili, A., Shehata, B., & Adarkwah, M. A. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*, 15. https://doi.org/10.1186/s40561-023-00237-x.

* Uddin, S. J., Albert, A., Ovid, A., & Alsharef, A. (2023). Leveraging ChatGPT to aid construction hazard recognition and support safety education and training. *Sustainability, 15*(9), 7121. https://doi.org/10.3390/su15097121.

Vargas-Murillo, A. R., de la Asuncion, I. N. M., & de Jesús Guevara-Soto, F. (2023). Challenges and opportunities of AI-assisted learning: A systematic literature review on the impact of ChatGPT usage in higher education. *International Journal of Learning, Teaching and Educational Research, 22*(7), 122–135. https://doi.org/10.26803/ijlter.22.7.7

* Vecchiarini, M., & Somià, T. (2023). Redefining entrepreneurship education in the age of artificial intelligence: An explorative analysis. *International Journal of Management in Education, 21*(3), Article 100879. https://doi.org/10.1016/j.ijme.2023.100879.

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly, 36*(1), 157–178. https://doi.org/10.2307/41410412

* von Garrel, J., & Mayer, J. (2023). Artificial Intelligence in studies—use of ChatGPT and AI-based tools among students in Germany. *Humanities and Social Sciences Communications, 10*, 799. https://doi.org/10.1057/s41599-023-02304-7.

* Waltzer, T., Cox, R. L., & Heyman, G. D. (2023). Testing the ability of teachers and students to differentiate between essays generated by ChatGPT and high school students. *Human Behavior and Emerging Technologies, 2023*, Article 1923981. https://doi.org/10.1155/2023/1923981.

* Wandelt, S., Sun, X., & Zhang, A. (2023). AI-driven assistants for education and research? A case study on ChatGPT for air transport management. *Journal of Air Transport Management, 113*, Article 102483. https://doi.org/10.1016/j.jairtraman.2023.102483.

* Wang, M., Wang, M., Xu, X., Yang, L., Cai, D., & Yin, M. (2023). Unleashing ChatGPT's power: A case study on optimizing information retrieval in flipped classrooms via prompt engineering. *IEEE Transactions on Learning Technologies, 17*, 629–641. https://doi.org/10.1109/TLT.2023.3324714.

* Wardat, Y., Tashtoush, M. A., AlAli, R., & Jarrah, A. M. (2023). ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia Journal of Mathematics, Science and Technology Education, 19*(7), Article em2286. https://doi.org/10.29333/ejmste/13272.

Wu, T., He, S., Liu, J., Sun, S., Liu, K., Han, Q. L., & Tang, Y. (2023). A brief overview of ChatGPT: The history, status quo and potential future development. *IEEE/CAA Journal of Automatica Sinica, 10*(5), 1122–1136. https://doi.org/10.1109/JAS.2023.123618

* Wu, T. T., Lee, H. Y., Li, P. H., Huang, C. N., & Huang, Y. M. (2024). Promoting self-regulation progress and knowledge construction in blended learning via ChatGPT-based learning aid. *Journal of Educational Computing Research, 61*(8), 3–31. https://doi.org/10.1177/07356633231191125.

* Xiao, Y., & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages, 8*(3), 212. https://doi.org/10.3390/languages8030212.

* Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies, 28*, 13943–13967. https://doi.org/10.1007/s10639-023-11742-4.

* Yilmaz, R., & Yilmaz, F. G. K. (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*, Article 100147. https://doi.org/10.1016/j.caeai.2023.100147.

Zamfiroiu, A., Vasile, D., & Savu, D. (2023). ChatGPT – a systematic review of published research papers. *Informatica Economica, 27*(1), 5–16. https://doi.org/10.24818/issn14531305/27.1.2023.01

Zhang, P., & Tur, G. (2023). A systematic review of ChatGPT use in K-12 education. *European Journal of Education.* https://doi.org/10.1111/ejed.12599. Advance online publication.

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

* Zou, M., & Huang, L. (2023). To use or not to use? Understanding doctoral students' acceptance of ChatGPT in writing through technology acceptance model. *Frontiers in Psychology, 14*, Article 1259531. https://doi.org/10.3389/fpsyg.2023.1259531.