Scilight

Journal of Educational Technology and Innovation
https://jeti.thewsu.org/index.php/cieti

Review

Facilitator or Barrier? A Systematic Review on the Relationship between Artificial Intelligence Technologies and the Development of Critical Thinking Skills

Haoming Lin [1,*], Wei Wei [1] and Handan Lu [2]

[1] Faculty of Applied Sciences, Macao Polytechnic University, Macao, China
[2] The Assessment and Qualifications Alliance (AQA), Surrey GU2 7XJ, UK
* Correspondence: haoming.lin@outlook.com

How To Cite: Lin, H., Wei, W., & Lu, H. (2025). Facilitator or Barrier? A Systematic Review on the Relationship between Artificial Intelligence Technologies and the Development of Critical Thinking Skills. *Journal of Educational Technology and Innovation*, *7*(2), 11–24. https://doi.org/10.61414/f2ah8r08

Received: 24 February 2025
Revised: 5 May 2025
Accepted: 25 May 2025
Published: 27 June 2025

Abstract: The advancement of Artificial Intelligence (AI) has garnered significant attention within the academic research community, reflecting the prevailing zeitgeist. However, there is a paucity of literature that has delved into its connection with the higher order thinking skills of human beings. The purpose of this systematic review is to investigate the relationship between AI utilization and the development of critical thinking (CT) in the field of education. A systematic literature search was performed in two databases, Web of Science and Scopus, with a focus on empirical studies related to AI and CT. The review process followed the PRISMA framework and adopted a bottom-up approach, Ultimately, the integrated review synthesized 21 eligible studies from a total of 649 articles. The systematic review identified three relationships between AI technologies and CT. The results suggest that AI technologies can help to enhance learners' CT skills (n = 13). However, excessive or inappropriate utilization of AI may hinder CT development (n = 7). Moreover, the cultivation of CT skills has been shown to positively influence the effectiveness of AI utilization (n = 4). This article represents the first literature review to delve into the reciprocal relationship between AI implementation and CT development within the education field, striving to illuminate the ways in which learners can enhance their higher-order thinking skills through more effective utilization of AI technologies.

Keywords: artificial intelligence; critical thinking; systematic review; higher-order thinking; education

1. Introduction

In recent times, with the popularization of computing technologies, numerous multifunctional Artificial Intelligence (AI) applications tailored to real-world scenarios for solving real-life problems have been developed and applied in various fields. Within educational contexts, the increased accessibility of diversified commercial AI tools (e.g., ChatGPT and Bard) has inspired students to learn beyond the constraints of time and space (Ji et al., 2023) and has stretched teachers' boundaries to explore new possibilities (Nazaretsky et al., 2022). The accessibility has also facilitated the studies as regards the feasibility of incorporating AI technologies into daily class activities. A cursory glance at the relevant literature on AI technologies reveals that scholars are addressing research issues regarding their impact on learners' behavior (Wang et al., 2022), affect (Lee et al., 2022), and cognition (Jin et al., 2023), including motivation, self-efficacy, learning performance, and communication skills. However, as a subset of cognition, higher-order thinking skills have received comparatively limited empirical

attention, with a recent review by Chu et al. (2022) indicating that only a paucity of studies have placed these skills at the center of their investigation into the effects of AI technologies.

Critical thinking (CT), as an essential component of higher order thinking skills (Ennis, 1987, Halpern, 1998, Schraw et al., 2011), is a vital attribute in the 21st century. It encourages students to trouble-shoot problems with out-of-the-box solutions (Anggraeni et al., 2023) and inspires teachers to gain more fine-grained insights into teaching abilities as well as academic skills (Yuan et al., 2022). Although teachers and students across various disciplines profit significantly from adopting CT in practical, numerous studies (Ahern et al., 2019, Yin et al., 2023, Huang and Sang, 2023) have found it difficult to define CT theoretically due to its multidimensionality and contextualization. For instance, CT can be generally perceived as clinical reasoning (Hong and Yu, 2017) and clinical decision-making ability (Ludin, 2018, Lee et al., 2017) in the context of healthcare and nursing. However, in the realm of English language learning, El Soufi and See (2019) would rather link the definition of CT with Toulmin's argumentation model and state that CT is about generating proper arguments. As a pioneer attempt to address this dissent, four camps of experts, including scientists from the social and physical science field, educationalist and philosophers, also weighed in on the Delphi Project (Facione, 1990), which has raised a clarion call to reconcile different definitions of CT from an interdisciplinary perspective. In their framework (Facione, 1990), CT encompasses a metacognitive character (i.e., self-regulation), interpretation, analysis, assessment, inference and elucidation. Subsequently, despite the growing body of literature (Dwyer et al., 2014, Suligoj et al., 2020) have mushroomed in number to explore the multifaceted nature of CT, the emergence of new technologies, such as AR, VR and AI, calls for the further infusion of new concepts such as information literacy (Wang and Wu, 2023) and digital literacy (Lim, 2023) into the CT framework, and the still-fuzzy conceptions of CT reflect the need for a systematic review aligned with the demands of digital transformation.

Three studies have been identified to provide an in-depth overview of the adoption of technology in the cultivation of CT. For example, two previous studies (Lu and Xie, 2022, Liang, 2023), which emphasize English language education, have suggested the effectiveness of technology-enhanced tools in assisting the improvement of CT skills as well as English proficiency. Another research conducted by Chou et al. (2019) also contributes to the literature on CT by analyzing its characteristics and trends in the e-learning environment. However, while it primarily focuses on technology-assisted learning activities, the specific details of technologies have not been carefully delineated in their paper. Moreover, none of these studies have illustrated the negative effect of technological tools on CT, suggesting a high probability of publication bias (Marks-Anglin and Chen, 2020)—meaning that positive results are more likely to be favored by publishers and thus more likely to be published. Similarly, this situation is in line with the observation made by El Soufi and See (2019), where scholars tend to report positive results and are therefore more passionate about detecting positive result.

While there are burgeoning publications focused on using technologies to support the training of CT skills, inasmuch as we know, there has been little research synthesis (Melisa et al., 2025) that discusses the relationships between the implementation of AI technologies and the development of CT skills. This gap in literature may prevent academia from understanding the full gamut of implications brought by the AI, given that these technologies have matured and penetrated into our daily learning and thinking activities. Given this context, CT skills were selected as the focus of this review due to their foundational role in decision-making, and problem-solving across disciplines. Moreover, unlike other cognitive or technical skills, CT is uniquely impacted by AI technologies, which have a twofold impact: they can both enhance analytical reasoning and potentially undermine independent judgment through automation and overreliance, an effect that makes the relationship between AI and CT particularly worthy of in-depth investigation. Additionally, Hao et al. (2024) emphasized that CT plays a pivotal role in fostering the responsible and effective use of AI, yet comprehensive research synthesis on this topic remains insufficient. Furthermore, the absence of an integrated definition of CT in previous studies of AI-assisted teaching and learning poses a risk to generalizing findings to other contexts. The current study serves as the primary point of contact in AI-assisted CT cultivation, aiming to address these research vacancies. To this end, we propose the following three questions:

RQ1. Can the implementation of AI technologies enhance learners' development of CT skills? If so, what might be the possible reasons?

RQ2. Can the implementation of AI technologies hinder students' development of CT skills? If so, what might be the possible reasons?

RQ3. Can the development of CT skills improve the effectiveness of using AI? If so, what might be the possible reasons?

2. Research Methodology

To conduct a systematic review about AI technologies and CT, we rigorously followed the guidance provided by the updated PRISMA framework (Page et al., 2021), which consisted of three main sections (see Figure 1): (1) literature identification, namely detecting the review sources of articles by searching key strings, (2) literature screening, namely filtering out the ineligible records by adopting a criteria-based screening strategy, (3) data analysis, namely scrutinizing and analyzing the remaining literature with a fine-tooth comb.

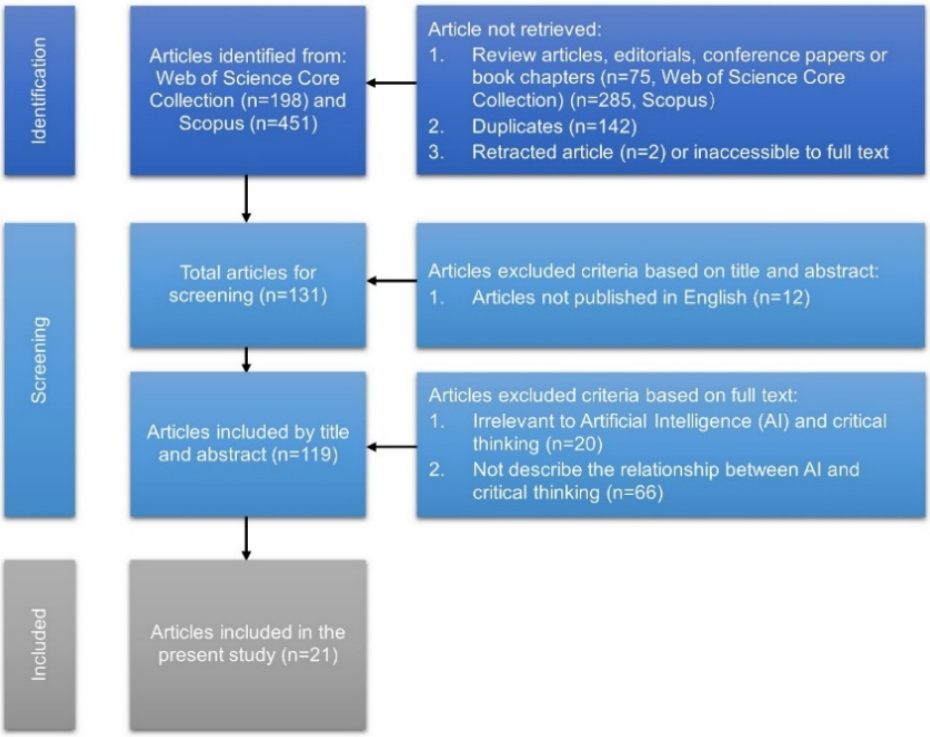

Figure 1. Flowchart of literature selection procedure.

2.1. Literature Identification

The search keywords in this study were "AI" or its full name "artificial intelligence" AND "critical thinking," and the operations were performed multiple times to minimize the risk of missing potential research. Two electronic literature search engines, Web of Science Core Collection (WOSCC) and Scopus, served as the main data sources because they could provide highly qualified records of peer-reviewed papers. The reliability of these sources was also validated by some review studies (Zhang et al., 2023, Yu and Zin, 2023).

The literature search was conducted in July 2023 without limiting the time range of the publications, allowing for a holistic view of the review process. This search yielded a total of 649 articles, including 198 articles from WOSCC and 451 articles from Scopus.

2.2. Literature Screening

The criteria for paper screening have been thoroughly executed and are presented in Table 1. The literature screening process was carried out following the approaches of Yang and Wei (2023) and Zhao (2024), and is illustrated as follows.

Initially, 649 articles served as the baseline for further selection. The first set of 360 articles (including 75 from WOSCC and 285 from Scopus) were excluded with the assistance of search engines (WOSCC and Scopus) using the filtering panel due to the violation of inclusion criteria 1 (only research articles are eligible). Afterward, 142 duplicates, 2 retracted records and 14 inaccessible articles were screened out prudently by the authors of the current study, leaving 131 articles for screening.

In the second screening phase, 12 entries of non-English papers were eliminated at once based on their titles and abstract, leaving 119 articles for intensive reading. Of these, 20 studies were judged as inappropriate for further inspection either because they did not involve AI applications or were unrelated to CT skills. A relatively large number of articles (n = 66) were removed because they lacked descriptions of the interrelationship between AI

and CT. The removal of 12 articles was also conducted because they did not provide empirical data, resulting in a total of 21 articles included for the final review.

Table 1. Inclusion and exclusion criteria.

No.	Inclusion Criteria	Exclusion Criteria
1	Peer-reviewed research articles	Review articles, editorials, conference papers and book chapters
2	The studies must involve some applications of AI.	Not related to AI applications
3	The studies must address individuals' CT.	Not related to CT
4	The studies must discuss the relationship between AI and CT.	Not discuss the relationship between AI and CT.
5	The studies must be published in English	Published other than English
6	The studies must be accessible to full text.	Not accessible to full text
7	The studies must be unretracted	Retracted studies
8	The studies must be empirical	Not provide empirical data

2.3. Data Analysis

We employed a two-step coding approach as advanced and applied by Chou et al. (2019). The included articles were first read for collecting and analyzing the basic information of articles (i.e., year of publication, context of the study, article sources, educational level, and field of discipline). Second, we read for evidence related to research questions (i.e., identifying AI names, affordances of AI, defining CT, assessing CT measurements and evidence, and exploring the relationship between AI and CT until it attained a state of saturation, at which point further data analysis appeared to generate no extra insights into the reviewed literature.

To ensure review validity, two researchers coded the literature separately using a bottom-up approach (Zou et al., 2022). This approach involved making extrapolations and classifications from the extracted information in the included articles. The researchers then met face-to-face to scrutinize the results of the analysis and resolve any divergences. In cases where consensus could not be reached between the two researchers, discussions were held via online meetings with a third researcher to achieve agreement.

3. Findings

3.1. Overview of the Reviewed Studies

With reference to the year of publication and the number of publications, there was a limited amount of research focusing on AI and CT before 2022 (year of 2020: n = 1; year of 2021, n = 0; year of 2022, n = 4). However, this Figure 2 saw an exponential increase to 16 in 2023, with 10 articles involving the use of ChatGPT.

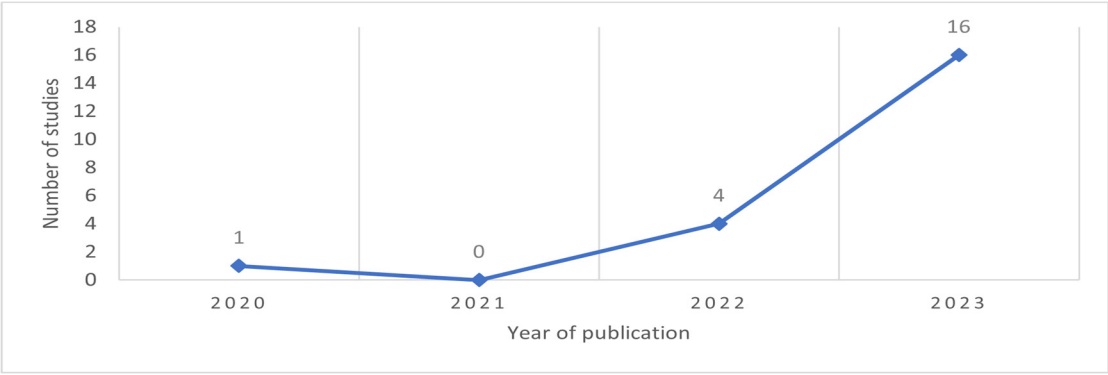

Figure 2. Year of publication.

These studies were conducted in a wide spread of contexts (see Figure 3), and nearly half of the studies were conducted to investigate the status quo of AI and CT in the USA (n = 4), Australia (n = 3), and Saudi Arabia (n = 2), while the rest of the studies were scattered from Asia (n = 6), Europe (n = 4), South America (n = 1), to Africa (n = 1).

A plethora of studies discussing the relationship between AI and CT were conducted in the USA and Australia, and Bearman et al. (2023) similarly noted a prevalence of AI-related studies in these two regions. This can be attributed to the fact that numerous revolutionary AI technologies are developed in the USA, while Australia also possesses suitable economic and technological conditions to implement new technologies.

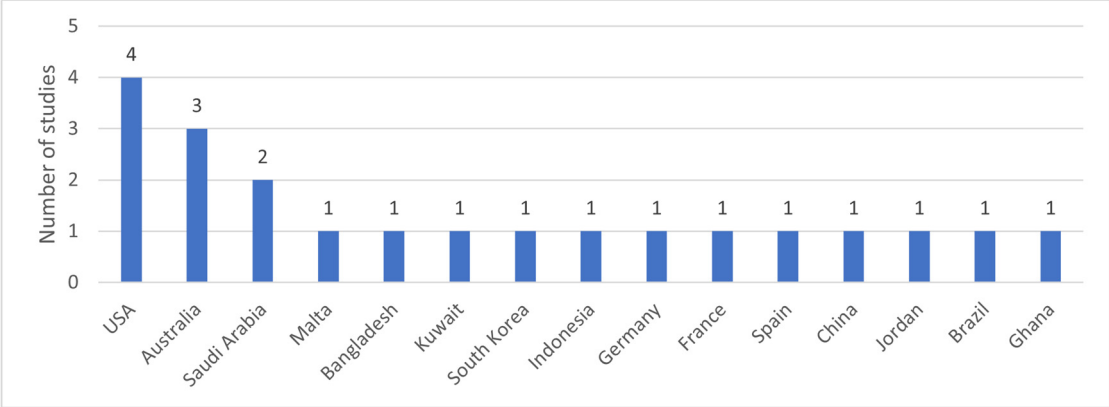

Figure 3. Context of studies.

Regarding the journal and article sources (Figure 4), numerous studies on AI and CT was published in *Education and Information Technologies* (n = 3), which focused on exploring a wide range of educational contexts associated with new technological advancements. Journals such as *Applied Sciences* (n = 2), *Frontiers in Education* (n = 2), and *International Journal of Medical Informatics* (n = 2) also featured several studies on AI and CT.

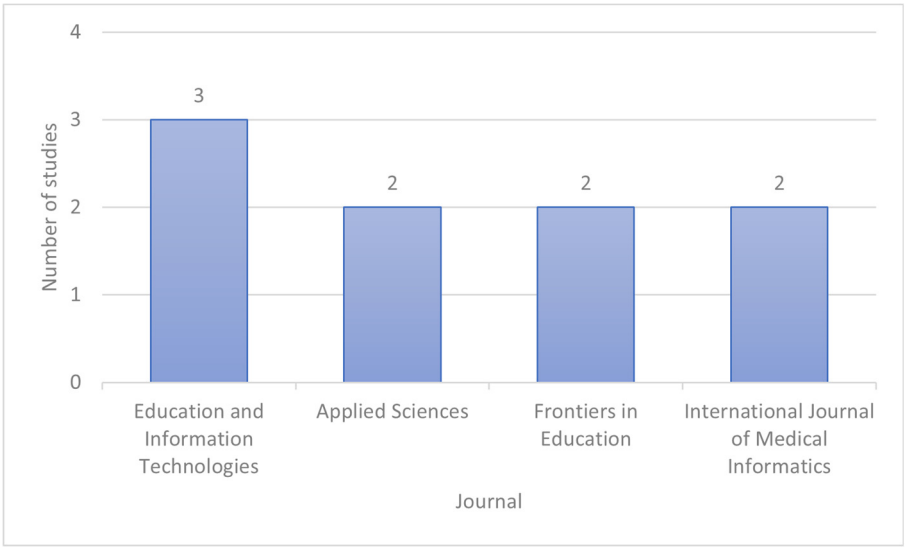

Figure 4. Top 4 article resources.

In terms of the discipline areas (see Figure 5), more than half of the studies focused on the discipline of STEAM education (n = 11), representing 52% of the total. Medical education (n = 5) and Language education (n = 3) followed closely, accounting for 24% and 14%, respectively. Tourism education (n = 1) and Journalism and media education (n = 1) each represented 5%.

After reviewing the literature, it is evident that a significant number of studies have primarily focused on assessing CT among university students rather than students at other levels, consistent with findings from several studies (Lu and Xie, 2022, Chou et al., 2019). Higher education institutions emerge as ideal settings for conducting such research due to the students' advanced cognitive capabilities and the ease of recruiting volunteers. Nonetheless, there remains an insufficient emphasis on nurturing CT skills among young learners, despite the imperative need for consistent and sustained development (Liang, 2023, Drennan, 2010).

In terms of data sources, self-reported data serve as the primary resources for numerous studies, a trend consistent with the findings of two studies (Liang, 2023, Yuan et al., 2022) that focused on investigating CT within the field of English education. Subsequently, content analysis has also emerged as a frequent method employed to assess CT-related responses generated by AI, and the result is corroborated by another study (Puig et al., 2020). Standardized CT tests, such as the California Critical Thinking Skills Test and Ennis-Weir Critical Thinking Essay Tests, have rarely been used as instruments to measure CT.

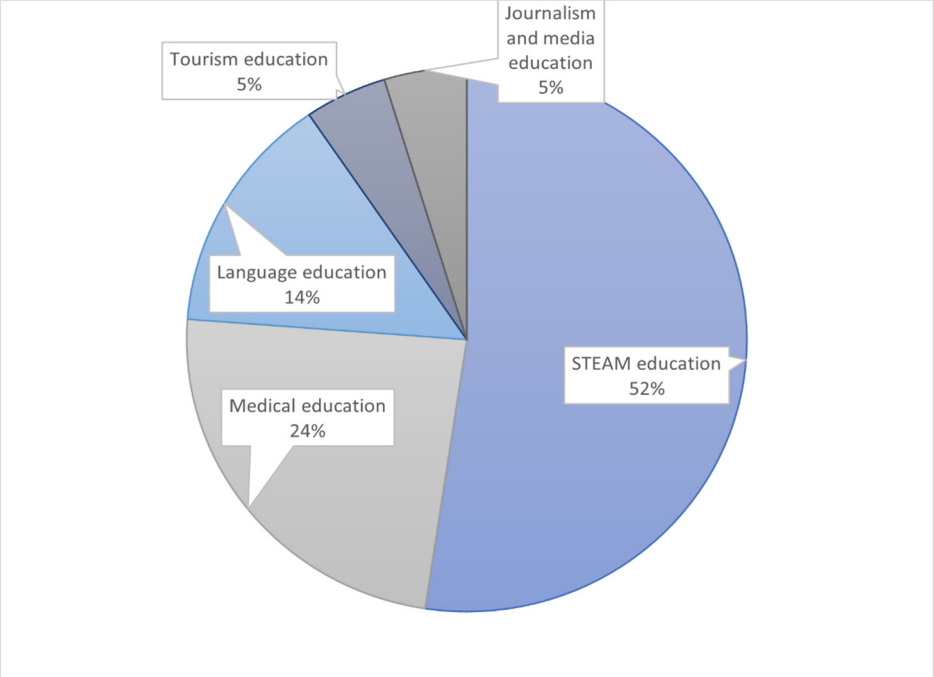

Figure 5. Discipline areas.

3.2. RQ1: Can the Implementation of AI Technologies Enhance Learners' Development of CT Skills? If So, What Might Be the Possible Reasons?

Our research found that AI technology implementation could have substantial potential for enhancing CT skill development. We can speculate two explanatory mechanisms why AI technologies can help improve the development of CT skills. First, most of the affordances of AI technologies contain elements that facilitate CT skills. For example, textual chatbot can inspire students' reflective thinking and enhance problem-solving skills by providing multidimensional viewpoints and applicable suggestions. Additionally, learning assistance tools can offer additional learning opportunities that foster creativity. Second, some AI technologies, such as clinical decision-making and automation systems, can simplify the process of performing repetitive tasks, allowing students to put their primary focus on more complex thinking activities.

13 (Table A1), out of the 21 examined studies, reported enhancements in CT skills. Of these studies, the frequently investigated AI technologies comprise: (1) Textual chatbots that generate responses and perform evaluations (e.g., ChatGPT and Replika, n = 5); (2) Clinical decision-making and automation systems (e.g., Healthcare AI, AiDental, and AI-Clinical Decision Support Systems, n = 3); (3) Learning assistance tools (e.g., Digital STEAM Creativity Tools, Riff platform, the Meeting Mediator tool, and AI algorithms for education, n = 3); (4) Game-based AI (e.g., Artbot, n = 1); and (5) Dynamic assessment systems (e.g., Automatic writing evaluation system, AWE, n = 1).

The affordances encompassed by the different types of AI technologies and the development of CT are also worth attention. Textual chatbots typically offer affordances such as providing feedback and answers, conducting evaluation, facilitating interaction, provoking discussions, and promoting collaboration. For instance, Rahman and Watanobe (2023) revealed that ChatGPT serves as a generator of huma-like responses and proficient in answering programming questions. Clinical decision-making and automation systems mainly function in automating healthcare work and clinical tasks, as well as conducting evaluations. For example, AI systems could help diagnose illness based on images and produce clinical reports automatically in their work (Aquino et al., 2023). Learning assistance tools emphasized their roles in promoting learning opportunities, recommending materials, provoking reflection and discussions, providing feedback, and conducting evaluations. Take the research employed by Henze et al. (2022) as an example: different tools have been used to assist students in programming, and one tool named DJI Tello Edu drone even allowed them to simulate the flight of a drone by testing the programming code within the tool. Finally, game-based AI and dynamic assessment systems both focus on provoking reflection and discussions, while dynamic assessment systems also hold additional capabilities in assigning tasks, providing feedback, and conducting evaluations.

Of these 13 studies, 5 defined CT skills to include reflection and reflective thinking, which relates to pondering how to cope with a task or contemplating what has been done after finishing a task (Ghanizadeh, 2017).

Meanwhile, 4 studies within the medical field integrated clinical decision-making ability into the CT framework. Clinical decision-making ability is defined as the consideration of different ideas and choices (Lee et al., 2017). Moreover, some studies also emphasized problem-solving skill and analytical ability, which refer to the ability to solve interdisciplinary problems that occur in real-life context (Dwyer et al., 2014) and the ability to clarify relationships between diversified information (Facione, 1990), respectively, as important components of CT. Other CT skills that AI has enhanced may include critical writing skill, scientific reasoning, and creative thinking.

The improvement of CT skills is predominantly evidenced by students' self-reported improvement in CT skills through surveys, interviews, comments, and learning analytics. A case in point is Liu et al. (2023), whose work adopted a 6-item questionnaire to evaluate students' CT tendencies, and the results suggested that peer-assessment assisted AI technology (AWE in particular) enables students to obtain higher scores in CT. This finding is also supported by interview data. Meanwhile, AI technologies' test performance (Nikolic et al., 2023) and their generation of CT-related responses (Vasconcelos and dos Santos, 2023, Al Ka'bi, 2023) were also important sources of evidence to support the claim of CT enhancement.

3.3. RQ2: Can the Implementation of AI Technologies Hinder Learners' Development of CT Skills? If So, What Might Be the Possible Reasons?

Our systematic review suggested that the utilization of AI technologies may have potential negative effects on the cultivation of CT skills. The top three identified negative effects on the development of CT skills with AI technologies were as follows: (1) Over-reliance on AI-generated content, involving the unquestioning acceptance of AI output without contextual consideration; (2) Amplification of pre-existing bias and dissemination of false information, intensifying (i.e., exacerbating existing unfair trends of thought and perpetuating false or misleading information); and (3) preventing students from deeper thinking activities, such as reasoning and reflection, or from engaging in skill-augmenting activities such as clinical skills development. For instance, Sallam et al. (2023) found that relying too much on AI could potentially hinder the cultivation of social skills and interaction abilities among clinical students, subsequently affecting their capacity to make informed decisions. Another example can be seen in the research conducted by Mohamed (2023), which indicated that the training of ChatGPT is dependent on pre-defined data, and the not-fully reliable and partial content produced by AI may inadvertently introduce bias and misinformation. This may pose challenges to the advancement of students' CT skills, especially during English language learning.

Six studies reported a negative impact of AI on CT skills which were associated with ChatGPT. Except for one study (Aquino et al., 2023) stated that Healthcare AI can replace the roles of experienced experts by automating clinical tasks, leading to reduced CT skills. In addition to that, one study (Baker et al., 2022) even reported weak or no correlation between the training time of ML algorithms and the accuracy of using CT skills to discriminate neuroimages generated by ML algorithms.

Among these seven studies (Table A2), the most frequently mentioned affordance was providing feedback and answers, followed by offering instructions. For example, Mohamed (2023) found that ChatGPT can serve as a supplementary language learning tool by engaging students in conversations and scaffolding their learning with step-by-step guidance. However, concerns have been raised about the accuracy of responses. Additionally, other functions of AI technologies include provoking discussions, promoting collaboration and assisting academic writing. While textual chatbots can offer immediate responses as well as personalized answers and significantly contribute to learners' knowledge development, their affordance of feedback generation is also highly correlated with over-simplified thinking processes, which may pose a threat to the development of CT skills.

In the current review, analytical ability and clinical decision-making ability were identified as key CT skills that were negatively impacted by AI. Furthermore, some studies also suggested that CT skills, including reflection and reflective thinking, creativity, critical writing ability, and quantitative literacy, have also been impaired. For instance, one study (Sánchez-Ruiz et al., 2023) demonstrated that a series of CT skills, such as analytical ability, reasoning, and quantitative literacy (defined as the ability to do math calculations), can be seriously hindered by excessive use of AI. Apart from analytical ability, Iskender (2023) identified that creativity, which refers to the ability to generate novel ideas (Sternberg and Halpern, 2020), and critical writing ability, defined as the capacity to express CT within written work (Liu, 2006), have also been hindered at the higher education level.

The negative evidence from the literature review against CT development was found either on self-reported data from students and teachers or the analysis of content generated by AI. As a typical instance, Aquino et al. (2023) used interviews to explore experts' understanding of the relationship between AI and CT. However, they detected both positive and negative effects of AI and in the end, they also pointed out that whether AI tends to have positive or negative impacts merits further investigation.

3.4. Can the Development of CT Skills Improve the Effectiveness of Using AI? If So, What Might Be the Possible Reasons?

The enhancement of CT skills is likely to enhance the efficacy of AI in supporting learning, teaching, and research. This effect can be explained by three key factors: (1) The merging of digital literacy and CT definition (2) The synergy between CT skills and AI functionality (3) The intersection of digital affordances and the development of CT skills.

As shown in Table A3, four studies demonstrating this effect are discussed in detail. In the work carried out by Lim (2023), digital literacy was defined as a contemporary manifestation of CT skill, characterized by the ability to collect, arrange, interpret, analyze, assess, and create information using digital technology. A large-scale survey was used as the instrument, and the results revealed that teachers with higher levels of CT skills, namely, digital literacy, tend to use AI technology for education. Notably, the development of CT skills could assist teachers in better grasping the digital affordances of AI, such as facilitating interaction and collaboration and enhancing teaching.

Cooper (2023) assessed the relationship between CT skills, namely analytical ability and decision-making ability, and the use of AI, specifically ChatGPT, through content analysis of AI-generated responses. He found that CT skills were essential to educators who want to adapt AI to maximize its functions in assisting learning and teaching, particularly within specific contexts.

Similarly, Pavlik (2023) also employed content analysis to investigate the relationship between CT skills, particularly creative thinking and critical writing ability, and the application of ChatGPT. The findings indicated that improving CT skills enables individuals to maximize AI's potential because AI generally lacks the ability to think in a critical and creative way,

Henze et al. (2022) deployed interviews to investigate the mutual correlation between CT skills, such as creative thinking ability and scientific reasoning, and the utilization of AI-assisted digital creativity tools for STEAM education. The results showed that secondary teachers attach high importance to the beneficial effects of developing CT skills when using AI technologies. This might be because these tools were purposefully designed to facilitate learning opportunities for CT skills development, and inherently, the enhancement of CT skills can enhance the efficacy of AI utilization.

We found that the vast majority of studies still rely on self-reported data from learners, obtained through surveys and interviews, or content analysis. Ultimately, from these studies, three explanations can be deduced to elucidate the mechanism by which the development of CT skills positively influences AI usage. First, concepts like digital literacy already infuse the ability to use digital technologies with CT, and unsurprisingly, the cultivation of CT skills naturally enhances the utilization of AI technologies. Second, some studies believed that AI technologies, by their nature, lack higher order thinking skills such as CT skills, which are exclusive to human beings. Consequently, as learners develop their CT skills, the deficiencies in AI can be supplemented by human intelligence, resulting in enhanced AI effectiveness. Finally, some AI tools are designed to facilitate the acquisition of CT skills, and their digital affordances align with the cultivation of CT skills, resulting in a symbiotic relationship between the two.

4. Discussion

This systematic review revealed three types of relationships between AI and CT have been identified: (1) AI can help learners develop their CT skills, (2) AI may have negative impacts on the development of learners' CT skills when used inappropriately, and (3) The cultivation of CT skills can also improve the effectiveness of AI in assisting learning, teaching, and research. These mixed results stem from varying definitions of CT skills and different types of AI technologies.

As to the first relationship, the findings reveal that digital affordances are closely connected to the facilitation of CT skills for providing more CT-related prompts and CT training activities to elicit learners' deeper thinking. Additionally, different types of AI technologies are likely to create diversified pathways to guide the CT development process. For example, AI in the medical field normally can streamline repetition, giving learners more time and space to enhance CT, and several studies (Sönmez, 2021; Liang, 2023) that focused on technology-enhanced CT cultivation have underscored the significance of time in the CT improvement process.

With reference to the negative effects of AI on CT cultivation, while most of these negative effects are associated with the use of ChatGPT, it would be unfair to label ChatGPT as the sole source of AI's negative impacts. A more reasonable explanation for this phenomenon is that ChatGPT stands out among AI applications due to its revolutionary processing capacity and performance, which has attracted significant attention to this field. More likely, the improper use of AI should be primarily blamed for causing these negative effects, such as

excessive reliance on AI output and a tendency to reduce social interaction. Nonetheless, it's important to acknowledge that AI can produce biased content, but the key challenge lies in how we address this issue using our CT skills (Pavlik, 2023).

The novelty of this study is also represented in the review of the positive effects of developing CT skills to enhance the effectiveness of AI utilization. And this relationship is associated with the confluence of several factors, such as the combination of digital literacy into CT model, complementarity of CT skills and AI usage, and the symbiosis between AI affordances and cultivation of CT skills.

Defining CT proves challenging due to its multidisciplinary nature. Despite its intricacies, the present study attempts to synthesize existing literature and propose that in the era of AI, CT may generally encompass the abilities of reflective thinking, creativity, analytical ability, problem-solving skills, decision-making ability, critical writing ability, and digital literacy. This also echoed with the call advocated by Long and Magerko (2020), who insisted that digital literacy (specifically AI literacy in their article) should be incorporated into the contemporary CT framework.

However, as it currently stands, most of the existing studies were found to lack the guidance of a CT framework, such as Bloom's Taxonomy or the Paul-Elder Critical Thinking Framework, which may limit the depth of their research. In contrast, Lu and Xie (2022) identified that a majority of the studies they reviewed incorporated a CT framework. The divergence in findings can be attributed to the broader context of their research, which encompasses Information Communication Technology, as well as their linkage of CT development within the domain of English language education—an area that has been extensively investigated (Yin et al., 2023, El Soufi and See, 2019). Given that AI is an emerging technology in education, there has been limited research in this area (Chiu et al., 2023). Consequently, researchers in this field may initially lack the guidance of established theories.

5. Conclusions and Implications for Future Research

This article serves as the pioneering work which presents a holistic understanding of the relationship between AI and CT by examining a vast array of AI technologies with their affordances and differentiated CT skills, ultimately identifying a bi-directional relationship. The findings indicate that generally the AI technologies can assist in developing learners' CT skills but immoderate as well as inappropriate utilization may trigger the negative effects of AI, leading to impediment to CT development. Moreover, the cultivation of CT skills has also been found to have a positive impact on enhancing the effectiveness of using AI.

Nevertheless, we must admit that the present review possesses three limitations. First, the number of included studies examining the relationship between AI and CT was relatively small. This may be attributed to the fact that only English-language articles were considered in this study, and some studies were filtered out because they were inaccessible for downloading electronic files. Second, although the present review comprehensively addresses the negative effects of AI technologies on CT development, discussions about other limitations of AI technologies have been relatively scarce. Given that AI is likely to become increasingly integrated into our daily learning activities, it is advisable to conduct a more comprehensive review with a specific focus on this aspect. Third, meta-analysis was not conducted in the present study because the available studies did not provide sufficient data for effect size calculation. For future researchers, it is recommended to use the method to extrapolate the symbiotic relationship between AI and CT.

In this systematic review, five implications have been discussed. The first implication is that scholars and educators should embrace AI technologies to assist their teaching and research. Simultaneously, they should provide guidance to help students use AI appropriately, taking into account both the positive and negative impacts on CT development. Second, practitioners need to strategize on how to integrate CT instruction into the future curricula that aim to enhance the effectiveness of AI utilization. Third, future studies are encouraged to adopt relevant theoretical frameworks when conducting empirical research to explore the relationship between AI and CT. Fourth, existing studies have primarily focused on developed countries and regions. We suggest that researchers also investigate the status of AI utilization and CT cultivation in other developing countries and regions for educational equity. Fifth, considering the growing recognition of the importance of CT development for younger learners in the AI era (O'Reilly et al., 2022, Yin et al., 2023), it is crucial to explore this aspect further. Last, future research is advisable to adopt more reliable and widely accepted tests to evaluate students' development of CT skills.

Funding

This research was funded by Macao Polytechnic University grant number RP/FCA-08-2023.

Conflicts of Interest

The authors declare no conflict of interest.

Appendix A

Table A1. Studies related to the application of AI technologies to enhance CT skills.

Studies	AI Technologies	Digital Affordances	Improved CT Skills	Measurement and Evidence
Aquino et al. (2023)/ Adult education	Healthcare AI	Automate healthcare work and clinical tasks	Clinical decision-making skills	Interview/Students' self-reported improvement of CT skills
Zammit et al. (2022)/ Primary and secondary education	ArtBot (Game-based AI)	Provoke reflection and discussions	Reflection, reflective thinking and problem solving skills	Learning analytics and students' feedback or freeform responses/ Students' self-reported improvement of CT skills
Rahman and Watanobe (2023)/Higher education	ChatGPT (Textual chatbots powered by AI)	Provide feedback and answers and conduct evaluation	Problem solving skills and analytical ability	Survey/ Students' self-reported improvement of CT skills
Ibna Seraj and Oteir (2022)/Higher education	Replika (Textual chatbots powered by AI)	Facilitate interaction	Analytical ability, reflection and critical writing skills	Survey/ Students' self-reported improvement of CT skills
Henze et al. (2022)/Secondary education	Digital STEAM Creativity Tools	Promote learning opportunities	Creative thinking ability and scientific reasoning	Interview/ Students' self-reported improvement of CT skills
Almazyad et al. (2023)/ Adult education	ChatGPT	Provide feedback and answers, provoke discussions, and promote collaboration	Clinical reasoning and clinical decision-making skills	Interview/ Students' self-reported improvement of CT skills
Mahrous et al. (2023)/ Higher education	AiDental (For removable partial denture design and include a Game - based AI)	Conduct evaluation, and automate healthcare work and clinical tasks	Clinical decision-making skills	Survey/Students' self-reported improvement of CT skills
Vasconcelos and dos Santos (2023)/Higher education	ChatGPT	Provide feedback and answers	Reflection and reflective thinking & problem solving skills	Content analysis/Identification of CT-related reponses generated by AI
Al Ka'bi (2023)/ Higher education	AI algorithm for higher education	Recommend materials and promote learning opportunities	Creativity and creative thinking	Learning analytics/ Detecting higher number of CT related suggestions and recommendations produced by AI
Nikolic et al. (2023)/ Higher education	ChatGPT	Provide feedback and answers	Reflection and reflective thinking	Tests & content analysis/ The great performance of AI on reflective and CT-based writing tests
Tsopra et al. (2023)/ Higher education	Artificial Intelligence-Clinical Decision Support Systems (AI and Web applications)	Automate healthcare work and clinical tasks	Analytical ability and clinical decision-making skills	Survey/ Students' self-reported improvement of CT skills
Porter and Grippa (2020)/ Higher education	Riff platform and the Meeting Mediator tool	Provoke reflection and discussions, provide feedback and conduct evaluation	Reflection and reflective thinking	Survey/Students' self-reported improvement of CT skills
Liu et al. (2023)/ Higher education	Automatic writing evaluation system	Provoke reflection and discussions, assign tasks, provide feedback and conduct evaluation	Reflection and and reflective thinking and critical writing skills	Survey and interview/ Students' self-reported improvement of CT skills

Table A2. Studies related to the application of AI technologies that hinder the development of CT skills.

Studies	AI Technologies	Digital Affordances	Hindered CT Skills	Causes of Hinderance	Measuremen and Evidence
Aquino et al. (2023)/ Adult education	Healthcare AI	Automate healthcare work and clinical tasks	Clinical decision-making skills	Decreased opportunities to maintain clinical skills and automation bias and confirmation bias	Interview/Students' self-reported hinderance of CT skills
Almazyad et al. (2023)/ Adult education	ChatGPT (Textual chatbots powered by AI)	Provide feedback and answers, provoke discussions, and promote collaboration	Clinical reasoning and clinical decision-making skills	Overreliance on AI-generated content	Interview/Students' self-reported hinderance of CT skills
Nikolic et al. (2023)/ Higher education	ChatGPT	Provide feedback and answers	Reflection and reflective thinking	Impediment to reflection process and reinforcement of biased content	Tests & content analysis/Identification of AI-generated responses in relation to barriers to CT skills
Iskender (2023)/ Higher education	ChatGPT	Provide feedback and answers and assist academic writing	Creativity, critical writing ability, and analytical ability	Overreliance on AI-generated content	Content analysis/ Identification of AI-generated responses in relation to barriers to CT skills
Mohamed (2023)/ Higher education	ChatGPT	Provide feedback and answers and offer instructions	Analytical ability and critical writing ability	Reinforcement of biased content	Interview/Teachers' self-reported hindrance to students' CT skills due to AI
Sánchez-Ruiz et al. (2023)/Higher education	ChatGPT	Provide feedback and answers and offer instructions	Quantitative literacy, quantitative reasoning and analytical ability	Overreliance on AI-generated content and impediment to reasoning process	Survey/Students' self-reported hinderance of CT skills
Sallam et al. (2023)/ Higher education	ChatGPT	Provide feedback and answers and offer instructions	Clinical reasoning, clinical decision-making skills and analytical ability	Overreliance on AI-generated content, reinforcement of biased content and lack of social interaction	Content analysis/ Identification of AI-generated responses in relation to barriers to CT skills

Table A3. Studies related to the development of CT skills to promote the use of AI.

Studies	AI Technologies	Digital Affordances	Identified CT Skills	Measurement and Evidence
Lim (2023)/ Preschool education	AI programmes in education	Facilitate interaction and collaboration and assist teaching	Digital literacy	Survey/Teachers' self-reported data reveals that CT improves the effectiveness of AI in assisting learning
Cooper (2023)/ Secondary education	ChatGPT	Scaffold learning and assist writing	Analytical ability and decision-making ability	Content analysis/Analysis of AI-generated responses shows that CT improves the effectiveness of AI in assisting learning and teaching
Henze et al. (2022)/ Secondary education	Digital STEAM Creativity Tools	Promote learning opportunities	Creative thinking ability and scientific reasoning	Interview/Teachers' self-reported emphasis on the positive effect of CT on improving the effectiveness of AI in assisting learning and teaching
Pavlik (2023)/ Higher education	ChatGPT	Provide feedback and answers	creative thinking and critical writing ability	Content analysis/Analysis of AI-generated responses shows that CT improves the effectiveness of AI in assisting learning, teaching and research

References

Ahern, A., Dominguez, C., McNally, C., O'Sullivan, J. J., & Pedrosa, D. (2019). A literature review of critical thinking in engineering education. *Studies in Higher Education*, *44*(5), 816–828.

Al Ka'bi, A. (2023). Proposed artificial intelligence algorithm and deep learning techniques for development of higher

education. *International Journal of Intelligent Networks*, *4*, 68–73.

Almazyad, M., Aljofan, F., Abouammoh, N. A., Muaygil, R., Malki, K. H., Aljamaan, F., Alturki, A., Alayed, T., Alshehri, S. S., Alrbiaan, A., Alsatrawi, M., Temsah, H. A., Alsohime, F., Alhaboob, A. A., Alabdulhafid, M., Jamal, A., Alhasan, K., Al-Eyadhy, A., & Temsah, M. H. (2023). Enhancing Expert Panel Discussions in Pediatric Palliative Care: Innovative Scenario Development and Summarization with ChatGPT-4. *Cureus*, *15*(4), e38249.

Anggraeni, D. M., Prahani, B. K., Suprapto, N., Shofiyah, N., & Jatmiko, B. (2023). Systematic review of problem based learning research in fostering critical thinking skills. *Thinking Skills and Creativity*, *49*, 101334.

Aquino, Y. S. J., Rogers, W. A., Braunack-Mayer, A., Frazer, H., Win, K. T., Houssami, N., Degeling, C., Semsarian, C., & Carter, S. M. (2023). Utopia versus dystopia: Professional perspectives on the impact of healthcare artificial intelligence on clinical roles and skills. *International Journal of Medical Informatics*, *169*, 104903.

Baker, M. R., Padmaja, D. L., Puviarasi, R., Mann, S., Panduro-Ramirez, J., Tiwari, M., & Samori, I. A. (2022). Implementing Critical Machine Learning (ML) Approaches for Generating Robust Discriminative Neuroimaging Representations Using Structural Equation Model (SEM). *Comput Math Methods Med*, *2022*, 6501975.

Bearman, M., Ryan, J., & Ajjawi, R. (2023). Discourses of artificial intelligence in higher education: A critical literature review. *Higher Education*, *86*(2), 369–385.

Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *4*, 100118.

Chou, T.-L., Wu, J.-J., & Tsai, C.-C. (2019). Research Trends and Features of Critical Thinking Studies in E-Learning Environments: A Review. *Journal of Educational Computing Research*, *57*(4), 1038–1077.

Chu, H.-C., Hwang, G.-H., Tu, Y.-F., & Yang, K.-H. (2022). Roles and research trends of artificial intelligence in higher education: A systematic review of the top 50 most-cited articles. *Australasian Journal of Educational Technology*, *38*(3), 22–42.

Cooper, G. (2023). Examining Science Education in ChatGPT: An Exploratory Study of Generative Artificial Intelligence. *Journal of Science Education and Technology*, *32*(3), 444–452.

Drennan, J. (2010). Critical thinking as an outcome of a Master's degree in Nursing programme. *Journal of Advanced Nursing*, *66*(2), 422–431.

Dwyer, C. P., Hogan, M. J., & Stewart, I. (2014). An integrated critical thinking framework for the 21st century. *Thinking Skills and Creativity*, *12*, 43–52.

El Soufi, N., & See, B. H. (2019). Does explicit teaching of critical thinking improve critical thinking skills of English language learners in higher education? A critical review of causal evidence. *Studies in Educational Evaluation*, *60*, 140–162.

Ennis, R. H. (1987). Critical thinking and the curriculum. *National Forum: Phi Kappa Phi Journal*, *65*(1), 28–31.

Facione, P. A. (1990). Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction. In *Research Findings and Recommendations*. The California Academic Press.

Ghanizadeh, A. (2017). The interplay between reflective thinking, critical thinking, self-monitoring, and academic achievement in higher education. *Higher Education*, *74*(1), 101–114.

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *American Psychologist*, *53*(4), 449–455.

Hao, Z., Fang, F., & Peng, J.-E. (2024). The integration of AI technology and critical thinking in English major education in China: Opportunities, challenges, and future prospects. *Digital Applied Linguistics*, *1*, 2256. https://doi.org/10.29140/dal.v1.2256.

Henze, J., Schatz, C., Malik, S., & Bresges, A. (2022). How Might We Raise Interest in Robotics, Coding, Artificial Intelligence, STEAM and Sustainable Development in University and On-the-Job Teacher Training? *Frontiers in Education*, *7*, 872637.

Hong, S., & Yu, P. (2017). Comparison of the effectiveness of two styles of case-based learning implemented in lectures for developing nursing students' critical thinking ability: A randomized controlled trial. *International Journal of Nursing Studies*, *68*, 16–24.

Huang, J., & Sang, G. (2023). Conceptualising critical thinking and its research in teacher education: A systematic review. *Teachers and Teaching*, *29*(6), 638–660.

Ibna Seraj, P. M., & Oteir, I. (2022). Playing with AI to Investigate Human-Computer Interaction Technology and Improving Critical Thinking Skills to Pursue 21stCentury Age. *Education Research International*, *2022*, 6468995.

Iskender, A. (2023). Holy or Unholy? Interview with Open AI's ChatGPT. *European Journal of Tourism Research*, *34*, 3414.

Ji, H., Han, I., & Ko, Y. (2023). A systematic review of conversational AI in language education: Focusing on the collaboration with human teachers. *Journal of Research on Technology in Education*, *55*(1), 48–63.

Jin, S.-H., Im, K., Yoo, M., Roll, I., & Seo, K. (2023). Supporting students' self-regulated learning in online learning using

artificial intelligence applications. *International Journal of Educational Technology in Higher Education*, *20*(1), 37.

Lee, D. S., Abdullah, K. L., Subramanian, P., Bachmann, R. T., & Ong, S. L. (2017). An integrated review of the correlation between critical thinking ability and clinical decision-making in nursing. *Journal of Clinical Nursing*, *26*, 4065–4079.

Lee, Y.-F., Hwang, G.-J., & Chen, P.-Y. (2022). Impacts of an AI-based chabot on college students' after-class review, academic performance, self-efficacy, learning attitude, and motivation. *Educational Technology Research and Development*, *70*(5), 1843–1865.

Liang, W. (2023). Towards a set of design principles for technology-assisted critical-thinking cultivation: A synthesis of research in English language education. *Thinking Skills and Creativity*, *47*, 101203.

Lim, E. M. (2023). The effects of pre-service early childhood teachers' digital literacy and self-efficacy on their perception of AI education for young children. *Education and Information Technologies*, *28*(10), 12969–12995.

Liu, C.-C., Liu, S.-J., Hwang, G.-J., Tu, Y.-F., Wang, Y., & Wang, N. (2023). Engaging EFL students' critical thinking tendency and in-depth reflection in technology-based writing contexts: A peer assessment-incorporated automatic evaluation approach. *Education and Information Technologies*, *28*(10), 13027–13052.

Liu, K. (2006). Annotation As an Index to Critical Writing. *Urban Education*, *41*(2), 192–207.

Long, D., & Magerko, B. (2020, April 25–30). *What is AI Literacy? Competencies and Design Considerations* [Conference session]. 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA.

Lu, D., & Xie, Y.-N. (2022). Critical thinking cultivation in TESOL with ICT tools: A systematic review. *Computer Assisted Language Learning*, *37*, 222–242.

Ludin, S. M. (2018). Does good critical thinking equal effective decision-making among critical care nurses? A cross-sectional survey. *Intensive and Critical Care Nursing*, *44*, 1–10.

Mahrous, A., Botsko, D. L., Elgreatly, A., Tsujimoto, A., Qian, F., & Schneider, G. B. (2023). The use of artificial intelligence and game-based learning in removable partial denture design: A comparative study. *Journal of Dental Education*, *87*(8), 1188–1199.

Marks-Anglin, A., & Chen, Y. (2020). A historical review of publication bias. *Research Synthesis Methods*, *11*(6), 725–742.

Melisa, R., Ashadi, A., Triastuti, A., Hidayati, S., Salido, A., Ero, P. E. L., Marlini, C., Zefrin, Z., & Al Fuad, Z. (2025). Critical Thinking in the Age of AI: A Systematic Review of AI's Effects on Higher Education. *Educational Process: International Journal*, *14*, e2025031.

Mohamed, A. M. (2023). Exploring the potential of an AI-based Chatbot (ChatGPT) in enhancing English as a Foreign Language (EFL) teaching: Perceptions of EFL Faculty Members. *Education and Information Technologies*, *29*(3), 3195–3217.

Nazaretsky, T., Ariely, M., Cukurova, M., & Alexandron, G. (2022). Teachers' trust in AI-powered educational technology and a professional development program to improve it. *British Journal of Educational Technology*, *53*(4), 914–931.

Nikolic, S., Daniel, S., Haque, R., Belkina, M., Hassan, G. M., Grundy, S., Lyden, S., Neal, P., & Sandison, C. (2023). ChatGPT versus engineering education assessment: A multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity. *European Journal of Engineering Education*, *48*(4), 559–614.

O'Reilly, C., Devitt, A., & Hayes, N. (2022). Critical thinking in the preschool classroom—A systematic literature review. *Thinking Skills and Creativity*, *46*, 101110.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, n71.

Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the Implications of Generative Artificial Intelligence for Journalism and Media Education. *Journalism and Mass Communication Educator*, *78*(1), 84–93.

Porter, B., & Grippa, F. (2020). A platform for ai-enabled real-time feedback to promote digital collaboration. *Sustainability*, *12*(24), 10243.

Puig, B., Blanco Anaya, P., & Bargiela, I. M. (2020). A Systematic Review on E-learning Environments for Promoting Critical Thinking in Higher Education. In M. J. Bishop, E. Boling, J. Elen, & V. Svihla (Eds.), *Handbook of Research in Educational Communications and Technology: Learning Design* (pp. 345–362). Springer International Publishing.

Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for Education and Research: Opportunities, Threats, and Strategies. *Applied Sciences*, 13(9), 5783.

Sallam, M., Salim, N., Barakat, M., & Al-Tammemi, A. (2023). ChatGPT applications in medical, dental, pharmacy, and public health education: A descriptive study highlighting the advantages and limitations. *Narra J*, *3*(1), e103.

Sánchez-Ruiz, L. M., Moll-López, S., Nuñez-Pérez, A., Moraño-Fernández, J. A., & Vega-Fleitas, E. (2023). ChatGPT Challenges Blended Learning Methodologies in Engineering Education: A Case Study in Mathematics. *Applied Sciences*, *13*(10), 6039.

Schraw, G., McCrudden, M. T., Lehman, S., & Hoffman, B. (2011). An overview of thinking skills. In G. Schraw, & D. H. Robinson (Eds.), *Assessment of higher order thinking skills* (pp. 19–45). Information Age Publishing.

Sönmez, E. (2021). Technology-Enhanced CT: A Systematic Review. *Thinking Skills and Creativity*, *41*, 100913.

Sternberg, R. J., & Halpern, D. F. (2020), *Critical Thinking in Psychology*, Cambridge University Press, Cambridge.

Suligoj, V., Zavbi, R., & Avsec, S. (2020). Interdisciplinary Critical and Design Thinking. *International Journal of Engineering Education*, *36*(1), 84–95.

Tsopra, R., Peiffer-Smadja, N., Charlier, C., Campeotto, F., Lemogne, C., Ruszniewski, P., Vivien, B., & Burgun, A. (2023). Putting undergraduate medical students in AI-CDSS designers' shoes: An innovative teaching method to develop digital health critical thinking. *International Journal of Medical Informatics*, *171*, 104980.

Vasconcelos, M. A. R., & dos Santos, R. P. (2023). Enhancing STEM learning with ChatGPT and Bing Chat as objects to think with: A case study. *Eurasia Journal of Mathematics*, *Science and Technology Education*, *19*(7), em2296.

Wang, Y.-M., Wei, C.-L., Lin, H.-H., Wang, S.-C., & Wang, Y.-S. (2022). What drives students' AI learning behavior: A perspective of AI anxiety. *Interactive Learning Environments*, *32*(6), 2584–2600.

Wang, Y., & Wu, Z. (2023). Adapting or adopting? Critical thinking education in the East Asian cultural sphere: A systematic integrative review. *Thinking Skills and Creativity*, *49*, 101330.

Yang, R., & Wei, W. (2023). A systematic review of language MOOCs: Research design, evidence, and implications. *Journal of Educational Technology and Innovation*, *5*(2), 33–47.

Yin, X., Saad, M. R. B. M., & Halim, H. B. A. (2023). A systematic review of critical thinking instructional pedagogies in EFL writing: What do we know from a decade of research. *Thinking Skills and Creativity*, *49*, 101363.

Yu, L., & Zin, Z. M. (2023). The critical thinking-oriented adaptations of problem-based learning models: A systematic review. *Frontiers in Education*, *8*, 1139987.

Yuan, R., Liao, W., Wang, Z., Kong, J., & Zhang, Y. (2022). How do English-as-a-foreign-language (EFL) teachers perceive and engage with critical thinking: A systematic review from 2010 to 2020. *Thinking Skills and Creativity*, *43*, 101002.

Zammit, M., Voulgari, I., Liapis, A., & Yannakakis, G. N. (2022). Learn to Machine Learn via Games in the Classroom. *Frontiers in Education*, *7*, 913530.

Zhang, R., Zou, D., & Cheng, G. (2023). A review of chatbot-assisted learning: Pedagogical approaches, implementations, factors leading to effectiveness, theories, and future directions. *Interactive Learning Environments*, *32*(8), 4529–4557.

Zhao, C. (2024). AI-assisted assessment in higher education: A systematic review. *Journal of Educational Technology and Innovation*, 6(4), 39–58.

Zou, D., Luo, S., Xie, H., & Hwang, G.-J. (2022). A systematic review of research on flipped language classrooms: Theoretical foundations, learning activities, tools, research topics and findings. *Computer Assisted Language Learning*, *35*(8), 1811–1837.