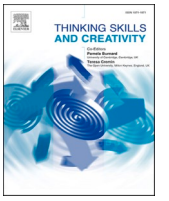
Thinking Skills and Creativity 57 (2025) 101850

Contents lists available at ScienceDirect

Thinking Skills and Creativity

journal homepage: www.elsevier.com/locate/tsc



Critical literacy in artificial intelligence assisted writing instruction: A systematic review

Tülin Arseven, Mazhar Bal *

Department of Turkish and Social Sciences Education, Akdeniz University, Antalya, Turkey

ARTICLE INFO

Keywords:
Writing skills
Artificial intelligence
Artificial intelligence-assisted writing
Critical literacy
K-12 education
Systematic review

ABSTRACT

This systematic review aims to examine the relationship between artificial intelligence (AI)-assisted writing instruction and critical literacy skills of K-12 students. As a result of searching Web of Science (WoS), Scopus, Education Resources Information Center (ERIC), and EBSCO databases in accordance with the criteria determined, 15 studies were included in the analysis. As a result of the descriptive analysis, five main themes were identified: critical content production and evaluation, metacognitive skills and self-regulation, ethical thinking and evaluation, analytical thinking and problem solving, motivation and self-efficacy for the AI-assisted writing process and critical literacy skills. Our findings show that students use AI tools not as a simple copy-paste tool, but as a supportive element to develop their thinking. Students question AI outputs, synthesize them with their own ideas and manage their writing processes more consciously. The findings of our study reveal that AI-assisted writing experiences increase students' ethical awareness of originality, academic honesty and creativity, improve their systematic thinking skills in complex writing tasks, and increase their motivation to write. However, the lack of any study at preschool level, the limited number of studies at primary school level, and the scarcity of qualitative studies indicate that more research is needed in this field.

1. Introduction

Writing is a fundamental skill closely linked to academic achievement, everyday communication, formal negotiations, and more. However, learning to write has been a challenging and complex cognitive process for most students due to a lack of contextual experience (Chen et al., 2024). The act of writing is an important tool for human self-expression in all its aspects (Kaya & Ateş, 2016). Writing is a powerful tool for thinking. It is a skill that covers all disciplines. It is important for an individual's career goals regardless of degree, program and profession (Tasic & Hannah, 2017). It requires the learner to synchronize the demands of idea generation, planning, text production, spelling, punctuation, grammar, self-monitoring, evaluation, and orthographic-motor integration (Berninger, 2019). It reflects an individual's thoughts and attitude about something, and writing also constitutes thinking (Sani, Kurniawati & Nurwanti, 2018). In this respect, it involves both cognitive and meta-cognitive skills (Larkin, 2009). Therefore, general knowledge and cognitive factors are important in developing writing skills (Barbot et al., 2012).

Writing requires the simultaneous control of many components such as organizing thoughts in a logical sequence, appropriate word choice, and compliance with grammar rules; it is the last language skill acquired and the most difficult to develop (Demirel, 2003; Kellogg & Whiteford, 2009). Developing students' critical thinking skills in writing is important for different educational contexts

* Corresponding author at: Department of Turkish and Social Sciences Education, Akdeniz University, Antalya, Turkey.
E-mail address: balmazhar@gmail.com (M. Bal).

https://doi.org/10.1016/j.tsc.2025.101850
Received 19 February 2025; Received in revised form 23 March 2025; Accepted 18 April 2025
Available online 22 April 2025

(Ngajie et al., 2018). Recent research shows that more and more attention is being paid to the development of young children's early writing skills (and students completing K-12 in the US are expected to be proficient in writing (Li & Zhang, 2015, Zhang, et al., 2015). Writing is considered a very important skill in education and work life). In the digital age, the meaningful realization of writing processes in the K-12 period has an important place for both the school and the future lives of individuals (Bal & Öztürk, 2025).

The emergence of artificial intelligence tools has also initiated changes in education systems. Indiscriminate use of AI can have a negative impact on oral and written communication and social skills (Mirón-Mérida & García-García, 2024). Thus, Blaizot et al. (2022), in their study evaluating the use of AI methods for systematic reviews, stated that extensive human validation is still necessary when applying AI methods. Purposeful use, on the other hand, can be beneficial in many ways. The integration of AI and IoT technologies in different fields has also attracted attention in the literature. For example, Lam et al. (2021) showed that the integration of AI and IoT technologies in healthcare during COVID-19 has made significant contributions to the real-time monitoring of the health status of older people and improving service quality. Applying similar approaches to education can lead to innovative learning experiences. For example, automatic writing assessment tools have positive contributions to writing skills (Parra & Calero, 2019). AI-assisted writing tools facilitate more effective writing performance improvements by providing automatic feedback on various aspects of writing, including word choices, grammar, and vocabulary beyond punctuation and spelling checks (Chen, 2023). Similarly, in different educational fields such as driver training, it has been shown that combining simulator-based training with AI analysis can effectively detect and correct problematic behaviors, highlighting the value of AI-assisted training in various learning contexts (Tang et al., 2025). Students receiving AI-assisted instruction show improved proficiency in various aspects of writing, including organization, coherence, grammar, and vocabulary (Song & Song, 2023). In this respect, artificial intelligence tools should take their place in the future of education instead of being banned/rejected (Lim, et.al, 2023). In addition, in relation to the purpose of our study, there are significant correlations between critical thinking skills, self-directed learning efficacy and AI-assisted writing (Shen & Teng, 2024). Artificial intelligence challenges traditional methods in developing critical thinking skills (Gonsalves, 2024). In addition, AI can both improve and complicate critical thinking in cognitive, affective and meta-cognitive domains (Gonsalves, 2024). The most common challenges to AI adoption are transparency, explainability, bias, data privacy, robustness, and reliability (Kokina et al., 2025).

1.1. Argumentation and critical thinking in artificial intelligence assisted writing

In the writing process, students may face higher-level challenges that require more complex cognitive processes beyond technical challenges such as grammar and word choice. These challenges include developing a coherent argument, constructing a logical structure, evaluating evidence, and conducting a critical analysis (Safitri & Adani, 2024). This is because writing skills involve basic writing processes, such as planning, language production, and revision, as well as the mental processes necessary for its creation and retention in working memory (Kellogg, 2008). This indicates that writing is a product of complex thought processes.

Critical thinking and argumentation skills are central to effective writing. Wu et al. (2024) showed that computational thinking improves critical thinking and problem solving skills by giving students the ability to analyze and evaluate text, identify patterns, and develop creative solutions to writing problems. This approach helps students better understand the structure and content of the text by breaking down complex writing tasks into smaller and more manageable chunks. A lack of these skills may cause students to experience difficulties in the writing process. Mallahi (2024) categorized the problems that students experience while composing argumentative texts as cognitive, grammatical, discursive, and psychological problems. According to him, these problems are associated with the fact that the argumentative writing genre is new and complex for students, their insufficient grammatical and discursive competence, and their inability to use writing strategies effectively. In addition, critical thinking plays an important role in the construction of argumentative discourse (Handayani 2016). Teng and Yue (2023) found that students who developed a certain level of metacognitive awareness were more inclined to think critically and this resulted in better writing performance. In addition, Yilmaz-Na and Sönmez's (2023) research on critical thinking and argumentation skills from a different perspective revealed that a computer-assisted argument-mapping tool improved students' critical thinking and ability to create better argument maps.

Different types of technologies have been the focus of other studies. In their study of AI technologies, Latifi et al. (2023) stated that these technologies can affect students' argumentative essay writing performance in both positive and negative ways. Although the negative aspects of AI-supported writing are mentioned in previous studies, the positive aspects are more prominent. For example, Kim et al. (2025) found that generative AI-supported writing benefits students in three areas: the writing process, writing performance, and affective domains. According to them, students indicated that their higher-order thinking skills, such as logical reasoning, critical questioning, and evaluating the content produced by generative AI, were also important in overcoming the challenges associated with AI. These studies focusing on AI technologies in writing instruction suggest that more detailed research is needed in areas where students' AI-assisted writing and critical thinking skills intersect.

There are some studies on the effects of using artificial intelligence on writing skills and critical thinking. Some of the studies (Chen, 2023; Chen & Teng, 2024; Song & Song, 2023) focused on artificial intelligence and writing. Gonsalves (2024) examined the relationship between AI and critical thinking. Many research articles (Larkin, 2009; Zhang, et al., 2015) have focused on writing skills and the writing process as a metacognitive skill. There are also review studies that examine the relationship between writing skills and artificial intelligence. For example, Albadarin et al. (2024) found that ChatGPT is a valuable tool for improving academic writing skills and language proficiency. Similarly, Feng Teng, (2024) studied the role of ChatGPT in the development of English as a foreign language writing skills. There are also studies directly related to critical thinking in AI-assisted writing processes (Janse van Rensburg, 2024; Liu, et al., 2023; Sharma, 2024 and so forth). However, the participant groups of these studies are not at the K-12 level. The participants of most of the studies on artificial intelligence and writing processes are university-level students. Our study is a systematic analysis of research articles on the effects of artificial intelligence tools on both writing and critical literacy skills at the K-12 level. It

fills the gap in the literature on the relationship between artificial intelligence and writing and critical thinking skills for school-age children by holistically addressing the studies that address the relationship between artificial intelligence and writing and critical thinking skills at the K-12 level with the systematic analysis method.

1.2. Purpose of the study and research question

This systematic review aims to examine in detail the studies on the relationship between AI-assisted writing instruction and critical literacy skills of K-12 students. In line with this purpose, the research question of the study is as follows:

- What is the relationship between AI-assisted writing instruction and critical literacy skills in the identified studies?

2. Method

2.1. Research design

This research is a systematic literature review designed and conducted in accordance with PRISMA 2020 guidelines (Page et al., 2021). In this way, a transparent, reproducible, and comprehensive synthesis of the research findings is presented. This systematic review is one of the pioneering studies to comprehensively examine the relationship between AI-assisted writing and cultural literacy in K-12 education.

2.2. Search strategy

The literature search for this systematic review was conducted using the WoS, Scopus, ERIC, and EBSCO databases. These databases were selected for several reasons: (1) they represent the most comprehensive collection of peer-reviewed educational research, covering both international and interdisciplinary publications; (2) WoS and Scopus are known for indexing high-impact journals with rigorous peer-review processes that ensure the quality of the included research; (3) ERIC covers research that is directly specific to our field, with a particular focus on educational research; and (4) EBSCO complements these by offering access to educational practice literature that may not be available through the other three databases included in the search. These databases provide a balanced coverage of theoretical, empirical, and practical research in educational technology without significant overlap. Other databases were excluded either because they contained predominantly non-peer-reviewed publications or because their content was largely duplicative of the work in our selected databases.

The keyword selection process was developed systematically based on the four conceptual pillars necessary for our research question. These conceptual domains are artificial intelligence, writing skills, critical literacy, and K-12 education. To conduct a comprehensive search for each conceptual domain, primary terms were identified, along with their common variations and related concepts. The artificial intelligence category includes both general terms (e.g., "artificial intelligence,'' "AI") and specific applications (e.g., "natural language processing,'' "generative AI") to capture the technological conceptual space related to writing instruction. The critical literacy category covers various aspects of higher order thinking skills (e.g. "critical thinking,'' "critical evaluation,'' "critical reasoning") to account for the different terminology used in educational contexts. The writing category focuses on instructional aspects, with terms such as "writing instruction" and "writing skills". The K-12 category included all relevant educational levels, from "kindergarten" to "high school, " to ensure that appropriate age ranges were captured. All keyword variations were validated through preliminary search. It was also developed in consultation with educational technology and literacy education subject matter experts to maximize the precision of the search strategy. The search index used in the databases is as follows:

> ALL= (("artificial intelligence" OR "AI" OR "intelligent systems" OR "AI applications" OR "AI technology" OR "AI systems" OR "AI tools" OR "artificial neural networks" OR "smart systems" OR "intelligent automation" OR "natural language processing" OR "NLP" OR "generative AI") AND ("critical literacy" OR "critical thinking" OR "critical reading" OR "critical analysis" OR "critical evaluation" OR "critical understanding" OR "critical pedagogy" OR "critical skills" OR "critical comprehension" OR "critical interpretation" OR "critical awareness" OR "critical interpretation" OR "critical reflection" OR "critical reasoning" OR "critical inquiry" OR "critical examination" OR "critical assessment" OR "critical discourse" OR "critical perspective" OR "critical approach" OR "critical media literacy" OR "critical digital literacy" OR "critical information literacy" OR "critical text analysis" OR "critical discourse analysis") AND (“writing” OR “writing education” OR “writing skill” OR “writing teaching”) AND ("K-12″ OR "kindergarten" OR "primary school" OR "elementary school" OR "middle school" OR "secondary school" OR "high school" OR "student*"))

The initial search were found 111 documents in WoS, 9725 in Scopus, 88 in ERIC, and 49 in EBSCO. Considering the number of documents, the number of documents accessed when "search within" is set as "All fields" in Scopus is quite high. In addition, as a result of the preliminary scanning, it was determined that there were studies far beyond the purpose of the study in these documents. For this reason, "search within" was set as "Article title, Abstract, Keywords". Thus, the documents directly serving the purpose of the study were searched. As a result of this search, 173 documents were found in Scopus. No date restrictions were made in the databases and documents up to 28.01.2025 were included. The distribution of the search results according to document types is as follows:

- WoS: 76 article, 27 prooceding paper, 14 early access, 2 editorial, 6 review article

- Scopus: 100 article, 46 conference paper, 11 book chapter, 4 review, 4 note, 4 book, 2 letter, 1 editorial, 1 conference review
- ERIC: 68 article, 9 conference paper, 10 report, 1 interview
- EBSCO: 44 article, 3 conference paper, 2 document

Only peer-reviewed research articles were included in the study. The reasons for this selection are: (1) peer-reviewed articles have undergone a rigorous academic review process, (2) transparent and detailed reporting of research methods, (3) based on original research findings, and (4) ease of access to full text. As a result of the screening according to these criteria, a total of 294 articles were found. Language restriction was applied for accurate analysis and interpretation of research findings. 75 articles in WoS, 97 in Scopus, 68 in ERIC and 42 in EBSCO are in English. After including only English articles, 202 articles were finally identified.

2.3. Inclusion and exclusion criteria

The inclusion and exclusion criteria used in this systematic review were developed in line with the PRISMA 2020 guidelines. The inclusion and exclusion criteria were carefully developed to ensure methodological rigor and relevance to our research question on critical literacy in AI-assisted writing instruction in K-12 education. Database selection criteria were established to focus on quality, peer-reviewed sources in internationally recognized journals with high impact factors. Publication type was restricted to full-text research articles to ensure comprehensive reporting of methods and findings, as other types of publications (e.g., conference proceedings, book chapters) can often provide abbreviated descriptions where methodological details are omitted. The language criterion (English only) was applied for practical reasons for accurate analysis and to avoid misinterpretation during translation. To ensure the quality and reliability of the findings for the methodological criteria, we only included studies with a clearly articulated research design and detailed data collection and analysis procedures. The participant criterion (K-12 students) was necessary to directly address our research focus on school-age students as developmental considerations in writing, and critical thinking skills differed significantly between K-12 and adult students. Finally, the content criteria focused on the intersection of AI, writing skills, critical literacy, and K-12 education to make a direct connection to our research question. These carefully formulated criteria allowed us to exclude indirectly relevant or methodologically weak studies that did not contribute significantly to answering our research question. However, it helped us identify the most relevant and methodologically sound studies that were most relevant to our research question. They are included in Table 1:

As seen in Table 1, studies were included or excluded according to the titles of database, publication type, language, method, participant, and content. Database, publication type, and language were performed automatically on the database. Method, participant and content were performed manually by the researchers.

2.4. Study selection process

The flow card for the systematic study selection process is shown in Fig. 1:

As shown in Fig. 1, four different databases were used. In the first search, 421 documents were identified. In the screening process, only 282 articles that were not written in English were evaluated. Subsequently, 103 irrelevant articles were excluded as a result of 102 duplicate, title and abstract searches. The full text of the remaining 75 articles was analysed in detail in line with the inclusion and exclusion criteria. At this stage, a total of 62 articles were eliminated, including 37 articles that did not cover K-12 students, 6 articles that did not focus on writing and digital literacy skills, 17 articles that were reviews, and 1 article whose full text could not be accessed. After a rigorous screening and evaluation process, 15 articles that met the criteria of our study were included in the final analysis. To ensure the objectivity of the selection process, the evaluations were conducted by two independent researchers and inter-rater reliability was calculated using Cohen's Kappa (Cohen, 1960) coefficient ($\kappa = 0.74$). We resolved disagreements between the reviewers by seeking the opinion of a third expert.

2.5. Data extraction

Within the scope of the systematic review, a comprehensive form development process was followed to ensure the reliability and validity of the data extraction process. PRISMA 2020 guidelines were taken as a basis for form development. Categories were created to

Table 1
Inclusion and Exclusion Criteria.

Features	Inclusion Criteria	Exclusion Criteria
Database	Articles indexed in WoS, Scopus, ERIC, EBSCO	Articles no indexed in WoS, Scopus, ERIC, EBSCO
Publication	Full-text research articles	Prooceding paper, book, book chapters, review article, conference paper, conference review, rapor
Language	Articles in English language only	Articles in langauges other than English language
Method	Clearly stated research Detailed data collection and analyses porcedures	Review article, Literature review and research not clearly
Participants	Participants consisting of K-12 students	Participants do not consist of K-12 students
Content	Studies related to artificial intelligence, writing skill, critical literacy, K-12 education	Studies unrelated to artificial intelligence, writing skill, critical literacy, K-12 education

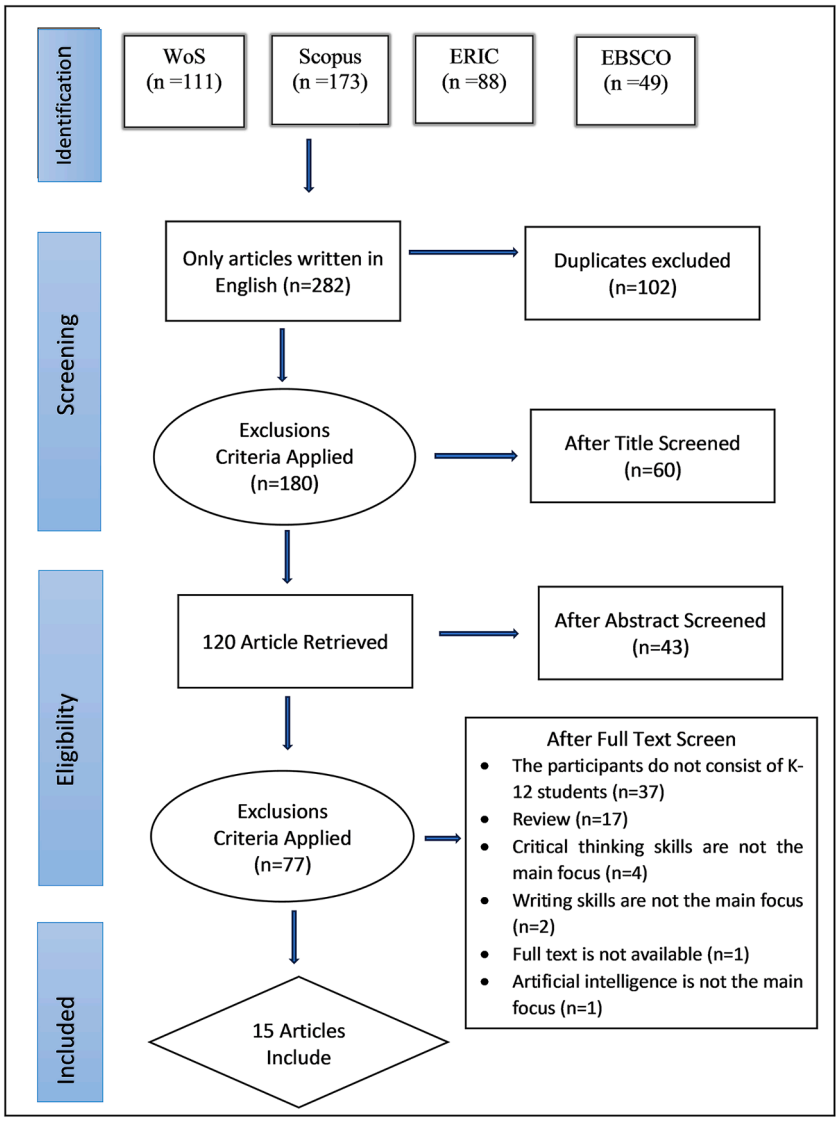


Fig. 1. Flow diagram of the article selection process.

examine the relationship between AI-assisted writing skills and critical literacy in K-12 education. The form development process was carried out in three stages. In the first stage, the draft form consisting of seven basic categories obtained from the literature review was evaluated by two experts. In the second stage, new subheadings were added and elaborated in line with the expert opinions. In the last stage, the form was pilot tested on three articles and the form was finalized by making the necessary adjustments.

This systematic development process resulted in a structured data extraction form that allows for a comprehensive examination of the relationship between AI-assisted writing and critical literacy in K-12 education. Two researchers analysed the articles independently. Inter-coder reliability for descriptive data was assessed by calculating the percentage of agreement (Miles & Huberman, 1994). For the points coded differently, the researchers came together and reached a consensus. Consensus was achieved at a level of 85 % in descriptive data. While creating the findings section, the data were classified thematically in line with the research question and this classification was made independently by two researchers and compared.

In order to assess reporting bias in the systematic review, inclusion/exclusion criteria were rigorously applied to prevent selection bias, two independent researchers' evaluation ($\kappa = 0.74$) and disagreements were resolved by a third expert opinion. We preferred to use Web of Science, Scopus, ERIC and EBSCO databases together to avoid reporting bias. Mixed Methods Appraisal Tool (MMAT) Version 2018 was used to control methodological bias. Good agreement ($\kappa = 0.80$) was obtained between researchers in the MMAT assessment. The majority of included studies (60 %) were of high quality, supporting the reliability of the findings of this systematic review. Studies of moderate quality also met acceptable methodological standards. There were no studies rated as low quality, indicating that the overall quality of the included studies was good. In addition, in order to minimize geographical and cultural bias, the existing situation is shown without any geographical limitation.

2.6. Quality assessment

The method quality of included studies was assessed using the Mixed Methods Appraisal Tool (MMAT) Version 2018 (Hong et al., 2018). The MMAT is a reliable tool designed to assess the method quality of qualitative, quantitative and mixed methods research in systematic reviews. The quality assessment of the studies was carried out by two independent assessors who are experts in the field. The assessors examined each study according to the MMAT criteria and scored each study on a 0–5 point scale. Cohen's Kappa coefficient was calculated to determine inter-rater reliability. As a result of the analysis, a good level of agreement ($\kappa = 0.82$) was found between the raters. In twelve of the fifteen articles (80 %), full agreement was observed between the raters, while in the remaining three articles, only one point difference was found. There were no assessments that differed by two or more points. As a result of the MMAT assessment, the methodological quality of the studies was categorized into three levels.

In the evaluation, the studies were classified as high quality studies (1st, 2nd, 3rd, 4th, 5th, 6th, 8th, 10th, 13th articles; $n = 9$) and medium quality studies (7th, 9th, 11th, 12th, 14th and 15th articles; $n = 6$). There were no studies categorized as low quality. The majority of the included studies (60 %) were of high quality, supporting the reliability of the findings of this systematic review. Medium-quality studies also met acceptable methodological standards. The lack of studies rated as low quality indicates that the overall quality of the included studies was good.

2.7. Data analysis and synthesis

In this study, we chose the descriptive analysis method to examine the collected information in an orderly manner and conducted our study in three stages. First, we carefully analysed the data extraction form to understand how the use of artificial intelligence in writing education affects students' critical thinking skills. Then, we examined the findings of the 15 studies we selected in detail. Finally, we grouped the information we obtained from all these examinations according to similar topics and grouped them under main headings. As a result of the descriptive analysis of the research, five main themes were identified that reveal the relationship between AI-assisted writing processes and critical literacy skills of K-12 students: (1) Critical Content Generation and Evaluation in AI-Assisted Writing Process, (2) Metacognitive Skills and Self-Regulation in AI-Assisted Writing Process, (3) Ethical Thinking and Evaluation in AI-Assisted Writing Process, (4) Analytical Thinking and Problem Solving in AI-assisted Writing Process, (5) Motivation and Self-Efficacy in AI-Assisted Writing Process. We explained the findings in each theme in an interconnected way and supported them with direct quotations from the studies. We then interpreted these findings in detail and discussed what they meant. In doing so, we both considered the purpose of our research and compared it with previous studies on the subject. To ensure the reliability of our study, two researchers first analysed the data separately. In the first evaluation using Miles and Huberman's formula, the agreement between us was 0.70. Then we came together to discuss the points where we disagreed and formed a common opinion. For some cases where we were undecided, we consulted the opinion of a third researcher. As a result of these interviews, the agreement between us increased to 0.85.

2.8. Trustworthiness and reliability

To ensure the validity of this systematic review, we considered the criteria of credibility, transferability, dependability, consistency and confirmability suggested by Lincoln and Guba (1985) for qualitative research.

First, we compared the analyses we conducted separately as two researchers using Miles and Huberman's formula. Then, we asked a third researcher experienced in the field of artificial intelligence and education to review our entire analysis process and findings. Finally, we supported our interpretations with quotations from the studies. We explained each step in detail so that our research could serve as an example for other studies. We explained in detail how each study addressed the relationship between AI-assisted writing instruction and critical thinking skills. We also transparently reported our entire systematic review process using the PRISMA 2020 checklist. For the consistency of our study, we recorded the entire research process in detail. We explained step by step which studies we selected and why, how we collected the data, and how we analysed them. Our agreement in our evaluations as two researchers increased from 70 % at the beginning to 85 % after discussions. We also clearly reported our methodological quality assessments using the MMAT. We stored our entire analysis process in digital format (OSF) so that it could be reviewed by others when needed. We followed the PRISMA 2020 guidelines to increase the credibility of our study and transparently reported each step. We conducted a comprehensive literature review and clearly stated which studies we included and excluded and why. We collected data using a standardized form and conducted the entire analysis process in a systematic way.

3. Results

In this section of the study, we present the findings of the systematic review. In Table 2, we present information about the studies included in the analysis.

Analysis of Table 2 reveals that the studies on the subject have intensified in recent years (especially between 2023–2025). This may be related to the widespread use of large language models such as ChatGPT in education. In terms of research designs, it is noteworthy that mixed method studies are predominant. The studies encompass various sample sizes, ranging from in-depth qualitative studies with 10 students to large-scale experimental studies with 1000 students. In terms of educational levels, it is seen that the studies are mostly concentrated at the high school level, there are no studies at the preschool level, and there are relatively fewer studies at the primary school level compared to the high school level. This may be related to the fact that the use of artificial

Table 2
Studies included in the review.

Authors & Year	Purpose	Research Design	K-12 Level & Sample Size	AI Tool	Writing Dimension	Critical Thinking Dimension
Lin et al. (2025)	To examine the effects of generative AI integration into digital multimodal composition on students' motivation, writing processes, and learning outcomes.	Mixed-methods	High School 11 students	Poe (ChatGPT, Claude, Llama)	Content creation, feedback and revision, multilingual support	Self-evaluation, peer assessment, reflective practice
Aladini et al. (2025)	To investigate the effects of AI-based writing tasks on L2 students' writing proficiency, mindfulness development, and grammatical knowledge.	Quasi-experimental	High School 561 students	ChatGPT	Self-directed writing tasks	Self-regulation, metacognitive thinking, reflective thinking
Donnelly et al. (2015)	To analyse student compositions using natural language processing to improve students' explanations of thermodynamic concepts.	Quasi-experimental	Grade 6 346 students	c-rater	Science writing	Explanation making, evidence use, comparison, reflection
Higgs and Stornaiuolo (2024)	To examine young people's views, experiences, and ethical concerns about AI use and AI-assisted writing.	Qualitative	High School 131 students	ChatGPT, MyAI, Grammarly	Creative writing	Critical evaluation of technology's social impacts, ethical thinking
Stornaiuolo et al. (2024)	To investigate how young people interact with generative AI in their daily writing practices.	Critical case study	High School 63 students	Character.ai	Creative writing and interaction	Critical understanding of platform dynamics, ethical evaluation
Crossley et al. (2016)	To evaluate writing quality by incorporating individual differences and linguistic features into automatic essay scoring models.	Quantitative	Grade 10 86 students	Writing Pal (W-Pal)	Persuasive writing	Perspective development, evidence-based argumentation, logical organization
McCarthy et al. (2018)	To examine the effects of metacognitive prompts in an intelligent tutoring system.	2 × 2 factorial experimental	High School 234 students	iSTART	Self-explanation writing	Comprehension monitoring, inference making, knowledge integration
Allen et al. (2019)	To examine relationships between linguistic flexibility, reading comprehension, and automated writing feedback.	Experimental	High School 131 students	W-Pal	Argumentative writing	Argument development, evidence presentation, perspective evaluation
Solis and Idul (2024)	To explore students' lived experiences, motivations, and insights regarding AI writing tools.	Phenomenological	High School 10 students	ChatGPT, Cici, Quillbot, Others	Academic writing	Content creation, editing, analysis, evaluation
Lamb et al. (2021)	To examine students' cognitive mechanisms and critical thinking skills during writing through learning using computational modeling.	Experimental	Grades 4–5 1000 students	ANN	Science Writing Heuristic	Inductive reasoning, problem-solving, deductive reasoning, decision-making
Zhao et al. (2024)	To develop and evaluate an AI curriculum for children in Taiwan.	Pre-post experimental	Grades 3–4 30 students	ChatGPT, Artbreeder, Boomy	AI-assisted writing projects	AI ethics, critical evaluation of safety and societal impacts
Ng et al. (2022)	To investigate the effectiveness of digital story writing in developing AI literacy among primary students.	Mixed method	Primary School 82 students	Various AI tools	Digital story writing	Analysis, evaluation, creation, ethical thinking
Herbold et al. (2023)	To systematically evaluate the quality of ChatGPT-generated writings compared to human writers.	Comparative analysis	Grades 11–13 270 essays	ChatGPT (GPT-3.5, GPT-4)	Argumentative composition	Logical structure, expression and comprehensiveness, language mastery
Wang et al. (2020)	To examine how students implement feedback received through an automated writing evaluation system.	Mixed method	Grades 5–6 143 students	eRevise	Text-based argument writing	Evidence use, analysis, evaluation, reasoning

(*continued on next page*)

Table 2 (*continued*)

Authors & Year	Purpose	Research Design	K-12 Level & Sample Size	AI Tool	Writing Dimension	Critical Thinking Dimension
Utami, Andayani, Winarni, & Sumarwati, 2023	To map perceptions, barriers, and optimization suggestions regarding AI use in academic writing teaching.	Mixed method	Grade 11 58 students	Two AI writing assistants	Process-based writing	Systematic thinking and critical analysis

intelligence tools requires higher level cognitive skills. The artificial intelligence tools used range from general purpose large language models (such as ChatGPT) to systems designed for specific purposes (such as Writing Pal, iSTART). The widespread use of ChatGPT is noteworthy in recent studies. One of the important facts is that AI-assisted writing instruction aims to develop not only technical writing skills but also higher-order thinking skills, ethical awareness, and digital literacy. This holistic approach draws attention to the role of AI in developing 21st century skills.

As a result of the descriptive analysis of the research, we identified five main themes that reveal the relationship between K-12 students' AI-Assisted writing processes and critical literacy skills: (1) Critical Content Generation and Evaluation in AI-Assisted Writing Process, (2) Metacognitive Skills and Self-Regulation in AI-Assisted Writing Process, (3) Ethical Thinking and Evaluation in AI-Assisted Writing Process, (4) Analytical Thinking and Problem Solving in AI-Assisted Writing Process, (5) Motivation and Self-Efficacy in AI-Assisted Writing Process. The findings related to these themes are presented below.

3.1. Critical content generation and evaluation in AI-Assisted writing process

Research under this theme examines how students manage their content generation and evaluation processes in their interactions with AI tools. Lamb et al. (2021) analysed the effect of learning through writing on critical thinking using artificial neural networks in their experimental research with 1000 students and determined that the contribution of critical reasoning in the experimental group (0.98) was significantly higher than the control group (0.04) as a result of 100,000 simulations. Lin et al. (2025) found that 36.4 % of the students (4 people) used AI to generate poems from keywords and synthesized AI outputs with their own ideas instead of directly copying them. Herbold et al. (2023), in a comparative study in which they analysed 270 compositions, found that GPT-4 outperformed students' writing by an average of 1 point on a 7-point Likert scale in terms of logical structure, language complexity, word richness and text connections. In their study with 131 students, Higgs and Stornaiuolo (2024) found that 91 % of students used AI writing tools more than once, but students perceived a lack of emotional depth and authentic voice in AI writing.

3.2. Metacognitive skills and self-regulation in AI-Assisted writing process

Within the scope of this theme, it was examined how the AI-assisted writing process affected students' metacognitive strategies and self-regulation skills. McCarthy et al. (2018), in their study conducted with 234 students, found that the quality of self-explanation in the iSTART system increased significantly compared to the control group ($F(1, 231) = 29.78$, $p < .001$, $\eta^2 p = .11$). In the transfer test, students showed higher performance especially in questions requiring inferential thinking ($t(231) = 2.30$, $p < .05$). In their study with 131 high school students, Allen et al. (2019) found that students showed linguistic flexibility in different writing tasks and made changes in four linguistic dimensions (narrativity, word concreteness, referential cohesion, and deep cohesion) during the revision process. It was also found that students with higher reading skills used less narrative and more analytic language. Wang et al. (2020), in a study of 143 students using the eRevise system, found that 20 % of students ($n = 29$) showed improvement in their ability to use evidence, with greater improvements (ES=0.39), especially among students who scored low on the first draft ($n = 95$). Of the students who received Level 3 feedback, 39 % showed slight improvement and 21 % showed significant improvement.

3.3. Ethical thinking and evaluation in AI-Assisted writing process

Under this theme, ethical awareness and evaluation skills developed by students in the AI-assisted writing process were examined. Higgs and Stornaiuolo (2024) found that students had three main ethical concerns in the use of AI: (1) concern that human creativity is threatened, (2) fear of loss of originality and authenticity of writing, and (3) concern of unfair academic advantage. Stornaiuolo et al. (2024) examined the Character.ai platform and found that students questioned the platform's ambiguous policies on data and copyrights, and had privacy and surveillance concerns when using the platform. Solis and Idul (2024), in their phenomenological research with 10 high school students, identified four main ethical concerns in students' use of AI: (1) becoming over-dependent on AI and weakening critical thinking skills, (2) loss of authenticity, (3) plagiarism concerns, and (4) academic integrity concerns. Ng et al. (2022) conducted a study with 82 elementary school students and found that 53 % of the students performed above average on the AI knowledge test ($M = 3.18$; SD = 0.95), and four students recognized the importance of AI ethics and reflected ethical concerns with data bias in their stories. Zhao et al. (2024), in their research with 30 primary school students, found that students gained awareness about AI ethics, safety, and developed critical thinking skills about the opportunities and challenges that AI technology will create in

their future professions.

3.4. Analytical thinking and problem solving in AI-Assisted writing process

Under this theme, students' analytical thinking and problem solving skills were examined in the AI-assisted writing process. Crossley et al. (2016) evaluated the critical thinking quality of the Writing Pal system on a 6-point scale in a study conducted with 86 students. They found that the system's model, which combines individual differences and text features, explained 52 % of the variance and provided 99 % close matching accuracy. Donnelly et al. (2015) conducted a study with 346 students and found that the scores of students who received correct guidance increased significantly (0.59 points, $p<.01$), but the scores of students who received incorrect guidance (25 out of 150 pairs) decreased. Lamb et al. (2021) found that neural network analysis showed high prediction success (84 % of correct results and 88 % of incorrect results) and that the model significantly improved students' performance on the Cornell Critical Thinking Test. Aladini et al. (2025), in a study conducted with 561 students, found that AI-assisted tasks improved students' critical thinking, self-regulation and problem solving skills.

3.5. Motivation and self-efficacy in AI-Assisted writing process

Under this theme, students' motivation and self-efficacy skills in the AI-assisted writing process were examined. Lin et al. (2025) measured students' motivation for using AI in four dimensions. They found high levels of satisfaction ($M = 4.00$, SD=0.50), attention ($M = 3.78$, SD=0.65), interest ($M = 3.72$, SD=0.55) and confidence ($M = 3.41$, SD=0.61). Utami, Andayani, Winarni, & Sumarwati, 2023 conducted a study with 58 students and found that 86 % of the students found that AI tools helped their writing performance, 65 % improved their writing skills, and 69 % improved their writing quality. In Wang et al.'s (2020) study, 80 % of the students reported that they understood the feedback "mostly" or "completely" and 75 % reported that they used "most" or "all" of the feedback. Zhao et al. (2024) found that 93 % (28/30) of students wanted to continue learning AI in the future and that the use of AI increased interest in technology (mean=5.83/7.0). As another positive contribution, Aladini et al. (2025) found in their large-sample study ($n = 561$) that AI-assisted tasks increased student autonomy and self-efficacy, and increased motivation and engagement levels in students.

4. Discussion

In this study, we aimed to examine the relationship between AI-assisted writing instruction and critical thinking skills of K-12 students with a systematic approach. As a result of the research, we determined that AI-assisted writing processes shape students' content generation and evaluation competencies, higher-order thinking and self-regulation skills, ethical awareness and evaluation approaches, analytical thinking and problem-solving capacities, and motivation and self-efficacy perceptions towards learning.

Our findings regarding students' content production and evaluation approaches in AI-assisted writing processes reveal that they create original content by integrating the opportunities offered by technology with their own thought processes. Faiz et al., (2023); Fathi and Rahimi (2024) also supported this finding and concluded that the use of artificial intelligence helps in the development of writing skills holistically with elements such as task completion, coherence, conjunction, vocabulary, grammatical variety and accuracy. The use of artificial intelligence in writing instruction will provide significant improvements in grammatical competence and its contribution to the quality of writing is also supported by different studies (Athanassopoulos et al., 2023). Another finding of our research is that artificial intelligence tools support students' critical thinking skills. Based on our findings, students' questioning artificial intelligence outputs and synthesizing them with their own ideas contribute to the development of evaluation and creation skills, which are the basic components of critical thinking. In relation to this finding, Al Shloul et al. (2024) emphasized that it is very important for students to interact critically with the content produced by AI in order to maintain the integrity and robustness of educational outcomes.

In terms of metacognitive skills and self-regulation, The analysis indicates that AI-assisted writing processes improve students' ability to manage and evaluate their own learning processes. This development deepens the critical thinking strategies that students use in the writing process. Linguistic flexibility and the use of evidence in editing processes were identified as important variables of critical thinking. Our finding regarding the contribution of artificial intelligence to evaluation skills is supported by many studies in the literature (Mintz, et al., 2023; Owan, et al., 2023; Pattada, 2024 and so forth). However, our study addresses a different point from the studies in the literature in terms of artificial intelligence, writing skills, critical literacy and evaluation. Supporting our finding, Al Shloul et al. (2024) addressed this multiple relationship within the limitation of scientific writing. According to them, AI-assisted writing is a powerful aid in drafting, revising and generating ideas, as well as a critical evaluation of the appropriateness of the content. Similar to our finding regarding the contribution of AI to students' metacognitive processes, Latifi & Noroozi (2021) demonstrated that during structured feedback processes, students not only learn the main components of effective writing in their field but also actively practice these components by providing and receiving quality feedback, thereby developing their critical evaluation skills.

Our findings regarding the ethical thinking and evaluation dimension show that students' awareness and critical approaches to the use of artificial intelligence in the writing process have improved. Thinking about the effects of artificial intelligence technology and developing ethical concerns strengthen the ethical dimension of critical thinking in students. This situation reveals that AI-assisted writing processes enable students to develop their critical thinking skills within an ethical framework. Corcuera (2024), which partially overlaps with our findings, concluded that artificial intelligence has a positive effect on writing skills. However, unlike our findings, Corcuera (2024) drew attention to the fact that artificial intelligence is overly addictive, makes students comfortable, and

plagiarism issues. Like Corcuera (2024), many studies in the literature, while stating that artificial intelligence contributes to writing skills, also draw attention to ethical concerns (AlSamhori & Alnaimat, 2024; Hosseini et al., 2023; Khup & Bantugan, 2025 and so forth). In this respect, as we stated in the findings of our study, it is necessary to strengthen students' critical literacy skills in order to minimize ethical concerns.

In terms of analytical thinking and problem solving skills, AI-assisted writing processes contribute to students' systematic thinking and reasoning skills. In this way, we concluded that students use critical thinking strategies more effectively in complex writing tasks. The findings demonstrate that artificial intelligence tools, when supported by appropriate pedagogical approaches accelerate the development of students' analytical thinking and critical evaluation skills. The findings of our study are supported by other studies in the literature. For example, Susanto et al. (2023), in a study conducted with secondary school students in Hong Kong, states that AI technology contributes to problem solving skills in the writing process and improves students' ability to communicate complex ideas effectively. Similarly, Khup and Bantugan (2025) in their study for junior high school students concluded that AI contributes to the generation of original ideas, grammar, vocabulary and text structure and improves students' creativity. However, in the same study, Khup and Bantugan (2025) reached a different conclusion from the findings of our study by stating that artificial intelligence also leads to problems such as overdependence, plagiarism and limited critical thinking.

Our findings regarding motivation and self-efficacy show that AI-assisted writing processes increase students' desire to learn and their self-confidence. Our findings show that this supportive situation enables students to use their critical thinking skills more willingly and effectively. In addition, this supportive learning environment provided by artificial intelligence tools contributes to students becoming more autonomous and self-confident individuals in developing their critical thinking skills. Szmyd and Mitera (2024), making similar points with our findings, state that artificial intelligence allows the pace and difficulty level to be adjusted according to the individual needs of each student, providing a higher level of engagement and motivation. Another study by Khup and Bantugan (2025) concluded that AI improves students' writing skills, increases their motivation and encourages engagement by providing personalized feedback. Recent advancements in generative artificial intelligence (Gen-AI) have significantly transformed the manner in which writing instruction can be supported by human intervention. Banihashem et al. (2024) conducted a study comparing peer-generated feedback with ChatGPT-generated feedback in university students' essay writing. Their findings revealed an interesting complementary relationship: while ChatGPT excelled in providing descriptive feedback about how essays were structured, human peers were more effective at identifying specific problems within the essays. Although their study focused on higher education rather than on K-12 students, their findings align with our observations regarding the importance of combining human guidance with AI tools. Similar to our finding that students benefit from critically evaluating AI outputs, Banihashem et al.'s (2024) study suggests that optimal educational outcomes occur when AI tools are integrated with human intervention rather than as standalone solutions. This integration helps develop both writing skills and higher-order thinking abilities, supporting our conclusion that the future of writing instruction lies in developing effective human-AI collaboration models across all educational levels. In addition, although our study did not specifically analyze demographic variables in AI-assisted writing processes, Noroozi et al.'s (2023) research showed that factors such as gender can significantly affect argumentative essay writing and peer review performance in online learning environments. This highlights the importance of considering demographic variables in future research on AI-enhanced writing instruction to gain a more comprehensive understanding of how different groups of students interact and benefit from these technologies.

5. Limitations

This systematic review has some methodological and content limitations. Methodologically, our study was limited to English articles in 4 databases. The exclusion of publications in other languages and other types of academic publications (book chapters, conference proceedings and so forth) may have limited the scope and universality of our study. A significant number of the studies we examined measured students' critical thinking skills in AI-assisted writing processes with different assessment tools. For this reason, it becomes difficult to compare the results of the studies. In addition, the analysed studies were limited to K-12 level, but there was not a balanced distribution among different age groups in the studies analysed for this study. No studies were found with pre-school students, and the number of studies with primary school students is quite limited compared to middle and high school students. Another limitation is that most of the studies evaluated students' critical thinking skills in AI-assisted writing processes with quantitative data. The scarcity of qualitative studies that examine students' thinking processes in depth has caused some dimensions of the subject to be insufficiently illuminated. There were also significant differences in sample sizes between the studies reviewed. For example, Solis and Idul (2024) provided qualitative data with only 10 students, whereas Lamb et al. (2021) provided larger-scale analyses with 1000 students. This makes it difficult to make comparisons in terms of generalizability and depth of the findings. Although studies with small samples can provide in-depth insights, the lack of statistical power can make it difficult to generalize the results. In addition, the variety of artificial intelligence tools used in the studies (ChatGPT, Writing Pal, iSTART, etc.) limits the consistent comparison of results. The different features and backgrounds of different tools may affect students' writing and critical thinking skills in various ways. The lack of detailed background information on the characteristics of the participants in the reviewed studies is also an important limitation. While these studies generally provide basic demographic information such as age, grade level, and sample size of the participants, they do not include other important characteristics that may affect the way students interact with artificial intelligence, critical thinking, and argumentation. Factors such as students' prior technological experiences, digital literacy levels, academic achievement levels, socioeconomic status, and language proficiency have not been reported in most studies. This deficiency makes it difficult to compare the experiences and performance of students with different characteristics in AI-assisted writing processes. Although some studies provided information on the cultural and educational contexts of participants, the depth and breadth of this information was limited.

5.1. Recommendations

Based on the findings of our study, recommendations are presented under three headings for teachers, researchers and policy makers. Teachers can organize writing activities to develop students' thinking skills. In these activities, students should be able to question and evaluate the texts they encounter and combine them with their own ideas. Teachers can use methods to help students plan and evaluate their writing processes. In classroom discussions, students can be made aware of writing in accordance with ethical values. School administrators and education policy makers can organize training programs to support the professional development of teachers in the process of AI-assisted writing. In these programs, teachers can be shown methods and techniques on how to improve students' critical thinking skills. In addition, necessary arrangements can be made to eliminate the technological infrastructure differences between schools and to ensure that all students have equal access to artificial intelligence technology. Additionally, education policymakers should consider developing practical implementation frameworks to address the equity challenges in AI tool deployment across different institutional settings. These frameworks should account for the varying levels of AI literacy among teachers, students, and institutions, ensuring that schools with limited technological resources or expertise are not left behind. We recommend a three-tiered approach: (1) assessment of institutional AI readiness, including infrastructure, teacher competencies, and student familiarity with AI tools; (2) differentiated professional development pathways for educators based on their current AI literacy levels; and (3) context-sensitive implementation guidelines that consider the cultural, socioeconomic, and geographical factors affecting AI integration. Such frameworks would help to identify potential barriers to equitable AI implementation and provide strategies to overcome them, ensuring that all students, regardless of their institutional context, can benefit from AI-assisted writing instruction while developing critical thinking skills. Researchers can develop common measurement tools to assess students' critical thinking skills in the process of AI-assisted writing. Studies examining preschool and primary school students' AI-assisted writing processes and the development of their critical thinking skills can be increased. Research can be conducted to provide a deeper understanding of how students think and make decisions while writing. There is also a need for research that tracks how students develop in the long term. It is also important to report participant characteristics in a more comprehensive and detailed manner in future studies. Factors that may affect students' interactions with AI and their critical thinking processes (such as digital literacy levels, technology experiences, cognitive development levels, socioeconomic status, language proficiency, and academic achievement levels) should be systematically examined and reported. Researchers can design studies comparing groups of students with different characteristics to examine the impact of demographic variables on AI-supported writing processes. Research that compares the AI-supported writing experiences of students in different socioeconomic and cultural contexts will be valuable for understanding the impact of digital inequalities and developing solutions. Another important research area is the development of standardized assessment tools to examine the impact of students' AI literacy levels on the development of critical thinking and argumentation skills.

6. Conclusion

This study, in which we examined the relationship between AI-assisted writing processes and critical thinking skills of K-12 students, reveals important results. Students utilize AI tools not merely for copy-pasting, but as supportive elements for developing their thoughts. In this respect, it is noteworthy that students create new texts by questioning these outputs and blending them with their own ideas instead of accepting artificial intelligence outputs as they are. Our research has shown that students manage their own writing processes more consciously. Students were more careful in revising their texts and were more attentive to evidence-based writing. In this way, students' thinking skills as well as their writing skills improve. Students' ethical awareness of technology use is also one of the important results of our study. Students are sensitive about originality, academic honesty and creativity. Another noteworthy result is that students think more systematically in complex writing tasks. Properly guided students can solve the problems they encounter in the writing process more effectively and produce more qualified texts. In addition, we found that students' self-confidence increased during the AI-assisted writing process. AI-assisted writing experiences reduce students' writing anxiety and make them more enthusiastic about writing.

The originality and contribution of this systematic review to the literature lies in its holistic approach to the complex relationship between AI-assisted writing and critical literacy at the K-12 level. While studies on AI and writing skills generally focus on the university level, our study is one of the first systematic analyses to comprehensively examine school-age children's interactions with these technologies. The five main themes we identified (critical content production and evaluation, metacognitive skills and self-regulation, ethical thinking and evaluation, analytical thinking and problem solving, motivation, and self-efficacy) provide a unique framework that reveals how AI is integrated with the development of 21st century skills. One of the most important contributions of our study is the empirical evidence that AI technologies can serve as active agents for K-12 students to develop critical thinking skills, rather than a passive auxiliary tool. Our findings offer an innovative perspective to educators and policymakers by emphasizing the importance of critical pedagogy in integrating AI tools into the classroom environment. This understanding is important for reshaping educational paradigms in the digital age and educating students as producers and critical consumers. Our research reveals that when supported with the right pedagogical approaches, AI-supported learning environments not only improve K-12 students' writing skills but also strengthen their higher-order cognitive skills, such as critical thinking, ethical reasoning, and self-regulation. These findings emphasize the importance of effectively utilizing the potential of these technologies to develop critical literacy skills, rather than banning artificial intelligence technologies in educational systems in the digital transformation process.

Funding

This study is not funded.

Ethics statement

Ethics approval was not required.

CRediT authorship contribution statement

Tülin Arseven: Writing – review & editing, Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Mazhar Bal:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation.

Declaration of competing interest

The author declares no potential conflicts of interest.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.tsc.2025.101850.

Data availability

The supplementary materials including evaluation process table is available as additional files at system.

References

Aladini, A., Ismail, S. M., Khasawneh, M. A. S., & Shakibaei, G. (2025). Self-directed writing development across computer/AI-based tasks: Unraveling the traces on L2 writing outcomes, growth mindfulness, and grammatical knowledge. *Computers in Human Behavior Reports, 17*, 1–14. https://doi.org/10.1016/j.chbr.2024.100566

Albadarin, Y., Saqr, M., Pope, N., & Tukiainen, M. (2024). A systematic literature review of empirical research on ChatGPT in education. *Discover Education, 3*(1), 60. https://doi.org/10.1007/s44217-024-00138-2

Allen, L. K., Likens, A. D., & McNamara, D. S. (2019). Writing flexibility in argumentative essays: A multidimensional analysis. *Reading and Writing, 32*, 1607–1634. https://doi.org/10.1007/s11145-018-9921-y

AlSamhori, A. F., & Alnaimat, F. (2024). Artificial intelligence in writing and research: Ethical implications and best practices. *Central Asian Journal of Medical Hypotheses and Ethics, 5*(4), 259–268. https://doi.org/10.47316/cajmhe.2024.5.4.02

Al Shloul, T., Mazhar, T., Qamar, A., Iqbal, M., Yasin Ghadi, Y., Shahzad, T., Mallek, F., & Hamam, H. (2024). Role of activity-based learning and ChatGPT on students' performance in education. *Computers and Education: Artificial Intelligence, 6*, 1–18. https://doi.org/10.1016/j.caeai.2024.100219

Athanassopoulos, S., Manoli, P., Gouvi, M., Lavidas, K., & Komis, V. (2023). The use of ChatGPT as a learning tool to improve foreign language writing in a multilingual and multicultural classroom. *Advances in Mobile Learning Educational Research, 3*(2), 818–824. https://doi.org/10.25082/AMLER.2023.02.009

Bal, M., & Öztürk, E. (2025). The potential of deep learning in improving K-12 students' writing skills: A systematic review. *British Educational Research Journal*. https://doi.org/10.1002/berj.4120

Banihashem, S. K., Kerman, N. T., Noroozi, O., Moon, J., & Drachsler, H. (2024). Feedback sources in essay writing: Peer-generated or AI-generated feedback? *International Journal of Educational Technology in Higher Education, 21*(23), 1–15. https://doi.org/10.1186/s41239-024-00455-4

Barbot, B., Tan, M., Randi, J., Santa-Donato, G., & Grigorenko, E. L. (2012). Essential skills for creative writing: Integrating multiple domain-specific perspectives. *Thinking Skills and Creativity, 7*(3), 209–223. https://doi.org/10.1016/j.tsc.2012.04.006

Berninger, V. W. (2019). *Reading and writing acquisition: A developmental neuropsychological perspective*. Routledge. https://doi.org/10.4324/9780429303821

... & Blaizot, A., Veettil, S. K., Saidoung, P., Moreno-Garcia, C. F., Wiratunga, N., Aceves-Martins, M., & Chaiyakunapruk, N. (2022). Using artificial intelligence methods for systematic review in health sciences: A systematic review. *Research Synthesis Methods, 13*(3), 353–362. https://doi.org/10.1002/jrsm.1553.

Chen, T. J. (2023). ChatGPT and other artificial intelligence applications speed up scientific writing. *Journal of the Chinese Medical Association, 86*(4), 351–353. https://doi.org/10.1097/JCMA.0000000000000900

Chen, Y. T., Li, M., Cukurova, M., & Jong, M. S. Y. (2024). Incorporation of peer-feedback into the pedagogical use of spherical video-based virtual reality in writing education. *British Journal of Educational Technology, 55*(2), 519–540. https://doi.org/10.1111/bjet.13376

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Corcuera, L. (2024). Uncovering QuillBot: Filipino senior high school students' experiences and factors influencing its use in enhancing language writing skills. *International Journal of Scholars in Education, 7*(2), 67–81. https://doi.org/10.52134/ueader.1497368

Crossley, S. A., Allen, L. K., Snow, E. L., & McNamara, D. S. (2016). Incorporating learning characteristics into automatic essay scoring models: What individual differences and linguistic features tell us about writing quality. *Journal of Educational Data Mining, 8*(2), 1–19. https://eric.ed.gov/?id=EJ1125807.

Demirel, Ö. (2003). *Türkçe Öğretimi*. Pegem A Yayıncılık.

Donnelly, D. F., Vitale, J. M., & Linn, M. C. (2015). Automated guidance for thermodynamics essays: Critiquing versus revisiting. *Journal of Science Education and Technology, 24*, 861–874. https://doi.org/10.1007/s10956-015-9569-1

Faiz, R., Bilal, H. A., Asghar, I., & Safdar, A. (2023). Optimizing ChatGPT as a writing aid for EFL learners: Balancing assistance and skill development in writing proficiency. *Linguistic Forum-A Journal of Linguistics, 5*(3), 24–37. https://doi.org/10.5281/zenodo.14756539

Fathi, J., & Rahimi, M. (2024). Utilising artificial intelligence-enhanced writing mediation to develop academic writing skills in EFL learners: A qualitative study. *Computer Assisted Language Learning*, 1–40. https://doi.org/10.1080/09588221.2024.2374772

Feng Teng, M. (2024). A systematic review of ChatGPT for English as a foreign language writing: Opportunities, challenges, and recommendations. *International Journal of TESOL Studies, 6*(3), 36–57. https://doi.org/10.58304/ijts.20240304
Gonsalves, C. (2024). Generative AI's impact on critical thinking: Revisiting Bloom's Taxonomy. *Journal of Marketing Education*. https://doi.org/10.1177/02734753241305980
Janse van Rensburg, J. (2024). Artificial human thinking: ChatGPT's capacity to be a model for critical thinking when prompted with problem-based writing activities. *Discover Education, 3*(1), 42. https://doi.org/10.1007/s44217-024-00113-x
Handayani, R. (2016). Students' critical thinking skills in a classroom debate. *Language and Language Teaching Journal, 19*(2), 132–140. https://doi.org/10.24071/llt.2016.190208
Herbold, S., Hautli-Janisz, A., Heuer, U., Kikteva, Z., & Trautsch, A. (2023). A large-scale comparison of human-written versus ChatGPT-generated essays. *Scientific Reports, 13*(1), 1–11. https://doi.org/10.1038/s41598-023-45644-9
Higgs, J. M., & Stornaiuolo, A. (2024). Being human in the age of generative AI: Young people's ethical concerns about writing and living with machines. *Reading Research Quarterly, 59*(4), 632–650. https://doi.org/10.1002/rrq.552
... & Hong, Q. N., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., & Pluye, P. (2018). The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for Information, 34*(4), 285–291. https://doi.org/10.3233/EFI-180221.
Hosseini, M., Resnik, D. B., & Holmes, K. (2023). The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts. *Research Ethics, 19*(4), 449–465. https://doi.org/10.1177/17470161231180449
Kaya, B., & Ateş, S. (2016). Üstbilişsel beceri odaklı yazma süreçlerinin dördüncü sınıf öğrencilerinin hikâye yazma becerisine etkisi. *Eğitim ve Bilim, 41*(187), 137–164. https://doi.org/10.15390/EB.2016.6752
Kellogg, R. T. (2008). Training writing skills: A cognitive developmental perspective. *Journal of Writing Research, 1*(1), 1–26. https://doi.org/10.17239/jowr-2008.01.01.1
Kellogg, R. T., & Whiteford, A. P. (2009). Training advanced writing skills: The case for deliberate practice. *Educational Psychologist, 44*(4), 250–266. https://doi.org/10.1080/00461520903213600
Khup, V. K., & Bantugan, B. (2025). Exploring the ımpact and ethical ımplications of ıntegrating AI-powered writing tools in junior high school English ınstruction: Enhancing creativity, proficiency, and academic outcomes. *International Journal of Research and Innovation in Social Science, 9*(3), 361–378. https://doi.org/10.47772/IJRISS.2025.903SEDU0022
Kim, J., Yu, S., Detrick, R., & Li, N. (2025). Exploring students' perspectives on Generative AI-assisted academic writing. *Education and Information Technologies, 30*(1), 1265–1300. https://doi.org/10.1007/s10639-024-12878-7
Kokina, J., Blanchette, S., Davenport, T. H., & Pachamanova, D. (2025). Challenges and opportunities for artificial intelligence in auditing: Evidence from the field. *International Journal of Accounting Information Systems, 56*, Article 1-22. https://doi.org/10.1016/j.accinf.2025.100734
Lam, H. Y., Ho, G. T. S., Mo, D. Y., & Tang, V. (2021). Enhancing data-driven elderly appointment services in domestic care communities under COVID-19. *Industrial Management & Data Systems, 121*(7), 1552–1576. https://doi.org/10.1108/IMDS-07-2020-0392
Lamb, R., Hand, B., & Kavner, A. (2021). Computational modeling of the effects of the science writing heuristic on student critical thinking in science using machine learning. *Journal of Science Education and Technology, 30*(2), 283–297. https://doi.org/10.1007/s10956-020-09871-3
Larkin, S. (2009). Socially mediated metacognition and learning to write. *Thinking Skills and Creativity, 4*(3), 149–159. https://doi.org/10.1016/j.tsc.2009.09.003
Latifi, S., Noroozi, O., & Banihashem, S. K. (2023). Scafolding argumentative essay writing: An online scripted peer feedback and peer feedforward module. In *International Conference on Education and New Developments (END) (2023)* (pp. 62–66). https://doi.org/10.36315/2023v2end013
Latifi, S., & Noroozi, O. (2021). Supporting argumentative essay writing through an online supported peer-review script. *Innovations in Education and Teaching International, 58*(5), 501–511. https://doi.org/10.1080/14703297.2021.1961097
Li, X., & Zhang, M. (2015). Rising to the challenge: What practicing teachers learned from a process-based writing project in a graduate capstone seminar. *Sage Open, 5*(2). https://doi.org/10.1177/2158244015590163
Lin, C. H., Zhou, K., Li, L., & Sun, L. (2025). Integrating generative AI into digital multimodal composition: A study of multicultural second-language classrooms. *Computers and Composition, 75*, 1–24. https://doi.org/10.1016/j.compcom.2024.102895
Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education, 21*(2), Article 1-13. https://doi.org/10.1016/j.ijme.2023.100790
Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic in-quiry*. Sage.
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. https://doi.org/10.1007/s10639-023-11697-6
Mallahi, O. (2024). Exploring the status of argumentative essay writing strategies and problems of Iranian EFL learners. *Asian-Pacific Journal of Second and Foreign Language Education, 9*(19), 1–26. https://doi.org/10.1186/s40862-023-00241-1
McCarthy, K. S., Likens, A. D., Johnson, A. M., Guerrero, T. A., & McNamara, D. S. (2018). Metacognitive overload!: Positive and negative effects of metacognitive prompts in an intelligent tutoring system. *International Journal of Artificial Intelligence in Education, 28*, 420–438. https://doi.org/10.1007/s40593-018-0164-5
Miles, M. B. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks.
Mintz, J., Holmes, W., Liu, L., & Perez-Ortiz, M. (2023). Artificial Intelligence and K-12 education: Possibilities, pedagogies and risks. *Computers in the Schools, 40*(4), 325–333. https://doi.org/10.1080/07380569.2023.2279870
Mirón-Mérida, V. A., & García-García, R. M. (2024). Developing written communication skills in engineers in Spanish: Is ChatGPT a tool or a hindrance? *Frontiers Education, 9*. https://doi.org/10.3389/feduc.2024.1416152
Ng, D. T. K., Luo, W., Chan, H. M. Y., & Chu, S. K. W. (2022). Using digital story writing as a pedagogy to develop AI literacy among primary students. *Computers and Education: Artificial Intelligence, 3*, 1–14. https://doi.org/10.1016/j.caeai.2022.100054
Ngajie, B. N., Li, Y., & Su, C. Y. (2018). Analyzing critical thinking elements in the argument structure of Non-English-major Chinese undergraduate students' writing. In *2018 Seventh International Conference of Educational Innovation through Technology (EITT)* (pp. 77–86). IEEE.
Noroozi, O., Banihashem, S. K., Taghizadeh Kerman, N., Parvaneh Akhteh Khaneh, M., Babaee, M., Ashrafi, H., & Biemans, H. J. (2023). Gender differences in students' argumentative essay writing, peer review performance and uptake in online learning environments. *Interactive Learning Environments, 31*(10), 6302–6316. https://doi.org/10.1080/10494820.2022.2034887
Owan, V. J., Abang, K. B., Idika, D. O., Etta, E. O., & Bassey, B. A. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *Eurasia Journal of Mathematics, Science and Technology Education, 19*(8), em2307. https://doi.org/10.29333/ejmste/13428
Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ (Clinical Research Ed.), 372, 1-9*. https://doi.org/10.1136/bmj.n71
Parra, G., L, Calero, S., & X. (2019). Automated writing evaluation tools in the improvement of the writing skill. *International Journal of Instruction, 12*(2), 209–226. https://doi.org/10.29333/iji.2019.12214a
Pattada, S. I. (2024). Exploring the effects of artificial intelligence tools on the essay-writing skills of higher secondary students. *International Journal of Smart Technology and Learning, 4*(1), 76–94. https://doi.org/10.1504/IJSMARTTL.2024.142178
Safitri, A., & Adani, S. (2024). Students' ability to develop critical thinking skills in English essay writing. *Edukatif: Jurnal Ilmu Pendidikan, 6*(4), 3440–3450. https://doi.org/10.31004/edukatif.v6i4.7335
Sani, S. M., Kurniawati, N., & Nurwanti, D. I. (2018). The use of reflective writing to improve students' writing and critical thinking skills. *Journal of Education and Practice, 6*(13), 331–335. https://doi.org/10.5220/0007166903310335
Sharma, D. (2024). Critical thinking and problem-solving in the age of ChatGPT: Experiential-bibliotherapy-blogging project. *Business and Professional Communication Quarterly, 87*(4), 630–653. https://doi.org/10.1177/23294906241254780

Shen, X., & Teng, M. F. (2024). Three-wave cross-lagged model on the correlations between critical thinking skills, self-directed learning competency and AI-assisted writing. *Thinking Skills and Creativity, 52*, Article 1-13. https://doi.org/10.1016/j.tsc.2024.101524

Solis, J., & Idul, M. (2024). Stories in the age of automation: A phenomenological study on the role of artificial ıntelligence on students' writing. *Psychology and Education: A Multidisciplinary Journal, 23*(1), 92–107. https://doi.org/10.5281/zenodo.13222155

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology, 14*, Article 1-14. https://doi.org/10.3389/fpsyg.2023.1260843

Stornaiuolo, A., Higgs, J., Jawale, O., & Martin, R. M. (2024). Digital writing with AI platforms: The role of fun with/in generative AI. *English Teaching: Practice & Critique, 23*(1), 83–103. https://doi.org/10.1108/ETPC-08-2023-0103

Susanto, H., Woo, D.J., .& Guo, K. (2023). The role of AI in human-aı creative writing for Hong Kong secondary students. https://doi.org/10.48550/arXiv.2304.11276.

Szmyd, K., & Mitera, E. (2024). The ımpact of artificial ıntelligence on the development of critical thinking skills in students. *European Research Studies Journal, 27*(2), 1022–1039. https://doi.org/10.35808/ersj/3876

Tang, Y. M., Zhao, D., Chen, T., & Fu, X. (2025). A systematic review of abnormal behaviour detection and analysis in driving simulators. *Transportation Research Part F: Traffic Psychology and Behaviour, 109*, 897–920. https://doi.org/10.1016/j.trf.2025.01.002

Tasic, B., & Hannah, J. (2017). Supporting written communication skills in statistics courses. In *EDULEARN17 Proceedings* (pp. 9695–9703). IATED.

Teng, M. F., & Yue, M. (2023). Metacognitive writing strategies, critical thinking skills, and academic writing performance: A structural equation modeling approach. *Metacognition and Learning, 18*(1), 237–260. https://doi.org/10.1007/s11409-022-09328-5

Utami, S. P. T., Andayani, A., Winarni, R., & Sumarwati, S. (2023). Utilization of artificial intelligence technology in an academic writing class: How do Indonesian students perceive? *Contemporary Educational Technology, 15*(4), 1–13. https://doi.org/10.30935/cedtech/13419

... & Wang, E. L., Matsumura, L. C., Correnti, R., Litman, D., Zhang, H., Howe, E., & Quintana, R. (2020). eRevis (ing): Students' revision of text evidence use in an automated writing evaluation system. *Assessing Writing, 44*, 1–19. https://doi.org/10.1016/j.asw.2020.100449.

Wu, T. T., Silitonga, L. M., & Murti, A. T. (2024). Enhancing English writing and higher-order thinking skills through computational thinking. *Computers & Education, 213*, 1–19. https://doi.org/10.1016/j.compedu.2024.105012

Yilmaz-Na, E., & Sönmez, E. (2023). Having qualified arguments: Promoting Pre-service teachers' critical thinking through deliberate computer-assisted argument mapping practices. *Thinking Skills and Creativity, 47*, 1–22. https://doi.org/10.1016/j.tsc.2022.101216

Zhang, C., Hur, J., Diamond, K. E., & Powell, D. (2015). Classroom writing environments and children's early writing skills: An observational study in Head Start classrooms. *Early Childhood Education Journal, 43*(4), 307–315. https://doi.org/10.1007/s10643-014-0655-4

Zhao, H. G., Li, X. Z., & Kang, X. (2024). Development of an artificial intelligence curriculum design for children in Taiwan and its impact on learning outcomes. *Humanities and Social Sciences Communications, 11*(1), 1–17. https://doi.org/10.1057/s41599-024-03839-z