



An AI-assisted critical thinking intervention to enhance undergraduate EFL learners' writing proficiency

Xue Yin ^a, Kun Dou ^{b,*}

^a School of Foreign Languages, South China University of Technology, Guangzhou, China

^b Department of Language and Literacy Education, Faculty of Education, University of Malaya, Kuala Lumpur, Malaysia

ARTICLE INFO

Keywords:

Critical thinking
EFL writing
Technology acceptance
ChatGPT
AI-enhanced education

ABSTRACT

This study investigates the effects of an AI-assisted critical thinking (CT)-oriented writing intervention supported with ChatGPT involving 250 undergraduate EFL learners from three public universities. The research measured changes in students' writing proficiency and their acceptance of ChatGPT in CT-oriented writing practices. A pretest-posttest design was adopted, with data collected through pre-post-writing tests ($n = 250$), an acceptance questionnaire ($n = 250$), and further semi-structured interviews ($n = 25$). The writing tests were evaluated across nine dimensions: clarity, accuracy, precision, relevance, depth, breadth, logic, significance, and fairness. The acceptance questionnaire, grounded in the Extended Unified Theory of Acceptance and Use of Technology (UTAUT2) model, assessed six constructs, i.e., performance expectancy, effort expectancy, hedonic motivation, habit, trust, and behavioral intention. The results showed that 1) The intervention significantly enhanced students' CT reflected in writing; 2) Positive shifts were observed in four scales of ChatGPT acceptance in a CT-oriented writing context, i.e., performance expectancy, effort expectancy, trust, and behavior intention, though only marginal increase in hedonic motivation and habit. Meanwhile, the study identified the primary aspects of ChatGPT influencing EFL learners' CT-oriented writing. These results demonstrate the potential of ChatGPT as a scaffolding tool for CT-oriented writing training in EFL contexts. Despite limitations such as the single-group experimental design and 8-week experimental period, the findings offer valuable insights for further research and development in AI-driven tools in EFL writing pedagogy. Future research could adopt a multi-group design and longitudinal assessments to further validate the intervention's sustained effects.

1. Introduction

English as a foreign language (EFL) learners often face significant challenges in developing English writing proficiency, particularly in mastering syntactic, lexical, and rhetorical conventions (Algarraedy & Mahyoob, 2023; Bai et al., 2021; Yin, 2024). Among the essential language competencies that EFL learners must acquire, writing is frequently regarded as one of the most demanding and intricate skills to master (Esmaeil Nejad et al., 2022; Irgin & Bilki, 2024; Ruan, 2005; Shida et al., 2018; Yih Mah et al., 2021; Zuckerman et al., 2023; X. Yin, Guo, & Li, 2024; X. Yin, Saad, & Halim, 2024).

The limited development of critical thinking (CT) skills among EFL learners often impedes their ability to produce high-quality written work (Bean & Melzer, 2021; Dong, 2017a; Liu & Yao, 2019; X. Yin et al., 2024). Numerous scholars (Akbari et al., 2018; Esmaeil Nejad et al., 2022; Halpern, 2014; Piaget, 1970; Sharadgah et al., 2019; Susilo et al.,

2021; Vygotsky, 1986; Wade, 1995; Yin & Saad, 2023; Zou et al., 2022) have documented a strong link between CT and the practice of writing across multiple studies and academic discussions. Empirical studies demonstrate that written texts operate as visible representations of CT capacities (Condon & Kelly-Riley, 2004; Fraser et al., 2022; Halpern, 2014; Kaviani & Heidar, 2020; Lin, 2019), and targeted CT training interventions have been shown to significantly enhance the logical organization and evidentiary robustness of EFL learners' written discourse (Akbari et al., 2018; Barnawi, 2011; Healey, 2024; Murtadho, 2021; Pei et al., 2017).

Recent studies indicate that the scaffolding environment has evolved from conventional classroom-based contexts to digital platforms, with technology being employed as a pivotal scaffolding mechanism (Banihashem et al., 2024; Lu & Xie, 2022; Yin et al., 2023). According to a systematic review conducted by Yin et al. (2023), which analyzed articles on CT-oriented EFL writing instruction indexed in Web of

* Corresponding author.

E-mail addresses: Sixpence_1@163.com (X. Yin), anastasiadou0131@hotmail.com (K. Dou).

Science and Scopus over the past decade, approximately 30 % of the authors incorporated online methodologies to foster interaction between educators and student peers.

The limited number of studies utilizing technology to facilitate writing exhibit several notable limitations. First, most CT-oriented research utilizing technology-optimized support in EFL writing classrooms focused on traditional web 2.0 technology, which featured online interaction and user involvement (Yin & Saad, 2023; Yin et al., 2023; Yin, 2024; X. Yin et al., 2024). For instance, engaging in immersed learning through playing online games (Stapa & Ibaharim, 2020) or self-regulated strategy use in writing through e-learning (Bai et al., 2021). Those studies predominantly utilized early iterations of chatbots with limited functionalities, such as delivering pre-programmed responses instead of providing tailored and adaptive feedback (Goda et al., 2014; Guo & Wan, 2023; Irgin & Bilki, 2024). Others primarily focused on the evaluation of written outputs (Zuckerman et al., 2023), including post-writing reflection (Su, Lin, & Lai, 2023) or the detection of writing errors (Algarady & Mahyoob, 2023). A limited number of studies have explored the application of advanced AI-driven language models to support thinking-oriented writing in the context of EFL learning.

Second, few scholars have explored the application of GPT-based tools to support writing, with existing studies primarily examining superficial outputs or feedback generated by ChatGPT. These investigations reveal that the tool lacks the capacity to pose guiding questions or scaffold students' CT, functioning merely as a "knowledgeable peer who is unfortunately not able to initiate conversations and guide thinking development" (Su et al., 2023, p. 8). Despite the advanced natural language processing capabilities of OpenAI's GPT-series models, their effectiveness is constrained by limited digital literacy and dependence on user input (Godwin-Jones, 2022). Consequently, explicit instruction and guidance are essential to equip students with the necessary skills to formulate meaningful queries and leverage ChatGPT effectively for reflective thinking during the writing process (Su et al., 2023). As evidenced by the existing literature, there is a notable gap in research examining how AI-driven language models can function as effective collaborative peers to support CT-oriented writing in EFL learning. To address this gap, the present study proposes a CT-oriented framework designed to facilitate EFL learners' collaboration with ChatGPT during the writing process. Specifically, AI-assisted Brainstorming Worksheets and AI-assisted Peer-review Checklists were provided to scaffold the writing process. These two worksheets aim to mitigate the limitations by guiding both the students and this knowledge peer to a structured, thinking-oriented training process during writing activities.

Besides the effectiveness of the CT-oriented writing framework supported by ChatGPT, learners' acceptance of using AI-assisted CT-oriented writing strategies was also investigated in this research. Such information is important since acceptance is a prerequisite for participation (Boudouaia et al., 2024; Cheung & Vogel, 2013; Salam, 2024; Sun & Zou, 2022). The intention of individuals to incorporate and utilize AI technology plays a pivotal role in enhancing the instructional quality and learning outcomes related to writing (Yan, 2023a; Zou & Huang, 2023).

Therefore, this study was an attempt to fill the gap by providing research insights on the AI-assisted methods for imparting CT skills in EFL writing. On the one hand, this research provides insights that ChatGPT can be used systematically and collaboratively to deepen higher-order thinking during the writing process. This study's findings have theoretical and conceptual implications regarding cultivating CT in higher education. On the other hand, the acceptance of the CT-oriented instructional framework supported by ChatGPT was also investigated as an indicator of learners' satisfaction when using this approach. Examining the sentiments of early adopters in the emerging market can offer valuable insights regarding potential obstacles and opportunities (Menon & Shilpa, 2023). It cultivated valuable perspectives for optimizing the utilization of ChatGPT and analogous generative AI tools in

writing pedagogy within higher education. However, due to some factors such as habit formation and enthusiastic adoption (Lavidas et al., 2024), it is imperative to underscore that educators must implement structured guidance to cultivate critical engagement with AI-driven technologies in writing contexts, thereby mitigating students' over-reliance on such tools and preserving the integrity of autonomous cognitive development. This study also highlights the pros and cons of the role of AI scaffolding in this research. Instructors should also consider the ethical problems that may occur during AI use in teaching practices, such as informing students of the rules of using ChatGPT during the training.

The research questions examined in the research were:

- 1) How does ChatGPT affect EFL learners' writing performance through CT-oriented writing intervention?
- 2) What changes will occur to EFL learners' acceptance of ChatGPT as a result of the intervention?

2. Literature review

2.1. Critical thinking instruction

CT instruction is an integral component of the pedagogical framework for CT, aimed at promoting the effective integration of CT skills within educational practices (Glodjo, 2018). It is related to how to teach CT, which refers to the teaching and learning process that aims to develop the ability to think clearly, logically, and independently (Fahim & Masouleh, 2012; McLaren, 2015; Monchinski, 2008). Like Paulo Freire's emphasis, critical pedagogy is intertwined with cultivating "critical consciousness" (Monchinski, 2008, p. 4), and thinking-oriented pedagogy usually takes technology as a keyword in wide-context education (Rasa et al., 2023). Therefore, critical pedagogy transforms marginalized individuals into a cohort characterized by heightened critical awareness. It encourages learners to act, interpret, analyze, evaluate, and infer information while justifying the underlying factors influencing those cognitive judgments (Facione, 2020).

Scholars propose that the implementation of CT instruction enhances academic achievements in the mathematics field (Chimmalee & Anupan, 2023), medical field (Sullivan et al., 2024), psychology field (Nielsen et al., 2022), etc. Research in the field of English language learning and teaching has shown that the application of CT instructional strategies can enhance various literacy skills, e.g. reading (Hastuti et al., 2022), writing (Susilo et al., 2021), listening (Shahani et al., 2022), and speaking (S. Yin et al., 2024).

Different types of CT instructions can be employed during the writing process. Abami's study identified and examined a range of generic and content-specific CT skills, ultimately determining that several effective types of instructions for promoting CT include generic, infusion, immersion, and mixed approaches within the CT curriculum. (Abrami et al., 2014; Ennis, 1989). Due to Yin's team, a large part of the researchers (Dong, 2017c; Ebadi & Rahimi, 2018; Gao et al., 2017; Irgin & Bilki, 2024; Kaviani & Heidar, 2020; Murtadho, 2021; Yang et al., 2013) proved the influential role of infused CT-oriented approach in EFL writing (Yin et al., 2023).

2.2. CT-oriented EFL writing

CT-oriented writing has several main features in: process-based, explicit instruction, and context-based activities (Abrami et al., 2014; Yin et al., 2023). Various strategies can be used in the CT-oriented writing approach, such as teacher or peer feedback (Avni & Finn, 2020; Irgin & Bilki, 2024; Lin & Xiang, 2019; Mall-Amiri & Sheikhy, 2014; Susilo et al., 2021), social interactive environment (Gao et al., 2017; Irgin & Bilki, 2024; Murtadho, 2021; Safari, 2020), explicit CT concepts cognitive learning (Ahangari & Sepehran, 2014; Chason et al., 2017; Jubier et al., 2024; Mok, 2010), & questioning techniques

(Azevedo et al., 2023; Etemadzadeh et al., 2013; Miri & Azizi, 2018; Wale & Bishaw, 2020).

Different types of CT-oriented strategies have been found to contribute to writing competencies. For example, stimulating students' cognitive processes by presenting thought-provoking questions that foster student involvement and effectiveness in augmenting CT skills in written works (Ebadi & Rahimi, 2018; Panadero & Lipnevich, 2022). Learners collaborate to get feedback from one another using guided worksheets or checklists for more thoughtful and analytical prose (Chason et al., 2017; Irgin & Bilki, 2024), encourage reflection on CT elements and their writing (Dong, 2017a), get better information processing (Akbari et al., 2018; Anderson et al., 2023), and actively involve students in learning (Susilo et al., 2021).

In this research, the CT-oriented approach is based on one infusion type of CT instruction (Dong, 2017a), which includes four main steps: CT-oriented brainstorming, drafting, CT-oriented peer review, and revision. The two CT-oriented activities are conducted to infuse the teaching and practice of CT skills in EFL learners' writing practices and peer interaction (Dong, 2017c). The structured exercise of CT-oriented brainstorming is intended to support students in generating ideas by systematically considering the eight essential elements (Paul & Elder, 2021) of CT relevant to the assigned activities. Developed by Dong (2017c), a CT-oriented brainstorming worksheet is provided based on the elements of thought in Paul and Elder (2012). Students were engaged in a discussion on the eight essential elements of CT by posing and responding to the relevant inquiries. The CT elements enabled individuals to thoroughly comprehend the assigned tasks, thereby aiding in effectively generating ideas for their writing (Dong, 2017c).

After completing their first drafts, students participated in a peer review session focused on CT skills. During these tasks, they were expected to analyze their classmates' essays and formulate questions to seek clarification or additional information. This process was guided by Dong's CT-oriented peer-review checklist, which is based on the intellectual standards of Paul and Elder (2012). The peer review checklist is a guiding tool to inform students of the criteria to evaluate and improve their ideas and remind them of the nature of good writing. With this peer review checklist, students are encouraged to identify problems in their peers' essays by asking the guiding questions when they find anything inaccurate, vague, irrelevant, etc. At the same time, these questions are supposed to trigger students' reflections on their thinking revealed from their writing (Dong, 2017c).

2.3. English writing facilitated with technology

Through engaging in meaningful discussions and peer interactions during the writing process, learners have the opportunity to construct knowledge from these exchanges, involving activities such as generating ideas, providing explanations, clarifying concepts, and integrating feedback from peers, as well as consulting with peers and instructors for further clarification and insight (Etemadzadeh et al., 2013; Irgin & Bilki, 2024; Lu & Xie, 2019; Panadero & Lipnevich, 2022; Rasa et al., 2023; Sharadgah et al., 2019). Previous research has applied online techniques in CT-oriented writing approaches, such as formative feedback schemes through online systems (Huang, 2016), debates or inquiry-based discussions (Wale & Bishaw, 2020), and network-based writing blogs (Loncar, Schams, & Liang, 2021). However, previous research mainly focuses on traditional Web 2.0 technology, with the help of social interactive platforms (Gao et al., 2017; Irgin & Bilki, 2024; Yang & Gamble, 2013; Yih Mah et al., 2021) or other applications such as online games (Stapa & Ibaharim, 2020). All of these only emphasize user interaction, while there is a lack of emerging technologies, such as the application of OpenAI's GPT-series models in CT-oriented writing education.

Only a few researchers have tried using GPT to assist with writing, but they only focused on the superficial use as an "information browser" to get direct answers without critically assessing the answers, which

cannot prompt guiding questions to scaffold students' thinking (T Aldulaijan & M Almalky, 2025, p. 22). A recent study that collaborated ChatGPT in argumentative writing classrooms summarized that it is only functionalized as a "knowledgeable peer who is unfortunately unable to initiate conversations and guide thinking development" (Su et al., 2023, p. 8). Recent research advocated that high-quality interaction through the use of social technology could improve learners' CT (Pang, 2022). Therefore, more research should be conducted to explore the application of ChatGPT in EFL writing, especially from a CT-oriented perspective.

Technology brings convenience and efficiency to EFL learners' writing, such as process-oriented works (Irgin & Bilki, 2024), generating online peer feedback (Panadero & Lipnevich, 2022) providing support for both the cognitive and metacognitive aspects of learning (Molenaar, 2022), boost their ability and productivity and save time (Al-Zahrani, 2023; Almassaad et al., 2024). However, Generative AI also presents risks for EFL learners' writing. Lodge et al. (2023) identified two distinct interaction patterns with AI tools: intellectual partner and information browser. The former represents a higher-order use of AI, in which students actively engage with AI by posing questions that require intellectual effort, particularly in CT-oriented writing practices. This approach involves selective use and critical evaluation of generated ideas (Dahri et al., 2024; T Aldulaijan & M Almalky, 2025). In contrast, the latter reflects a more superficial use, where students treat Generative AI as a mere search engine, retrieving direct answers without critically assessing their validity (Lodge, Yang, Furze, & Dawson, 2023; Vargas-Murillo et al., 2023).

2.4. ChatGPT acceptance

To ensure the successful adoption of ChatGPT in higher education, extensive research has been conducted to determine the critical aspects influencing students' acceptance of ChatGPT. Observed from these studies, the most widely used models to examine the adoption and behavioral intention of using ChatGPT are the Technology Acceptance Model (TAM) (Greiner et al., 2023; Rahman et al., 2023), the Unified Theory of Acceptance and Use of Technology (UTAUT) (Menon & Shilpa, 2023; Strzelecki, 2023b), and the Unified Theory of Acceptance and Use of Technology (UTAUT2) (Strzelecki, 2023a).

A growing number of studies have been conducted to explore the threats and benefits regarding the use of ChatGPT in EFL or ESL contexts, such as EFL writing (Boudouaia et al., 2024; Salam, 2024; Wang, 2025). Under the context of digital learning, Liu and Ma (2023) employed the TAM and generated empirical insights into EFL learners' use of ChatGPT. Their study revealed that learners tended to have a higher degree of behavioral intention if they held positive attitudes regarding the effectiveness of ChatGPT, which further impacted their actual use of ChatGPT in English learning in informal settings. Testing the use and acceptance of ChatGPT in EFL writing, experimental research by Boudouaia et al. (2024) provided results that students from the experimental group performed better in EFL writing practice, demonstrating that students' use of ChatGPT boosted their writing skills, which further influenced their acceptance. Zooming into using ChatGPT during EFL students' writing process, Salam (2024) specifically explored students' learning experiences with ChatGPT in a specific course. His research revealed that most participants preferred using ChatGPT as an academic writing assistant rather than directly generating complete essays.

Recent research shows that the UTAUT2 model is widely applied to test new technologies in higher education contexts. While ChatGPT is gaining popularity as an efficient learning tool that can boost motivation and learning outcomes, it should not be assumed that all learners will adopt ChatGPT and reach a high degree of acceptance, especially in specific EFL writing disciplines. Thus, taking the acceptability of ChatGPT into account would be crucial in the EFL education field, especially measuring it from the user's side. Based on the above, ChatGPT acceptance may include the essential constructs of the UTAUT2. We modified the predictors of the original UTAUT2 by

removing social influence, facilitating conditions, and price value, as interventions will be implemented in this research.

3. Methods

3.1. Participants

Convenience and purposive sampling were performed to recruit participants for this research. According to several researchers (Hair et al., 2018; Memon et al., 2020; Suhr, 2006), for questionnaire measurement, the recommended ratio of items to respondents is 1:5, with an even better effect achieved at a ratio of 1:10 (Rahi & Alnaser, 2019). Given that there were 24 measurement items contained in this research, a sample size of at least 120 is required for a meaningful result, and ideally, 240 samples for an even better effect. Based on this, 250 first-year undergraduate participants from three public universities in Hebei province, China, including 135 male students (54 %) and 115 female students (46 %), were recruited online to guarantee a more trustworthy result for this research. They were in an intermediate B2 level based on the Common European Framework of Reference for Languages (CEFR), which means they have the ability to communicate easily in a clear and detailed manner in most situations. Although they are in different majors, learning English is compulsory for them, and every week, students have two lessons of 120 min. One compelling rationale for intentionally enrolling first-year students is that the initial year of higher education presents a valuable opportunity for enhancing students' capacity for CT. According to a famous Chinese expert in CT, the educational setting aimed at fostering CT abilities in the initial phases of the university is more effective in catering to the requirements of students possessing lower to intermediate levels of CT proficiency (Wen et al., 2018).

Chinese researchers identified that writing entails the capacity to articulate written communication through the collaborative process of language and thinking ability (Dong, 2017c; Gao & Wen, 2017), and educators wanted to find out if a CT-oriented writing intervention that was implemented through large language models (i.e. ChatGPT) would have positive impacts on their writing proficiency and behavior of using technology into EFL writing. Then, they could implement an intervention on a larger scale if there were positive results.

3.2. Procedures

The study utilized a quasi-experimental methodology. The intervention was implemented within the extra English writing classes attended by the participants. The study employed a mixed research methodology utilizing a pretest-posttest research design to examine potential enhancements in participants' proficiency in CT as demonstrated in their writing and their acceptance of incorporating ChatGPT in CT-oriented EFL writing tasks. The intervention was administered to the students at their own university.

A proficient English writing and CT instructor served as the trainer of the intervention by delivering CT-oriented writing lessons of the same design in the three public universities, thereby enhancing the validity of the intervention. Another lecturer, a pre-service English teacher, facilitated the students' utilization of ChatGPT to address challenges encountered during the classroom intervention. In order to guarantee the availability and convenience of the amenities, instructional sessions for participants focused on writing were conducted in multimedia classrooms equipped with internet access and computer terminals. Meanwhile, a pretest involving 25 students was conducted before the full-scale data collection. During the pretest, both quantitative responses and qualitative feedback on item clarity and relevance were collected. Based on students' feedback, the wording of several items was refined to ensure that students could accurately interpret questions related to CT in the context of EFL writing.

The participants conducted the pre-writing assessment and

acceptance survey in the classroom one week prior to the intervention. Subsequently, they completed the post-writing evaluation along with the same acceptance questionnaire immediately after the intervention. The researchers, comprising the authors of this study, the instructor, and the research assistant, engaged in multiple rounds of talks to guarantee the appropriateness of the lesson plans, instructional materials, writing prompts, task arrangements, and questionnaires for the intervention.

In addition, further interviews were undertaken with the three designated public universities, resulting in a total of 25 interviewees. The interview questions for participants were designed to explore their experience of the AI-assisted brainstorming and peer-review activities regarding the effects of ChatGPT on their learning of CT and EFL writing.

3.3. The intervention

The main content of the intervention classes was to allow the participants to practice CT-oriented writing, collaborating with ChatGPT by employing the AI-assisted CT-oriented brainstorming worksheet and AI-assisted CT-oriented peer-review checklist. The intervention adopted infused CT instruction, which consisted of the teacher's instruction, students' collaborative discussion with ChatGPT, note-taking, and writing. Except for the pretest (week 1) and posttest (week 8), the intervention consisted of a two-hour lesson from week 2 to week 7 (6 weeks), which was divided into four stages: 1) brainstorming, guided by AI-assisted CT-oriented brainstorming worksheet using ChatGPT, 2) drafting, 3) peer-reviewing, guided by AI-assisted CT-oriented peer-review checklist, and 4) Revising, grounded in the framework of Paul and Elder's elements of thought and CT intellectual standards (Dong, 2017c; Paul & Elder, 2021). Throughout the eight-week interventions,

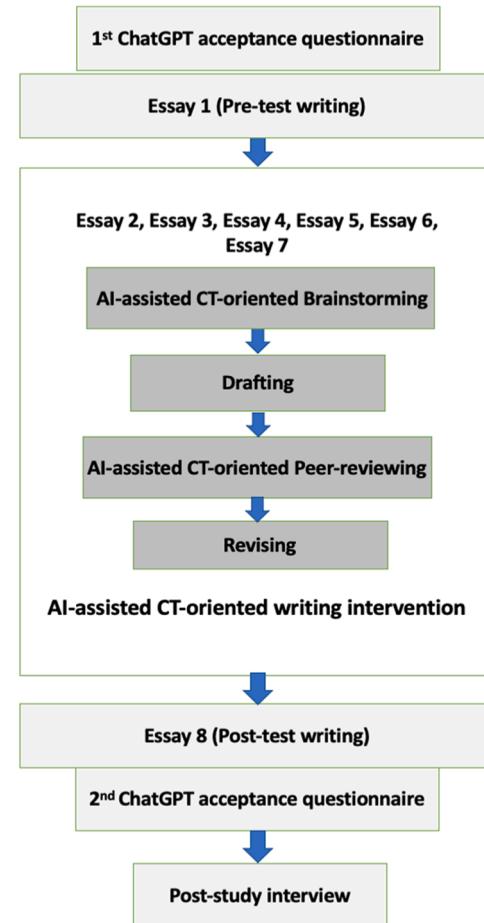


Fig. 1. Research procedure.

every student was equipped with a free version of the ChatGPT account. Fig. 1 shows some parts of the intervention.

At the beginning of the study, the pre-test essay was finished by the original CT-oriented instruction guided by brainstorming and peer-reviewing worksheets (Dong, 2017c) by real peers. The difference between the innovative intervention and the original CT approach is the scaffolding role of ChatGPT guided by an AI-assisted brainstorming worksheet and AI-assisted peer-review checklist during the brainstorming and peer-review stages. The instructor introduced and taught how to use the two worksheets before the writing practices.

3.3.1. AI-assisted Brainstorming worksheet

Based on Paul and Elder's eight elements of thought (Paul & Elder, 2021), there are three phases guided by the AI-assisted brainstorming worksheet (See Appendix 1), including 1) shaping and generating, 2) expanding and supporting, and 3) verifying and finishing.

The questions will help determine the purpose of the writing, identify key questions to answer, choose information to include, clarify key concepts, generate viewpoints, and make appropriate inferences, assumptions, and implications to draw a convincing conclusion (Dong, 2017c; Paul & Elder, 2021). Considering all of these essential components, it is probable that students will formulate concepts on what and how to compose their first-version drafts and demonstrate a more profound comprehension of the relationship between the CT components and their writing.

Guided by these questions, students took notes of the key elements they needed to identify and mark in the *Notes* column. Once they took notes, they made clear the ChatGPT input and started to wait for the answer generated. They selected and wrote answers they needed from the output generated by ChatGPT and then began to write their draft inspired by their virtual knowledgeable peer.

3.3.2. AI-assisted Peer-review checklist

After completing their first draft, students were required to use the nine intellectual standards to evaluate their own essays through the interaction with ChatGPT. Through the corresponding prompt, ChatGPT was required to identify problems concerning *clarity, accuracy, precision, etc.* It also gave examples and provided more details or further explanations. Based on the questions and answers, the students would understand their writing problems and how to improve their CT perspective. By using the AI-assisted peer-review checklist (See Appendix 2) in this study, the participants were supposed to be able to familiarize themselves with these CT intellectual standards and gain a deeper understanding of their connections to quality writing.

Following the completion of their initial draft, students were required to apply the nine intellectual standards to assess their own essays by engaging with ChatGPT. Utilizing the given prompt, ChatGPT was tasked with identifying problems of clarity, accuracy, precision, etc. It also included examples and supplied additional details to better understand the problems regarding the nine CT standards. Through examining the inquiries and responses, students can gain insights into their writing challenges and strategies for enhancement from a CT lens. Using the AI-facilitated peer-review checklist, participants were expected to become more familiar with the CT criteria and better comprehend how these standards relate to high-quality writing.

Here is an example of ChatGPT output feedback using the AI-assisted peer-review checklist. For instance, in terms of the last CT standard - *fairness*, learners got helpful feedback from ChatGPT with the relative prompt:

Do you think my following thesis writing is clear to understand?
Could you give some examples of why it is not clear?

Respect for All Languages:

- Avoid phrases that could be interpreted as culturally insensitive or biased, such as suggesting that a language may not be "very

civilized." This implies an unfair judgment about languages and their speakers. All languages are complex and developed systems of communication that are part of a culture's identity and history.

Avoiding the Notion of 'Unsuitable' Languages:

- The idea of eliminating "unsuitable languages" is problematic because it implies some languages are inherently less valuable or necessary than others. This is a form of linguistic discrimination and fails to recognize the equal value of all languages.

3.4. The instruments

Data were collected from participants using a test, questionnaire, and semi-structured interview. First, argumentative English writing tests were conducted to measure participants' EFL writing performance.

3.4.1. Pretest and posttest writing

Before the study, a pre-test essay was conducted to write an IELTS argumentative essay with a non-AI-supported CT-oriented approach scaffolded by peers. After a six-week AI-assisted CT-oriented intervention training scaffolded by ChatGPT, a post-test essay without AI was also conducted immediately after the one-semester intervention to see whether the intervention could positively change students' writing proficiency. Two lecturers independently assessed the pre- and post-writing samples to mitigate bias, utilizing a rubric developed by Dong (2017b) to assess specific CT in EFL writing.

3.4.2. ChatGPT acceptance questionnaire

Another main data collection instrument was a self-administered ChatGPT acceptance questionnaire (See Appendix 3), designed and sent to participants online to measure students' AI-assisted CT-oriented strategy use in EFL writing. Based on prior research on technology acceptance, especially within the information system (IS)/information technology (IT) field, the questionnaire was adapted to suit the CT-oriented EFL writing context. This questionnaire was divided into two sections. The first section included three items to get students' basic information (i.e., gender and major) and their previous experience with ChatGPT. In the second section, there were twenty-four items in total, which were used to measure the six constructs presented in Table 1. Each variable was measured by four items, adapting from previous research in the IS/IT field and AI field (Polyportis & Pahos, 2024; Venkatesh et al., 2012). A 5-point Likert scale ranging from "1 =strongly disagree" to "5 =strongly agree" was used when designing the questionnaire to score the response. As this research is located in the context of Chinese higher education, the questionnaire was translated into Chinese to avoid misunderstandings among the participants. Meanwhile, to avoid missing information, all items were subjected to a back-translation process by bringing in an independent translator.

Before conducting any statistical analysis, reliability was examined to guarantee the results' consistency and reduce measurement error. Thus, Cronbach's alpha was initially assessed for all constructs. As suggested by Hair et al. (2018), a scale is generally considered reliable if the value of Cronbach's alpha for each construct is equal to or greater than 0.70. Table 2 shows that Cronbach's alpha values of this questionnaire are 0.913 (pre-test) and 0.909 (post-test), both greater than the critical value of 0.7, meaning its degree of reliability is good. In addition, Cronbach's alpha of every single construct in the research model was also above 0.7, suggesting that each latent variable in the questionnaire measured by this model had a relatively good level of reliability.

To verify the accuracy of the data, the researchers calculated the Kaiser-Meyer-Olkin (KMO) values of sample adequacy and performed Bartlett's test of sphericity (Table 3). It can be observed that the KMO value is 0.9 (pre-test) and 0.898 (post-test), and the *p-value* of Bartlett's test is less than 0.05 on both pre-test and post-test, indicating a high validity of the measurement scale. Therefore, based on the results of the

Table 1
Measurement Scales.

Construct	Code	Items	Reference
Performance expectancy	PE1	I would find ChatGPT useful in improving my CT-oriented EFL writing skills.	Polyportis & Pahos, (2024)
	PE2	Using ChatGPT enhances my EFL writing motivation.	
	PE3	Using ChatGPT might increase my productivity in CT-oriented EFL writing.	
	PE4	Using ChatGPT helps me to accomplish my EFL writing assignments more quickly and with high quality.	
Effort expectancy	EE1	It is easy for me to learn how to use ChatGPT for CT-oriented EFL writing as a student.	Polyportis & Pahos, (2024)
	EE2	My interaction with ChatGPT is clear and understandable in the process of using it to practice CT-oriented EFL writing.	
	EE3	Using ChatGPT to assist CT-oriented EFL writing is easy for me.	
	EE4	Developing CT-oriented EFL writing skills using ChatGPT is easy to achieve	
Hedonic motivation	HM1	I feel fun using ChatGPT to develop CT-oriented EFL writing skills.	Venkatesh et al., (2012)
	HM2	I enjoy using ChatGPT to practice CT-oriented EFL writing.	
	HM3	Using ChatGPT for CT-oriented EFL writing training makes me feel relaxed and happy.	
	HM4	Using ChatGPT makes my CT-oriented EFL writing process enjoyable and interesting.	
Habit	HA1	Using ChatGPT to train CT-oriented EFL writing skills has become my habit.	Venkatesh et al., (2012)
	HA2	I feel addicted to using ChatGPT to practice CT-oriented EFL writing.	
	HA3	I frequently use ChatGPT to assist my CT-oriented EFL writing practices.	
	HA4	I always prefer to consult ChatGPT first when critical thinking is needed in EFL writing.	
Trust	TR1	I believe that ChatGPT is trustworthy.	Polyportis & Pahos, (2024)
	TR2	I trust ChatGPT when I use it to develop CT-oriented EFL writing skills.	
	TR3	I do not doubt the content that ChatGPT provides when I use it to practice CT-oriented EFL writing.	
	TR4	I believe using ChatGPT for cultivating CT-oriented writing will not put me at risk.	
Behavioral intention	BI1	I intend to continue using ChatGPT as a scaffolding tool in future CT-oriented EFL writing activities.	Polyportis & Pahos, (2024)
	BI2	I will use ChatGPT to develop CT-oriented EFL writing skills in the future.	
	BI3	I will always try to use ChatGPT for CT-oriented EFL writing practice.	
	BI4	I strongly recommend my classmates use ChatGPT to cultivate critical thinking ability in EFL writing.	

reliability and validity test, the questionnaire of this research is suitable for further analysis.

3.4.3. Semi-structured interview

In addition, the semi-structured interview contained seven open-

Table 2
Cronbach's Alpha of the questionnaire used.

Item	Pre-test	Post-test	Number of items
PE	0.821	0.819	4
EE	0.846	0.868	4
HM	0.841	0.815	4
HB	0.865	0.813	4
TR	0.781	0.788	4
BI	0.812	0.818	4
Overall	0.913	0.909	24

Table 3
Bartlett's Test and KMO.

	Pre-test	Post-test
KMO Measure of Sampling Adequacy	0.900	0.898
Bartlett's Test of Sphericity	Approx. Chi-Square	2750.302
df	276	276
Sig.	.000	.000

ended questions, which could be divided into four sections: factors affecting students' acceptance of AI-assisted CT-oriented writing, students' perception of brainstorming and peer review, students' perceived AI interaction, and students' attitudes towards different types of CT-oriented writing.

3.5. Minimizing potential threats

In the pre-post-test design, internal validity refers to the extent to which observed changes from the pretest to the posttest can be confidently attributed to the treatment itself, rather than to extraneous variables (Reichardt, 2019). Threats to internal validity arise when factors other than the intervention potentially account for the observed differences, thereby inflating or attenuating the estimated treatment effect. A critical consideration in establishing causal inference is the presence of covariation—that is, whether changes in the independent variable (the treatment) are systematically associated with changes in the dependent variable (Creswell & Creswell, 2023). With the aim of getting a valid cause-and-effect result for the target m-CTI treatment, various factors are considered for minimizing the potential threats, such as pretest-posttest or pretest-treatment interaction, maturation, history, and morality.

3.5.1. Pretest-posttest or pretest-treatment interaction

Testing threats occur when participants' prior exposure to the pretest or external variables leads to improved scores that are unrelated to the actual effects of the intervention (Fraenkel et al., 2023). To mitigate this threat, our design implemented an eight-week interval between pretest and posttest administrations, and questionnaire items were automatically randomized at each administration to discourage rehearsal effects. Prior research on pretest sensitivity indicates that even when the same higher-order thinking items are re-administered, respondents do not reliably recall their original answers. Facione (1991) found that over the course of a semester, instrument effects are minimal because participants cannot precisely remember their prior choices. Additionally, distinct writing prompts were employed at posttest to eliminate direct practice effects. And a one-week interval between pretest and intervention was implemented to attenuate memory effects. Prior studies (MacLeod et al., 2018; Welhaf et al., 2022) suggested that such intervals reduce the salience of pretest content in writing-focused interventions. Overall, these measures ensure that observed gain scores more accurately reflect the treatment effect rather than artifacts of repeated testing.

3.5.2. Maturation

Maturation refers to natural developmental or experiential changes

in participants that occur over time and can confound treatment effects. In this study, only 6 weeks (2-hour classes once a week) of intervention and data collection was not long enough to cause the maturation threats in this research. The intervention duration was minimized to reduce natural skill development unrelated to the treatment. Accordingly, maturation threats to internal validity in this research are considered negligible.

3.5.3. History

To mitigate potential confounding effects arising from historical variables, rigorous measures were implemented to ensure strict adherence to experimental protocol. Participants were systematically excluded from participating in analogous CT interventions or related extracurricular activities throughout the study duration. During the data collection phase, participants received explicit written instructions mandating immediate disclosure of any concurrent enrollment in external CT-related educational programs or skill development initiatives. This disclosure protocol enabled the principal investigator to promptly identify and eliminate non-compliant subjects from the final cohort analysis, thereby preserving the internal validity of the experimental design.

3.5.4. Morality

Participant attrition, such as withdrawal or reduced engagement, can introduce bias and compromise study validity (Fraenkel et al., 2023). In this study, any student who left the course would have been omitted from the analysis; however, because the intervention was integrated into the regular semester schedule and participants could receive extra credit for completion, all enrolled students remained, resulting in no attrition.

Additionally, we acknowledge the inherent limitations of a single-group design. However, this approach was strategically chosen for the following reasons.

- Exploratory focus: As an initial investigation into ChatGPT's role in CT-oriented EFL writing, the design aligns with recommendations for pilot studies in emerging AI-driven educational research.
- Practical constraints: Resource limitations (e.g., institutional approvals, matching control groups across three universities) necessitated a focused intervention.
- Mitigation measures: All participants received identical instruction and ChatGPT access, minimizing external variability.

4. Data analysis

The paired *t*-test was used to compare the data collected before and after the AI-assisted CT-oriented writing intervention in writing proficiency and ChatGPT acceptance. Following quantitative data collection, further semi-structured interviews were conducted to investigate students' perceptions of AI-powered CT-oriented writing tools. Since the interviews were conducted in Chinese, the Chinese transcriptions were translated into English for data analysis. Qualitative data collected from the semi-structured interview was analyzed in an inductive way (Thomas, 2006). After identifying and omitting the redundant information, the transcripts were summarized to categorize the useful information into the same topics as the questionnaire.

4.1. Results

4.1.1. Quantitative results

4.1.1.1. Writing performance before and after the intervention. As shown in Table 4, the participants' mean score of their CT-oriented writing practice was lower than that of the most recent practice. The paired *t*-test result showed a significant difference between students' pre-post-test writing performance ($p < .05$). The mean for students' writing scores

Table 4

Comparison between participants' pre-post-writing performance.

	Pre-test	Post-test	t-Value	Sig. (2-tailed)
	Mean (SD)	Mean (SD)		
Writing proficiency	26.13 (3.08)	29.58 (3.16)	-1.452	.003

Note: $p < .05$

increased from 26.13 (SD=3.08) to 29.58 (SD=3.16).

The result suggests that intervention supported by the two AI-assisted worksheets using ChatGPT is helpful in improving students' writing quality in terms of CT.

4.1.1.2. ChatGPT acceptance before and after the intervention. Based on the UTAUT 2 model, students were asked to evaluate their acceptance of several statements regarding their use of ChatGPT for CT-oriented EFL writing. After analyzing the collected data using SPSS 25, the mean value (Table 5) was calculated and a *t*-test (Table 6) was conducted to make a comparison between pre-test and post-test. Compared to those of the pre-test, the mean value of each construct has all improved in the post-test, indicating that participants' overall acceptance of using ChatGPT for CT-oriented EFL writing activities has become higher.

Observing from the results, students' performance expectancy ($M=10.592$, $SD=4.38$) in using ChatGPT for CT-oriented writing activities performed at a high level in the post-test. Combining the *t*-test results in Table 6 ($t = 6.75$, $p < 0.01$, Cohen's $d=0.43$), it can be found that there was a significant improvement after the intervention was conducted. Same with PE, students' effort expectancy ($M=11.616$, $SD=4.52$) and trust ($M=10.316$, $SD=3.57$) also increased in the post-test. The *p* value of these two constructs were both below 0.05, indicating a significant improvement. Students' behavioral intentions to use ChatGPT for CT-oriented EFL writing in the future became higher as well, from the pre-test ($M=8.536$, $SD=3.69$) to the post-test ($M=11.244$, $SD=4.33$). Besides, although there was a slight increase in students' hedonic motivation ($M=8.972$, $SD=4.04$) and habit ($M=10.512$, $SD=4.40$) of using ChatGPT after the intervention, it was still lower compared to other constructs.

4.1.2. Qualitative analysis

The interview revealed in-depth themes from interviewees, enabling a deeper understanding of the concepts identified in the questionnaire analysis. More precisely, an analysis was conducted on the themes outlined in Table 7.

The interview demonstrates ChatGPT affects EFL learners' CT-oriented writing in several main aspects: CT concepts understanding, Model discourse embedded with CT elements, information, and feedback, as well as logic. In terms of CT concepts understanding, 72 % of interviewee agreed that ChatGPT could deepen their understanding of CT items, as AI could provide an environment that can show how CT is reflected in writing instead of only knowing the definition of CT.

Table 5

Descriptive Analysis.

Item		Mean	Std. Deviation	N
PE	Pre-test	8.6280	3.77394	250
	Post-test	10.5920	4.37700	250
EE	Pre-test	9.2280	4.19409	250
	Post-test	11.6160	4.51910	250
HM	Pre-test	8.5640	3.86285	250
	Post-test	8.9720	4.04036	250
HB	Pre-test	9.3320	4.18999	250
	Post-test	10.5120	4.40083	250
TR	Pre-test	8.1920	3.31289	250
	Post-test	10.3160	3.57144	250
BI	Pre-test	8.5360	3.69096	250
	Post-test	11.2440	4.32612	250

Table 6
T-test of pre-test and post-test.

Pair	Item	t	Sig. (2-tailed)	Cohen's d
1	PrePE-PostPE	-6.750	.000	0.426882281
2	PreEE-PostEE	-7.649	.000	0.483764122
3	PreHM-PostHM	-1.801	.073	0.113911116
4	PreHB-PostHB	-4.289	.000	0.271230074
5	PreTR-PostTR	-8.182	.000	0.517464424
6	PreBI-PostBI	-9.136	.000	0.577786786

Table 7
Theme analysis of interviews.

Categories	Sample excerpts	Amount	Percentage
In-depth CT concepts understanding	<i>Now I know how CT is reflected in writing, instead of only knowing the definition.</i>	18	72 %
Model discourse with CT elements	<i>AI helps me how to relate and combine different elements in a systematic way by providing model discourse following each prompt.</i>	17	68 %
Information and feedback	<i>I prefer AI-assisted CT-oriented writing since ChatGPT can provide more effective information and feedback than my real peers.</i>	16	64 %
Logic	<i>I really fancy AI-assisted CT-oriented writing since I think I can be more logical with the help of AI to proceed based on guided questions.</i>	21	84 %
AI interaction	<i>It is efficient and effective, though it requires more patience and is a bit boring for the whole process.</i>	9	36 %

Meanwhile, 68 % of the interviewees expressed the role of ChatGPT in providing model discourse embedded with CT elements. A1 said: "I think ChatGPT is quite good since its output is like great models for us to reflect how CT works infused in the writing process". Similar to B2: "Those CT elements are quite abstract to operate, especially for me, who is not good at writing. While ChatGPT showed me what it is like in the writing process with detailed and clear material and discourse". Furthermore, 64 % of interviewees claimed that one of the main reasons they prefer ChatGPT is it can provide effective information and feedback. For example, C3 advised: "I really recommend to use ChatGPT in CT-oriented writing, as it is quite knowledgeable than my real peers. Usually, my peers cannot give me too much useful feedback during the peer-review stage". Also, B1 pointed out: "Sometimes we (with real peers) kept silent during the brainstorming stage, and it seems we do not know how to think and lack of ideas". Another significant word that they almost mentioned is logic. 84 % of them thought and expected they could be more logical with the help of AI.

However, 36 % of interviewees still complained about the several stages that must be finished. For instance, A3 said: "Although it is quite effective, I think it is a bit boring and requires more patience to finish the whole process with so many prompts."

4.2. Discussion

This study explores the impact of the ChatGPT on EFL learners' writing proficiency through CT-oriented writing intervention and EFL learners' ChatGPT acceptance as a result of the intervention. This section will discuss significant findings by integrating quantitative and qualitative results.

Firstly, the quantitative results comparing writing tests before and after the intervention showed that ChatGPT significantly improved participants' writing proficiency more significantly than traditional CT-oriented writing with real peers. In alignment with prior studies (Guo et al., 2022; Irgin & Bilki, 2024; Su et al., 2023), the utilization of an AI-enhanced, CT-infused intervention supports the idea that ChatGPT

can effectively function as a virtual tutor or knowledgeable peer in aiding with concept organization, outline development, and feedback. However, previous research lacked empirical data to prove the role of ChatGPT in improving writing abilities, such as Su et al. (2023), who examined the potential of ChatGPT to scaffold at different stages of writing to provide support.

In addition, previous research focused more on automatic evaluation systems, which required students to provide an initial draft first and then seek feedback from ChatGPT. As Su et al. (2023) mentioned, ChatGPT is limited under such a framework since ChatGPT can only generate content upon request and cannot prompt guiding questions to scaffold students' thinking. Hence, besides the advantages of guiding "thinking development" (Su et al., 2023, p. 8), the two AI-assisted CT-oriented worksheets, grounded in Paul and Elder's CT framework, address this deficiency and have the potential to significantly enhance ChatGPT's generative capacity, enabling it to provide more detailed feedback, such as "clearer organization", "explicit linking of sub-claims to the main claim", and "specific evidence" (Su et al., 2023, p. 8), which is also reflected in the qualitative results. Unlike recent studies, this research employed a mixed-methods approach to enhance the validity of the results, rather than relying on a single-method design, as seen in Irgin and Bilki (2024).

Secondly, consistent with our prediction, a significant improvement in students' acceptance of using ChatGPT for EFL writing activities was found after the intervention. The findings support the original UTAUT 2 (Venkatesh et al., 2012) and echo previous research within the EFL and AI-assisted learning context. Prior studies have identified AI tools as valuable aids in English writing activities (Bai et al., 2022; Beck & Levine, 2023; Zou & Huang, 2023). The results of the present study are similar to and complement those studies by indicating that students generally held a more positive attitude toward using ChatGPT as a scaffolding of CT-oriented EFL writing. It also indicated that students were generally satisfied with their ChatGPT use experience and were willing to integrate it with their further CT-oriented EFL writing practices. One possible explanation for students' strong behavioral intention could be ChatGPT's writing affordances (Zou & Huang, 2023). Thus, training and guiding students to use AI tools such as ChatGPT in the EFL writing context can somehow change their opinions of education technologies, promote the integration of technology and writing, and influence their intentions to use technologies for EFL writing. Nevertheless, researchers are worried about the misuse of ChatGPT since students can use it to generate high-quality writing assignments or even cheat with it (Susnjak & McIntosh, 2024). As Tsai et al. (2024) mentioned, using ChatGPT to revise essays does not reveal students' writing competence. Although ChatGPT effectively provides in-time feedback to EFL students, researchers should consider the potential unfairness when evaluating writing assignments (Tsai et al., 2024).

Generally speaking, all indicators of students' acceptance of using ChatGPT in a CT-oriented EFL writing context showed an increasing trend after the intervention. However, an interesting finding of this research was that, unlike other constructs, students' hedonic motivations and habits in using ChatGPT barely increased after the intervention, which can also be found in the qualitative results. On the one hand, this is similar to the research findings of Rahman et al. (2023), that enjoyment had no substantial impact on participants' attitudes towards ChatGPT usage, and Maican et al. (2023) that hedonic motivation was not shown to be significant when students using AI. This finding indicates that when using ChatGPT as a writing assistant, Chinese EFL students value its effectiveness and usefulness more than its enjoyable elements. On the other hand, the relatively insignificant change in students' ChatGPT use habits could be attributed to the duration of the intervention in this research. Forming a habit requires repeated practice that lasts for a certain period while being exposed to only eight lessons of intervention might not be sufficient enough for students to form a habit of using ChatGPT constantly.

While the new approach may present challenges, such as

overreliance on AI, skill degradation, and the limitations of ChatGPT ([T Aldulaijan & M Almalky, 2025](#)), only a few students reported these concerns during the interview. This can largely be attributed to the instructor's proactive role in ensuring the appropriate use of ChatGPT. Before the experiment, the instructor explicitly outlined the objectives of the intervention and established clear guidelines for the ethical and effective use of AI tools ([T Aldulaijan & M Almalky, 2025](#), p. 20). Additionally, the two worksheets facilitated a structured, systematic approach to thinking, thereby enhancing students' awareness ([Aldossary, Aljindi, & Alamri, 2024](#); [Vargas-Murillo et al., 2023](#)). Another factor mitigating AI overreliance was the conceptualization of ChatGPT as an intellectual partner within a scaffolding framework. As [Lodge et al. \(2023\)](#) observed, student interactions with generative AI can foster co-regulation, wherein students actively seek regulatory support, such as requesting clarification of concepts. During the intervention, students engaged with ChatGPT collaboratively, analyzing how different CT components (e.g., organization, induction, summarization) applied to various writing contexts across diverse topics. Although some scholars have questioned ChatGPT's effectiveness in scaffolding CT development ([Chan & Hu, 2023](#); [Vargas-Murillo et al., 2023](#)), numerous studies ([Faisal, 2024](#); [Samala et al., 2024](#); [Tarankanti & Davidson, 2023](#)) have demonstrated its potential to enhance CT by broadening perspectives and providing insights within instructional design processes ([DaCosta & Kinsell, 2024](#); [Dahri et al., 2024](#)). These findings highlight the effectiveness of this approach in improving students' CT and writing performance. Rather than using ChatGPT merely as an information retrieval tool ([Lodge et al., 2023](#)), this in-depth learning strategy enables learners to monitor, evaluate, and refine their learning processes ([Dai et al., 2023](#); [Lodge et al., 2023](#); [T Aldulaijan & M Almalky, 2025](#)).

Additionally, the qualitative interviews were not merely supplementary but served two critical functions in complementing the quantitative findings. First, they facilitated triangulation by aligning emergent themes—such as students' emphasis on improved "logic" and exposure to model discourse with "CT elements". These are in line with the quantitative enhancements in CT standards (e.g., clarity, depth, and fairness). This alignment underscores the intervention's consistency across measurement modalities. Second, the interviews uncovered mechanisms behind ChatGPT's efficacy, revealing its role as a scaffold for self-regulated learning. Students highlighted ChatGPT's role in scaffolding self-regulated learning (e.g., "ChatGPT guided me to ask better questions"), elucidating how the tool enhanced CT beyond mere posttest scores. Future studies will expand the qualitative scope to track longitudinal behavioral changes, yet the current findings substantiate ChatGPT's potential as a cognitive partner in CT-oriented writing pedagogy.

This study yields three key pedagogical implications for integrating CT into EFL writing instruction through AI-assisted frameworks. First, educators should explicitly train students to formulate prompts grounded in [Paul and Elder's \(2021\)](#) critical thinking elements. Second, instructors ought to incorporate an intellectual standards-based rubric (e.g., clarity, logical coherence, fairness) during AI-assisted peer evaluation, encouraging learners to critically juxtapose AI-derived feedback with peer assessments to cultivate evaluative metacognition. Third, pedagogical approaches must position ChatGPT as a collaborative cognitive tool rather than an authoritative reference, requiring students to articulate rationales for incorporating or rejecting AI-suggested revisions. To optimize implementation, targeted teacher training programs should address prompt engineering, ethical usage protocols, and adaptive scaffolding strategies for heterogeneous proficiency levels. By systematically embedding these practices across iterative writing phases, educators can ensure AI tools function as complementary scaffolds that enhance, rather than undermine, learners' CT autonomy and writing proficiency.

5. Conclusion

The present study successfully implemented an AI-assisted CT-oriented writing intervention that guided the students to apply CT elements to EFL writing. It utilized ChatGPT, a more advanced knowledge peer for each participant, and was carried out in three public universities located in Hebei Province, China. By comparing the students' writing performance and questionnaire before and after the intervention, we found that the students improved their specific CT in writing (i.e. clarity, accuracy, precision, relevance, depth, breadth, logic, significance, and fairness). Furthermore, there was a notable increase in the students' acceptance of ChatGPT use in English writing.

These results provide additional evidence for the application of generative AI in language acquisition and introduce a novel approach for incorporating ChatGPT into EFL writing education: merging AI with CT-oriented instructions and learning. The students' increased writing achievement and acceptance mirror findings from related research across different fields, underscoring the utility and prospective relevance of incorporating ChatGPT in second language writing education ([Yan, 2023b](#)), and proved the necessity of developing and applying "thinking tools" in future-oriented pedagogies ([Rasa et al., 2023](#), p. 1164). But what should be noted, ChatGPT served solely as a supplementary tool in CT-oriented training within EFL writing instruction. Its primary role was to support the CT development process by acting as a knowledgeable peer, assisting students in applying higher-order thinking skills during brainstorming and peer review. With ChatGPT's guidance, students demonstrated a clearer understanding of the elements of thought and CT standards embedded in their writing.

Furthermore, potential challenges in implementing this AI-assisted CT-oriented approach, such as ethical concerns and the risk of misinformation, must be addressed. Instructors should clearly define the objectives of assignments to determine the appropriateness of ChatGPT usage ([T Aldulaijan & M Almalky, 2025](#)), establish for the responsible use of AI tools and enhance students' awareness ([Aldossary, Aljindi, & Alamri, 2024](#); [Vargas-Murillo et al., 2023](#)). Additionally, students should develop regulation skills to critically assess AI-generated content. For instance, the notes column in the worksheets enabled students to verify the accuracy and reliability of the information.

6. Limitations and future work

Some limitations should be considered. First, there is only one intervention group without a post-intervention follow-up assessment. It is imperative to administer a follow-up test after a one-month interval to evaluate the persistence of the intervention's impact in future research. The one-group-pre-post research design is limited in the result validity, such as the positive impact of other training courses or the continuous effects from the CT-oriented approach on their improvement of writing proficiency. Although this study included only one intervention group, the researcher implemented multiple measures to control potential confounding variables. To address potential confounding factors, rigorous measures were implemented, including standardization of instructional delivery through a consistent instructor, systematic control of participants' baseline characteristics (e.g., English proficiency, academic background, and prior critical thinking training), and methodological safeguards to mitigate threats to internal validity (e.g., pretest-posttest contamination, historical biases, and morality concerns). Such controls enhance the study's external and internal validity. Moreover, previous studies that employed two-group designs often treated ChatGPT as a general classroom aid rather than integrating it into a structured instructional intervention ([T Aldulaijan & M Almalky, 2025](#)). For instance, in [Elkhodr et al. \(2023\)](#), students used ChatGPT as a search engine when needed rather than following a structured set of prompts or procedures. In contrast, several high-quality studies have successfully employed a single-group intervention model, including a 14-week academic writing course conducted by [Irgin and Bilki \(2024\)](#) and a

one-semester course by T Aldulaijan and M Almalky (2025) with specific rubrics. These studies designed structured interventions with certain procedures that learners were required to follow. The present experiment aims to evaluate the effectiveness of the target AI-assisted CT-oriented instructional model that incorporates structured procedures and prompts, ensuring that students engage systematically with the intervention. Given the complex and higher-order nature of CT (Paul & Elder, 2021; Wen et al., 2018), mere exposure to AI-generated knowledge without structured, in-depth interaction is unlikely to yield significant CT improvements. Additionally, prior research has demonstrated that instrument effects do not occur over a one-semester treatment period, as participants are unlikely to remember their previous choices with sufficient precision to influence their responses on the post-test (Facione, 1991; Facione, 2020). Therefore, a carefully designed instructional model can generate valid results even within a single-group intervention framework. Indeed, the inclusion of a control group in traditional pen-and-paper formats and an experimental group that received AI-supported instruction would have highlighted the impacts of ChatGPT in a more targeted manner. Future research is recommended to incorporate a delayed assessment and employing a control group if the conditions permit.

Second, despite the data collection encompassing multiple universities and a diversified sample including respondents from three different institutions, it is important to acknowledge that the sample may still be influenced by self-reporting bias (Lavidas et al., 2022), thus imposing certain limitations on the research. Regarding this limitation, future research is suggested to include longitudinal assessments to examine whether these initial changes in perception are sustained over time or incorporate behavioral measures (such as usage data). Meanwhile, to increase the external validity of the results, future research should strive to replicate the research framework by employing a multi-site methodology encompassing participants from a broader range of universities or in different international contexts (Bryman, 2016; Lavidas et al., 2024).

In addition, the research involved participants in different majors,

including law, finance, marketing, insurance, economics, etc. While answering the questionnaire, no identifiable information was included (such as names), which limits this current research to link students' ChatGPT acceptance levels with their specific essay evaluations. Thus, to comprehensively assess the influence of ChatGPT on intention and behavior in EFL learning, it is advantageous to explore the moderating variables, such as field of study and English level, to enhance the precision of their impact estimation. Future research is suggested to include identifiable information to further explore the relationship between acceptance level and writing performance. In addition, although the sample of this research may have a nested structure (i.e., students within schools), the small number of clusters (i.e., only three schools) limited our ability to apply multilevel modeling. Thus, future research with a larger number of higher-level units is encouraged to adopt multilevel modeling for more accurate estimation of group-level effects. Furthermore, although this study utilized AI-assisted CT-oriented writing as a training tool only during the six-week intervention period, and learners did not use ChatGPT during the pretest or posttest writing tasks, the approach follows a process-oriented writing framework. Consequently, it is possible to observe a gradual improvement in the quality of these essays, which may provide valuable insights for future research.

CRediT authorship contribution statement

Xue Yin: Writing – original draft, Investigation. **Kun Dou:** Writing – original draft.

Declaration of Competing Interest

No potential conflict of interest was reported by the author(s).

Acknowledgments

The authors would like to thank all teachers and students participating in this project.

Appendix 1. Sample AI-assisted Brainstorming Worksheets

Phase One: Shaping and generating	Notes	ChatGPT prompt	Answers
PURPOSE: to write an argumentative essay X = an undergraduate student MAJOR QUESTION: Whether traveling to or working in a foreign country is the only reason to learn a foreign language? OTHER QUESTIONS: Is there any other reason someone should learn a foreign language? MAIN CONCEPTS: reasons for learning a foreign language TARGET VIEWPOINTS (Inspired by answers from ChatGPT in the latest procedure): "Practical Necessity," "Cultural Enrichment," and "Educational and Cognitive Benefits"	As an undergraduate student (X), I want to write an argumentative essay (PURPOSE) Discuss (the) "whether traveling to or working in a foreign country is the only reason to learn a foreign language" (MAJOR QUESTION) and "Is there any other reason someone should learn a foreign language" (OTHER QUESTION) Please provide several possible VIEWPOINTS around the 'reasons for learning a foreign language' (MAIN CONCEPTS)		1. **"Practical Necessity Viewpoint": - **Travel and Work Abroad**: This viewpoint emphasizes the practical necessity of learning a foreign language for ease of communication in travel or professional settings abroad. - **Expatriates and Immigrants**: Discuss how expatriates and immigrants benefit from learning the local language for social integration and daily life activities. 2. **Cultural Enrichment Viewpoint**: - **Access to Cultural Heritage**: Learning a language opens the door to understanding a culture's literature, music, film, and traditions. - **Cross-Cultural Understanding**: Explain how language learning promotes empathy and understanding between people of different cultures, reducing stereotypes and prejudice. 3. **Educational and Cognitive Benefits Viewpoint**: - **Cognitive Advantages**: Discuss research showing that bilingualism or multilingualism can enhance cognitive abilities, such as problem-solving and memory. - **Academic Opportunities**: Learning a foreign language can provide access to educational resources, scholarships, and study abroad programs.

(continued on next page)

(continued)

Phase One: Shaping and generating	Notes	ChatGPT prompt	Answers
			<p>4. **Personal Development Viewpoint**:</p> <ul style="list-style-type: none"> - **Personal Growth**: Language learning can be a fulfilling personal challenge that boosts confidence and self-esteem. - **Social Networking**: Discuss how knowing multiple languages can help build diverse social networks and foster friendships across cultures. In your essay, you may choose to focus on a few of these viewpoints or explore them all. Each viewpoint provides a different perspective on the importance and benefits of learning a foreign language beyond the traditional view of travel and work necessity. <p>The assumptions behind the categories "Practical Necessity," "Cultural Enrichment," and "Educational and Cognitive Benefits" in the context of learning a foreign language are based on the various intrinsic and extrinsic motivations that drive individuals to take up a new language. Here's what each category assumes regarding reasons for learning a foreign language:</p> <ol style="list-style-type: none"> 1. **Practical Necessity**: <ul style="list-style-type: none"> - This assumes that there is a functional requirement or need for language proficiency in daily life or work environments. It is often driven by immediate needs or goals such as: <ul style="list-style-type: none"> - Navigating daily activities in a country where the language is spoken. - Enhancing career prospects where language skills are a requirement or an asset. - Communicating effectively in personal or professional contexts. 2. **Cultural Enrichment**: <ul style="list-style-type: none"> - The assumption here is that language is a key to experiencing and understanding different cultures more authentically. It is based on the belief that language learning can: <ul style="list-style-type: none"> - Provide a richer, more nuanced appreciation of a culture's customs, values, and traditions. - Allow for personal growth and fulfillment through the exploration of diverse cultural expressions. - Foster respect and appreciation for cultural differences and commonalities. 3. **Educational and Cognitive Benefits**: <ul style="list-style-type: none"> - This category assumes that learning a foreign language has positive effects on the mind and educational outcomes. It suggests that language learning can: <ul style="list-style-type: none"> - Enhance cognitive abilities such as problem-solving, multitasking, and creativity. - Contribute to academic success by opening up access to resources and research in the foreign language. - Support cognitive health and potentially offer protective effects against cognitive decline. <p>Each of these assumptions underlines a different set of benefits that justify the time and effort invested in learning a foreign language, whether for immediate practical needs, cultural exploration, or long-term cognitive and educational gains.</p> <p>Certainly, each viewpoint is supported by various pieces of information and research that validate the assumptions about the reasons for learning a foreign language:</p> <ol style="list-style-type: none"> 1. **Practical Necessity**: <ul style="list-style-type: none"> - **Supporting Information**: - **Economic Benefits**: Studies have shown that bilingual employees can earn between 5 % and 20 % more per hour than those who speak only one language. - **Immigrant Integration**: Research indicates that language proficiency is a significant factor in
Phase Two: Expanding and supporting	<p>Inspired by answers from ChatGPT and summarize the needed ASSUMPTION:</p> <p>Practical Necessity: There is a functional requirement or need for language proficiency in daily life or work environments.</p> <p>Cultural Enrichment: Language is a key to experiencing and understanding different cultures more authentically.</p> <p>Educational and Cognitive Benefits: Learning a foreign language has positive effects on the mind and educational outcomes</p>	<p>What are "Practical Necessity", "Cultural Enrichment", and "Educational and Cognitive Benefits" (TARGET VIEWPOINTS) assuming regarding the reasons for learning a foreign language (MAIN CONCEPTS)?</p> <p>Regarding "Practical Necessity", "Cultural Enrichment", and "Educational and Cognitive Benefits" (TARGET VIEWPOINTS) followed by "there is a functional requirement or need for language proficiency in daily life or work environments", "language is a key to experiencing and understanding different cultures more authentically", and "learning a foreign language has positive effects on the mind and educational outcomes" around the reasons for learning a foreign language (MAIN CONCEPTS), please</p>	

(continued on next page)

(continued)

Phase One: Shaping and generating	Notes	ChatGPT prompt	Answers
		provide some information to support each viewpoint.	<p>the economic and social integration of immigrants, affecting job opportunities and community involvement.</p> <ul style="list-style-type: none"> - Travel**: Being able to speak the local language can greatly improve the travel experience, from navigating transportation to interacting with locals and understanding cultural nuances. 2. Cultural Enrichment**: - Supporting Information**: - Cultural Insight**: Language learners often gain a deeper understanding of cultural idioms, humor, and connotations that are not easily translated, providing a richer cultural experience. - Cultural Participation**: Knowing a foreign language allows individuals to participate in cultural events, festivals, and traditions that would otherwise be inaccessible. - Heritage and Identity**: For people with a family background from a different linguistic community, learning that language can help strengthen personal identity and familial bonds. 3. Educational and Cognitive Benefits**: - Supporting Information**: - Cognitive Development**: Bilingualism has been linked to enhanced executive functions, such as attention control, problem-solving, and the ability to switch between tasks, according to a body of cognitive research. - Academic Performance**: Proficiency in a foreign language can facilitate learning in other academic areas, as language learning improves overall reading abilities and comprehension. - Brain Health**: Some studies suggest that bilingualism may delay the onset of Alzheimer's disease and dementia, possibly due to increased brain reserve and the mental exercise involved in managing two or more languages. <p>Each viewpoint is backed by empirical evidence and theoretical frameworks that demonstrate the multifaceted benefits of learning a foreign language, which go beyond simple communication needs and extend into personal growth, intercultural understanding, and cognitive resilience.</p> <p>The best conclusion based on the provided supporting information is that learning a foreign language is not solely for the purpose of traveling to or working in a foreign country. While these are certainly significant and practical reasons, the benefits of language learning are far more extensive and multifaceted.</p> <ol style="list-style-type: none"> 1. Economic Benefits** and Immigrant Integration**: illustrate that language proficiency can lead to better job prospects, higher earnings, and greater social inclusion, regardless of whether one is traveling or working abroad. 2. Cultural Enrichment** through Cultural Insight**, Cultural Participation**, and the connection to Heritage and Identity**: emphasizes that language learning enables individuals to engage with and appreciate the cultural nuances, participate more fully in cultural practices, and connect with their ancestral roots, enhancing personal growth and intercultural understanding. 3. Educational and Cognitive Benefits** such as Cognitive Development**, Academic Performance**, and Brain Health**: show that learning a foreign language contributes to cognitive skills, academic achievement, and long-term brain health, benefits that extend well beyond the immediate utility for travel or work. <p>In conclusion, while travel and work are</p>
Phase three: Verifying and finishing	INFORMATION needed provided by ChatGPT:	<p>1. Practical Necessity**:</p> <ul style="list-style-type: none"> - Supporting Information**: - Economic Benefits**: Bilingual employees may earn more per hour than those who speak only one language. - Travel**: Being able to speak the local language can greatly improve the travel experience, from navigating transportation to interacting with locals and understanding cultural nuances. <p>2. Cultural Enrichment**:</p> <ul style="list-style-type: none"> - Supporting Information**: - Cultural Insight**: Language learners often gain a deeper understanding of cultural idioms, humor, and connotations that are not easily translated, providing a richer cultural experience. - Heritage and Identity**: For people with a family background from a different linguistic community, learning that language can help strengthen personal identity and familial bonds. <p>3. Educational and Cognitive Benefits**:</p> <ul style="list-style-type: none"> - Supporting Information**: - Cognitive Development**: Bilingualism has been linked to enhanced executive functions, such as attention control, problem-solving, and the ability to switch between tasks, according to a body of cognitive research. - Academic Performance**: Proficiency in a foreign language can facilitate learning in other 	<p>Based on the following supporting INFORMATION, what is the best conclusion regarding whether traveling to or working in a foreign country is the only reason to learn a foreign language and is there any other reason someone should learn a foreign language (QUESTION)?</p> <p>The best conclusion based on the provided supporting information is that learning a foreign language is not solely for the purpose of traveling to or working in a foreign country. While these are certainly significant and practical reasons, the benefits of language learning are far more extensive and multifaceted.</p> <ol style="list-style-type: none"> 1. Economic Benefits** and Immigrant Integration**: illustrate that language proficiency can lead to better job prospects, higher earnings, and greater social inclusion, regardless of whether one is traveling or working abroad. 2. Cultural Enrichment** through Cultural Insight**, Cultural Participation**, and the connection to Heritage and Identity**: emphasizes that language learning enables individuals to engage with and appreciate the cultural nuances, participate more fully in cultural practices, and connect with their ancestral roots, enhancing personal growth and intercultural understanding. 3. Educational and Cognitive Benefits** such as Cognitive Development**, Academic Performance**, and Brain Health**: show that learning a foreign language contributes to cognitive skills, academic achievement, and long-term brain health, benefits that extend well beyond the immediate utility for travel or work. <p>In conclusion, while travel and work are</p>

(continued on next page)

(continued)

Phase One: Shaping and generating	Notes	ChatGPT prompt	Answers
	academic areas, as language learning improves overall reading abilities and comprehension.		compelling reasons to learn a foreign language, they represent just a fraction of the myriad advantages that come with language proficiency. The broad scope of benefits, including economic, social, cultural, educational, and cognitive, demonstrates that language learning is a valuable endeavor for personal and professional development in a globally interconnected world. The interpretation and conclusion drawn from the supporting information suggest several significant implications for individuals, educational institutions, policymakers, and society at large:
	Summarize the INTERPRETATION AND INFERENCE provided by ChatGPT, and take notes of the implications provided: 1. **Educational Policy and Curriculum Development**: - Educational systems should prioritize foreign language education not only as a practical skill but also as a critical component of a well-rounded education that offers cognitive and cultural benefits. **Personal and Professional Advancement**: - Individuals are encouraged to view language learning as an investment in their personal and professional development, leading to increased job opportunities, cultural engagement, and potential health benefits.	What are the most significant implications can be provided based on the above INTERPRETATION AND INFERENCE ?	1. **Educational Policy and Curriculum Development**: - Educational systems should prioritize foreign language education not only as a practical skill but also as a critical component of a well-rounded education that offers cognitive and cultural benefits. 2. **Workforce and Economic Development**: - Employers and workforce development programs may consider incentivizing foreign language learning to cultivate a more versatile and competitive workforce, recognizing the economic advantages and higher earning potential for multilingual employees. 3. **Cultural Policy and Social Cohesion**: - Governments and community organizations might implement policies and programs that encourage language learning as a means of promoting cultural diversity, fostering social cohesion, and supporting immigrant communities. 4. **Personal and Professional Advancement**: - Individuals are encouraged to view language learning as an investment in their personal and professional development, leading to increased job opportunities, cultural engagement, and potential health benefits. These implications highlight the importance of recognizing and promoting the comprehensive benefits of foreign language learning beyond immediate practical needs, underscoring its value for individual growth, social integration, and the cultivation of a culturally and cognitively rich society.

Appendix 2. AI-assisted Peer-review Checklist

No.	CT standards	Explanations	ChatGPT Prompt
1	Clarity	Understandable, the meaning can be grasped; to free from confusion or ambiguity, to remove obscurities.	Do you think my following thesis writing is clear to understand? Could you give some examples of why it is not clear?
2	Accuracy	Free from errors, mistakes, or distortions; true, correct.	Do you think my following thesis writing is free from errors, mistakes, or distortions? If not, could you point out the incorrect descriptions?
3	Precision	Exact to the necessary level of detail, specific.	Do you think my following thesis writing is precise and specific? If not, please help me find where cannot reflect enough precision.
4	Relevance	Bearing upon or relating to the matter at hand; implies a close logical relationship with, and importance to, the matter under consideration.	Regarding the “relevance” standard, do you think my following thesis can satisfy this criterion? If not, please show why I’m not relevant to the opinion “Do you agree or disagree with this opinion that language die is not important because life will be easier if there are fewer languages in the world?”
5	Depth	Containing complexities and multiple interrelationships, implies thoroughness in thinking through the many variables in the situation, context, idea, or question.	Regarding the “depth” standard, do you think my following thesis satisfies this criterion, containing complexities and multiple interrelationships that imply thorough thinking? If not, please tell the reasons with examples.
6	Breadth	Encompassing multiple viewpoints, comprehensive in view, wide-ranging, and broadminded in perspective.	Regarding the “breadth” standard, do you think my following thesis satisfies this criterion, encompassing multiple viewpoints, comprehensive in view, wide-ranging, and broadminded in perspective? If not, please tell the reasons with examples.

(continued on next page)

(continued)

No.	CT standards	Explanations	ChatGPT Prompt
7	Logic	The parts make sense together, no contradictions; in keeping with the principles of sound judgment and reasonability.	Regarding the “logic” standard, do you think my following thesis satisfies this criterion, in keeping with the principles of sound judgment and reasonability? If not, please tell the reasons with examples.
8	Significance	Having importance, being of consequence; having considerable or substantial meaning.	Among the topic “Do you agree or disagree with this opinion that language die is not important because life will be easier if there are fewer languages in the world?”, do you think my following thesis satisfies this criterion “significance: having importance, being of consequence, having considerable or substantial meaning”? If not, please tell the reasons with examples.
9	Fairness	Free from bias, dishonesty, favoritism, selfish-interest, deception, or injustice.	Does my following thesis free from bias, dishonesty, favoritism, selfish-interest, deception, or injustice? If not, please tell the reasons with examples.

Appendix 3. Questionnaire

Questionnaire on University Students' Acceptance of Using ChatGPT for CT-Oriented EFL Writing

Dear Sir/Madam:

Thank you for taking the time to complete this questionnaire. The information you provide will be used for academic research only and will be kept confidential. The purpose of this questionnaire is to gain a better understanding of students' perceptions of using ChatGPT for critical thinking (CT)-oriented EFL writing. There will be 32 questions in total, and please answer truthfully based on your experience. Thank you for your participation.

By clicking “I agree” below you are indicating that you are at least 18 years old, have read and understood this consent form, and agree to participate in this research study.

- Agree
- Disagree

Please print a copy of this page for your records.

Section A: Demographic profile

1. Gender:

- a. Male
- b. Female

2. Major:

- a. English major
- b. Non-English major

3. Have you used ChatGPT for CT-oriented EFL writing before?

- a. Yes
- b. No

Section B: Participants' acceptance of ChatGPT for CT-oriented EFL writing training

For each statement, please choose the number that best matches a description of yourself when you use ChatGPT for CT-oriented EFL writing practice (1. Strongly disagree 2. Disagree 3. Neutral 4. Agree 5. Strongly agree).

Note: PE=Performance Expectancy, EE=Effort Expectancy, HM=Hedonic Motivation, HB=Habit, TR=Trust, BI=Behavioral intention

Performance Expectancy (PE)

PE1. I would find ChatGPT useful in improving my CT-oriented EFL writing skills.

1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree

PE2. Using ChatGPT enhances my CT-oriented EFL writing motivation.

1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree

PE3. Using ChatGPT might increase my productivity in CT-oriented EFL writing.

1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree

PE4. Using ChatGPT helps me to accomplish my EFL writing assignments more quickly and with high quality.

1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree

Effort Expectancy (EE)

EE1. It is easy for me to learn how to use ChatGPT for CT-oriented EFL writing practices as a student.

1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree
- EE2. My interaction with ChatGPT is clear and understandable in the process of using it to practice CT-oriented EFL writing.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree
- EE3. Using ChatGPT to assist CT-oriented EFL writing is easy for me.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree
- EE4. Developing CT-oriented EFL writing skills using ChatGPT is easy to achieve.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree

Hedonic Motivation (HM)

- HM1. I feel fun using ChatGPT to develop CT-oriented EFL writing skills.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree
- HM2. I enjoy using ChatGPT to practice CT-oriented EFL writing.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree
- HM3. Using ChatGPT for CT-oriented EFL writing training makes me feel relaxed and happy.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree
- HM4. Using ChatGPT makes my CT-oriented EFL writing practice enjoyable and interesting.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree

Habit (HA)

- HB1. Using ChatGPT to train CT-oriented EFL writing skills has become my habit.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree
- HB2. I feel addicted to using ChatGPT to practice CT-oriented EFL writing.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree
- HB3. I frequently use ChatGPT to assist my CT-oriented EFL writing practices.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree
- HB4. I always prefer to consult ChatGPT first when critical thinking is needed in EFL writing.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree

Trust

- TR1. I believe that ChatGPT is trustworthy.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree
- TR2. I trust ChatGPT when I use it to develop CT-oriented EFL writing skills.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree
- TR3. I do not doubt the content that ChatGPT provides when I use it to practice CT-oriented EFL writing.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree
- TR4. I believe using ChatGPT for cultivating CT-oriented writing will not put me at risk.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree

Behavioral Intention (BI)

- BI1. I intend to continue using ChatGPT as a scaffolding tool in future CT-oriented EFL writing activities.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree
- BI2. I will use ChatGPT to develop CT-oriented EFL writing skills in the future.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree
- BI3. I will always try to use ChatGPT for CT-oriented EFL writing practice.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree
- BI4. I strongly recommend my classmates use ChatGPT to cultivate critical thinking ability in EFL writing.
1. Strongly Disagree2. Disagree3. Neutral4. Agree5. Strongly Agree

Appendix 4. Interview Questions

- 1) Do you think CT-oriented writing supported by ChatGPT is useful in improving your thinking during English writing practices?
- 2) Do you enjoy CT-oriented writing practices supported by ChatGPT compared to traditional CT-oriented writing practices?
- 3) How do you think your interaction with CT-oriented writing practices supported by ChatGPT?
- 4) What is the role of ChatGPT in your brainstorming compared with the traditional CT-oriented approach by peers?
- 5) What is the role of ChatGPT in your peer review compared with traditional CT-oriented peer-review interaction? Which one is more trustworthy?
- 6) Are you willing to try to use AI tools such as ChatGPT for CT-oriented writing in the future? Any suggestions?
- 7) Do you think AI-assisted CT-oriented writing will replace human writers in the future? Why?

References

- Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2014). Strategies for teaching students to think critically: a meta-analysis. *Review of Educational Research*, 85(2), 275–314. <https://doi.org/10.3102/0034654314551063>
- Ahangari, S., & Sepehran, H. (2014). The Effect of Intertextuality on Iranian EFL Learners' Critical Writing. *Iranian Journal of Language Teaching Research*, 2(1), 85–98. (<https://files.eric.ed.gov/fulltext/EJ1127399.pdf>).
- Akbari, M., Seifoori, Z., & Ahour, T. (2018). Enhancing comprehension and production of argumentation through critical thinking awareness-raising. *Lingue - Rivista di lingue e Culture moderne*, 16(2), 111–130. <https://doi.org/10.7358/ling-2017-002-self>
- Al-Zahrani, A. M. (2023). The impact of generative AI tools on researchers and research: Implications for academia in higher education. *Innovations in Education and Teaching International*, 61(5), 1029–1043. <https://doi.org/10.1080/14703297.2023.2271445>
- Aldossary, A. S., Aljindi, A. A., & Alamri, J. M. (2024). The role of generative AI in education: Perceptions of Saudi students. *Contemporary Educational Technology*, 16 (4). <https://doi.org/10.30935/cedtech/15496>
- Algaraady, J., & Mahyoob, M. (2023). ChatGPT's capabilities in spotting and analyzing writing errors experienced by EFL learners. *Arab World English Journal*, 9(9), 3–17. <https://doi.org/10.24093/awej/call9.1>
- Almassaad, A., Alajlan, H., & Aleibaikan, R. (2024). Student perceptions of generative artificial intelligence: Investigating utilization, benefits, and challenges in higher education. *Systems*, 12(10). <https://doi.org/10.3390/systems12100385>
- Anderson, R. C., Chaparro, E. A., Smolkowski, K., & Cameron, R. (2023). Visual thinking and argumentative writing: A social-cognitive pairing for student writing development. *Assessing Writing*, 55. <https://doi.org/10.1016/j.awsv.2023.100694>
- Avni, S., & Finn, H. B. (2020). Meeting the needs of english language learners in co-requisite courses at community college. *Community College Journal of Research and Practice*, 45(8), 560–574. <https://doi.org/10.1080/10668926.2020.1727383>
- Azevedo, E., Ramos, A., Araújo, R., Valério, C., & Mesquita, I. (2023). How can tailored questions foster reflection in preservice teachers? A year-long action research study. *European Physical Education Review*, 30(3), 458–474. <https://doi.org/10.1177/1356336x231217405>
- Bai, B., Wang, J., & Zhou, H. (2021). An intervention study to improve primary school students' self-regulated strategy use in English writing through e-learning in Hong Kong. *Computer Assisted Language Learning*, 35(9), 2265–2290. <https://doi.org/10.1080/09588221.2020.1871030>
- Bai, C., Chen, Z., & Liu, X. (2022). Case Study Research in Tesla (China) Marketing Strategy Application During Covid-19 [JONKOPING University]. JONKOPING.
- Banishashem, 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(1). <https://doi.org/10.1186/s41239-024-00455-4>
- Barnawi, O. (2011). Finding a place for critical thinking and self-voice in college English as a foreign language writing classrooms. *English Language Teaching*, 4(2), 190–197. <https://doi.org/10.5539/elt.v4n2p190>
- Bean, J.C., & Melzer, D. (2021). Engaging Ideas: The Professor's Guide to Integrating Writing, Critical Thinking, and Active Learning in the Classroom (3rd ed.). Jossey-Bass.
- Beck, S. W., & Levine, S. R. (2023). Backtalk: ChatGPT: A powerful technology tool for writing instruction. *Phi Delta Kappan*, 105(1), 66–67. (<https://nyuscholars.nyu.edu/en/publications/backtalk-chatgpt-a-powerful-technology-tool-for-writing-instruction>).
- Boudouaia, A., Mouas, S., & Kouider, B. (2024). A study on ChatGPT-4 as an innovative approach to enhancing English as a foreign language writing learning. *Journal of Educational Computing Research*, 62(6), 1509–1537. <https://doi.org/10.1177/07356331241247465>
- Bryman, A. (2016). Social Research Methods. Oxford University Press.
- Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00411-8>
- Chason, L., Loyet, D., Sorenson, L., & Stoops, A. (2017). An approach for embedding critical thinking in second language paragraph writing. *TESOL Journal*, 8(3), 582–612. <https://doi.org/10.1002/tesj.288>
- Cheung, R., & Vogel, D. (2013). Predicting user acceptance of collaborative technologies: An extension of the technology acceptance model for e-learning. *Computers Education*, 63, 160–175. <https://doi.org/10.1016/j.compedu.2012.12.003>
- Chimmalee, B., & Anupan, A. (2023). The effects of using flipped cloud learning with advancing mathematical thinking approaches on undergraduate students' mathematical critical thinking. *Turkish Online Journal of Distance Education*, 24(3), 240–260. <https://doi.org/10.17718/tojde.1125894>
- Condon, W., & Kelly-Riley, D. (2004). Assessing and teaching what we value: The relationship between college-level writing and critical thinking abilities. *Assessing Writing*, 9(1), 56–75. <https://doi.org/10.1016/j.asw.2004.01.003>
- Creswell, J.W., & Creswell, J.D. (2023). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (6th ed.). SAGE.
- DaCosta, B., & Kinsell, C. (2024). Investigating media selection through ChatGPT: An exploratory study on generative artificial intelligence in the aid of instructional design. *Open Journal of Social Sciences*, 12(04), 187–227. <https://doi.org/10.4236/jss.2024.124041>
- Dahri, N. A., Yahaya, N., Al-Rahmi, W. M., Aldraiweesh, A., Alturki, U., Almutairy, S., & Soomro, R. B. (2024). Investigating AI-based academic support acceptance and its impact on students' performance in Malaysian and Pakistani higher education institutions. *Education and Information Technologies*, 29(14), 18695–18744. <https://doi.org/10.1007/s10639-024-12599-x>
- Dai, Y., Liu, A., & Lim, C. P. (2023). Reconceptualizing ChatGPT and generative AI as a student-driven innovation in higher education. *Procedia CIRP*, 119, 84–90. <https://doi.org/10.1016/j.procir.2023.05.002>
- Dong, Y. (2017a). Critical Thinking in Second Language Writing: Concept, Theory, Teaching and Assessment. Foreign Language Teaching and Research Press.
- Dong, Y. (2017b). A rating scale for assessing critical thinking in L2 writing: Exploration and practice. *Foreign Language Education in China (Quarterly)*, 10(1), 23–30.
- Dong, Y. (2017c). Teaching and assessing critical thinking in second language writing: An infusion approach. *Chinese Journal of Applied Linguistics*, 40(4), 431–451. <https://doi.org/10.1515/cjal-2017-0025>
- Ebad, S., & Rahimi, M. (2018). An exploration into the impact of WebQuest-based classroom on EFL learners' critical thinking and academic writing skills: a mixed-methods study. *Computer Assisted Language Learning*, 31(5–6), 617–651. <https://doi.org/10.1080/09588221.2018.1449757>
- Elkhodr, M., Gide, E., Wu, R., & Darwish, O. (2023). ICT students' perceptions towards ChatGPT: An experimental reflective lab analysis. *STEM Education*, 3(2), 70–88. <https://doi.org/10.3934/steme.2023006>
- Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational Researcher*, 18(3), 4–10. <https://doi.org/10.3102/0013189x018003004>
- Esmail Nejad, M., Izadpanah, S., Namaziandost, E., & Rahbar, B. (2022). The mediating role of critical thinking abilities in the relationship between English as a foreign language learners' writing performance and their language learning strategies. *Frontiers in Psychology*, 13, 1–11. <https://doi.org/10.3389/fpsyg.2022.746445>
- Etemadzadeh, A., Seifi, S., & Far, H. R. (2013). The role of questioning technique in developing thinking skills: The ongoing effect on writing skill. *Procedia - Social and Behavioral Sciences*, 70, 1024–1031. <https://doi.org/10.1016/j.sbspro.2013.01.154>
- Facione, P.A. (1991). *Using the California Critical Thinking Skills Test in Research, Evaluation, and Assessment*. C. A. Press. (<https://files.eric.ed.gov/fulltext/ED337498.pdf>).
- Facione, P.A. (2020). *Critical Thinking: What It is and Why it Counts*. The California Academic Press LLC. (<https://www.insightassessment.com/wp-content/uploads/1/a/pdf/whatwhy.pdf>).
- Fahim, M., & Masouleh, N. S. (2012). Critical thinking in higher education: A pedagogical look. *Theory and Practice in Language Studies*, 2(7). <https://doi.org/10.4304/tpls.2.7.1370-1375>
- Faisal, E. (2024). Unlock the potential for Saudi Arabian higher education: a systematic review of the benefits of ChatGPT. *Frontiers in Education*, 9, 1–12. <https://doi.org/10.3389/feduc.2024.1325601Faisal>
- Fraenkel, J.R., Wallen, N.E., & Hyun, H.H. (2023). How to Design and Evaluate Research in Education (11th ed.). McGraw Hill.
- Fraser, M., Wotring, A., Green, C. A., & Eady, M. J. (2022). Designing a framework to improve critical reflection writing in teacher education using action research. *Educational Action Research*, 32(1), 43–59. <https://doi.org/10.1080/09650792.2022.2038226>
- Gao, X., Asmawi, A., & Samuel, M. (2017). Critical peer feedback for business English writing through Qzone blogs: A mechanism among Chinese undergraduates. *GEMA Online® Journal of Language Studies*, 17(1), 39–54. <https://doi.org/10.17576/gema-2017-1701-03>
- Gao, X., & Wen, Q. F. (2017). Effects of critical thinking skills via linguistic factors on L2 writing performance. *Foreign Language Learning Theory and Practice*, 4, 44–50.
- Glodjo, T. (2018). Critical Pedagogy. In *The TESOL Encyclopedia of English Language Teaching* (pp. 1–6). John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118784235.eelt0914>
- Goda, Y., Yamada, M., Matsukawa, H., Hata, K., & Yasunami, S. (2014). Conversation with a chatbot before an online EFL group discussion and the effects on critical thinking. *Information and Systems in Education*, 13(1), 1–7. <https://doi.org/10.12937/ejsise.13.1>
- Greiner, C., Peisl, T. C., Höpfel, F., & Beese, O. (2023). Acceptance of AI in semi-structured decision-making situations applying the four-sides model of communication—An empirical analysis focused on higher education. *Education Sciences*, 13(9). <https://doi.org/10.3390/educsci13090865>
- Guo, K., & Wan, D. (2023). To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. *Education and Information Technologies*, 29, 1. <https://doi.org/10.1007/s10639-023-12146-0>
- Guo, K., Wang, J., & Chu, S. K. W. (2022). Using chatbots to scaffold EFL students' argumentative writing. *Assessing Writing*, 54, 1–6. <https://doi.org/10.1016/j.awsw.2022.100666>
- Hair, J.F., Black, W.C., Baibin, B.J., & Anderson, R.E. (2018). Multivariate data analysis (8th ed.). Cengage Learning EMEA.
- Halpern, D.F. (2014). Thought and Knowledge: an introduction to critical thinking (5th ed.). Psychology Press.
- Hastuti, D. P., Kristina, D., & Setyaningsih, E. (2022). Developing thinking reading Classrooms: Insight from pre-service EFL teachers' lesson plans and classroom instructions. *Register Journal*, 15(2), 201–221. <https://doi.org/10.18326/rjt.v15i2.201-221>
- Healey, B. (2024). Teaching an embodied thinking process for narrative writing 'with grammar in mind': An analysis of writing conferences and teacher talk. *Literacy Research and Instruction*, 64(2), 203–228. <https://doi.org/10.1080/19388071.2024.2345898>
- Huang, S.Z. (2016). An Action Research on Fostering Critical Thinking in EFL Writing through Formative Assessment and Feedback. *Contemporary Educational Research Quarterly*, 24(1), 81–105. <https://doi.org/10.6151/CERQ.2016.2401.03>
- Irgin, P., & Bilki, Z. (2024). Students' perceptions of online peer feedback in process-oriented L2 writing: A qualitative inquiry. *Studies in Educational Evaluation*, 83. <https://doi.org/10.1016/j.stueduc.2024.101403>

- Jubier, M. M., Baharum, H. I., & Said, S. M. (2024). Exploring the predictive value of self-regulation and thinking skills on the writing performance of Iraqi advanced EFL Learners on TOEFL iBT. *The Asia-Pacific Education Researcher*. <https://doi.org/10.1007/s40299-024-00906-6>
- Kaviani, M., & Heidar, D. M. (2020). The effectiveness of critical thinking on enhancing productive skills among Iranian EFL pre-intermediate learners. *Applied Research on English Language*, 9(3), 303–324. <https://doi.org/10.22108/are.2019.118748.1489>
- Lavidas, K., Papadakis, S., Manesis, D., Grigoriadou, A. S., & Gialamas, V. (2022). The effects of social desirability on students' self-reports in two social contexts: Lectures vs. lectures and lab classes. *Information*, 13(10). <https://doi.org/10.3390/info13100491>
- Lavidas, K., Voulgaris, I., Papadakis, S., Athanassopoulos, S., Anastasiou, A., Filippidi, A., & Karacapilidis, N. (2024). Determinants of humanities and social sciences students' intentions to use artificial intelligence applications for academic purposes. *Information*, 15(6). <https://doi.org/10.3390/info15060314>
- Lin, M., & Xiang, X. (2019). Integrating critical thinking into an EFL writing curriculum. *In Thinking Skills and Creativity in Second Language Education*, 95–130. <https://doi.org/10.4324/9781315098920-6>
- Lin, W. (2019). A Study on the development of english literacy and the evaluation of critical thinking. *Journal of Fujian Institute of Education*, 4, 96–99.
- Liu, G., & Ma, C. (2023). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, 1–14. <https://doi.org/10.1080/17501229.2023.2240316>
- Liu, X., & Yao, T. (2019). The Cultivation of college students' critical thinking ability based on task-based cooperative writing. *Journal of Language Teaching and Research*, 10(3), 557–568. <https://doi.org/10.17507/jltr.1003.20>
- Lodge, J. M., Yang, S., Furze, L., & Dawson, P. (2023). It's not like a calculator, so what is the relationship between learners and generative artificial intelligence? *Learning: Research and Practice*, 9(2), 117–124. <https://doi.org/10.1080/23735082.2023.2261106>
- Loncar, M., Schams, W., & Liang, J.-S. (2021). Multiple technologies, multiple sources: Trends and analyses of the literature on technology-mediated feedback for L2 English writing published from 2015–2019. *Computer Assisted Language Learning*, 36 (4), 722–784. <https://doi.org/10.1080/09588221.2021.1943452>
- Lu, D., & Xie, Y. (2019). The effects of a critical thinking oriented instructional pattern in a tertiary EFL argumentative writing course. *Higher Education Research Development*, 38(5), 969–984. <https://doi.org/10.1080/07294360.2019.1607830>
- Lu, D., & Xie, Y.-N. (2022). Critical thinking cultivation in TESOL with ICT tools: A systematic review. *Computer Assisted Language Learning*, 37(1–2), 222–242. <https://doi.org/10.1080/09588221.2022.2033788>
- MacLeod, S., Reynolds, M. G., & Lehmann, H. (2018). The mitigating effect of repeated memory reactivations on forgetting. *NPJ Science Learn*, 3, 9. <https://doi.org/10.1038/s41539-018-0025-x>
- Maican, C. I., Sumedrea, S., Tecau, A., Nichifor, E., Chitu, I. B., Lixandroiu, R., & Bratucu, G. (2023). Factors influencing the behavioural intention to use AI-generated images in business. *Journal of Organizational and End User Computing*, 35 (1), 1–32. <https://doi.org/10.4018/joeuc.330019>
- Mall-Amiri, B., & Sheikhy, F. (2014). The comparative impact of autonomy and critical thinking on EFL learners' writing achievement. *Theory and Practice in Language Studies*, 4(5), 903–916. <https://doi.org/10.4304/tpls.4.5.903-916>
- McLaren, P. (2015). Life in schools: An introduction to critical pedagogy in the foundations of education. Routledge.
- Memon, M. A., Ting, H., Cheah, J.-H., Thurasamy, R., Chuah, F., & Cham, T. H. (2020). Sample size for survey research: Review and recommendations. *Journal of Applied Structural Equation modeling*, 4(2) (i–xx).
- Menon, D., & Shilpa, K. (2023). "Chatting with ChatGPT": Analyzing the factors influencing users' intention to Use the Open AI's ChatGPT using the UTAUT model. *Helijon*, 9(11), Article e20962. <https://doi.org/10.1016/j.helijon.2023.e20962>
- Miri, F., & Azizi, D. B. (2018). The effect of teaching critical thinking on Iranian EFL learners' essay writing. *Theory and Practice in Language Studies*, 8(5), 509–515. <https://doi.org/10.17507/tpls.0805.08>
- Mok, J. (2010). The new role of English language teachers: Developing students' critical thinking in Hong Kong secondary school classrooms. *The Asian EFL Journal Quarterly*, 12(2), 262–287. (<https://www.asian-efl-journal.com/main-editions-new/the-new-role-of-english-language-teachers-developing-students-critical-thinking-in-hong-kong-secondary-school-classrooms/index.htm>).
- Molenaar, I. (2022). The concept of hybrid human-AI regulation: Exemplifying how to support young learners' self-regulated learning. *Computers and Education: Artificial Intelligence*, 3. <https://doi.org/10.1016/j.caei.2022.100070>
- Monchinski, T. (2008). Critical pedagogy and the every day classroom. Springer.
- Murtadho, F. (2021). Metacognitive and critical thinking practices in developing EFL students' argumentative writing skills. *Indonesian Journal of Applied Linguistics*, 10 (3), 656–666. <https://doi.org/10.17509/jjal.v10i3.31752>
- Nielsen, T., Martínez-García, I., & Alastor, E. (2022). Exploring first semester changes in domain-specific critical thinking. *Frontiers in Education*, 7. <https://doi.org/10.3389/feduc.2022.884635>
- Panadero, E., & Lipnevich, A. A. (2022). A review of feedback models and typologies: Towards an integrative model of feedback elements. *Educational Research Review*, 35. <https://doi.org/10.1016/j.edurev.2021.100416>
- Pang, Y. (2022). The role of web-based flipped learning in EFL learners' critical thinking and learner engagement. *Frontiers in Psychology*, 13, 1–12. <https://doi.org/10.3389/fpsyg.2022.1008257>
- Paul, R., & Elder, L. (2012). Critical Thinking Tools for Taking Charge of Your Learning and Your Life (third ed.). Pearson Education.
- Paul, R., & Elder, L. (2021). Critical thinking: Tools for taking charge of your learning and your life. Rowman & Littlefield.
- Pei, Z., Zheng, C., Zhang, M., & Liu, F. (2017). Critical thinking and argumentative writing: Inspecting the association among EFL Learners in China. *English Language Teaching*, 10(10). <https://doi.org/10.5539/elt.v10n10p31>
- Piaget, J. (1970). Genetic epistemology. Columbia University Press.
- Polyportis, A., & Pahos, N. (2024). Understanding students' adoption of the ChatGPT chatbot in higher education: the role of anthropomorphism, trust, design novelty and institutional policy. *Behaviour Information Technology*, 1–22. <https://doi.org/10.1080/0144929X.2024.2317364>
- Rahi, S., Alnaser, F. M., & Abd Ghani, M. (2019). Designing survey research: recommendation for questionnaire development, calculating sample size and selecting research paradigms. *Economic and Social Development: Book of Proceedings*, 1157–1169.
- Rahman, M. S., Sabbir, M. M., Zhang, J., Moral, I. H., & Hossain, G. M. S. (2023). Examining students' intention to use ChatGPT: Does trust matter? *Australasian Journal of Educational Technology*, 39(6), 51–71. <https://doi.org/10.14742/ajet.8956>
- Rasa, T., Lavonen, J., & Laherto, A. (2023). Agency and transformative potential of technology in students' images of the future. *Science Education*, 33(5), 1145–1169. <https://doi.org/10.1007/s11191-023-00432-9>
- Reichardt, C. S. (2019). *Quasi-Experimentation: A guide to design and analysis*. The Guilford Press.
- Ruan, Z. (2005). A metacognitive perspective on the growth of self-regulated EFL student writers. *Reading, Working Papers in Linguistics*, 8, 175–202. (<https://www.reading.ac.uk/internal/appling/wp8/Ruan.pdf>).
- Safari, P. (2020). Constructing an emancipatory learning environment in Iranian English classes through dialogue journal writing as an educational tool. *Education 3-13*, 49 (5), 618–634. <https://doi.org/10.1080/03004279.2020.1757736>
- Salam, U. (2024). The Integration of ChatGPT in English for foreign language course: Elevating AI writing assistant acceptance. *Computers in the Schools*, 1–21. <https://doi.org/10.1080/07380569.2024.2446239>
- Samala, A. D., Sokolova, E. V., Grassini, S., & Rawas, S. (2024). ChatGPT: A bibliometric analysis and visualization of emerging educational trends, challenges, and applications. *International Journal of Evaluation and Research in Education (IJERE)*, 13 (4). <https://doi.org/10.11591/ijere.v13i4.28119>
- Shahani, S., Chalak, A., & Tabrizi, H. H. (2022). Impact of critical thinking instruction through flipped teaching on Iranian EFL learners' listening comprehension. *Turkish Online Journal of Distance Education*, 23(2), 236–251. <https://doi.org/10.17718/tojde.1096565>
- Sharadgah, T., Sa'di, R., & Ahmad, H. (2019). Promoting and assessing EFL college students' critical thinking skills through argumentative essay writing. *Arab World English Journal*, 10(4), 133–150. <https://doi.org/10.24093/awej.vol10no4.11>
- Shida, N., Osman, S., Abdullah, A. H., & Ismail, N. (2018). Critical thinking dispositions among polytechnic students: why does it matter? *International Journal of Engineering Technology*, 7(25), 357–361. <https://doi.org/10.14419/ijet.v7i3.25.17597>
- Stapa, S. H., & Ibahrim, N. I. (2020). The use of edutainment in promoting higher order thinking skills in ESL writing among Malaysian university students. *Arab World English Journal*, 6, 49–66. <https://doi.org/10.24093/awej.call6.4>
- Strzelecki, A. (2023a). Students' acceptance of ChatGPT in higher education: An extended unified theory of acceptance and use of technology. *Innovative Higher Education*, 1–23. <https://doi.org/10.1007/s10755-023-09686-1>
- Strzelecki, A. (2023b). To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interactive Learning Environments*, 1–14. <https://doi.org/10.1080/10494820.2023.2209881>
- Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, 57, 1–11. <https://doi.org/10.1016/j.aw.2023.100752>
- Suhr, D.D. (2006). Exploratory or confirmatory factor analysis? Cary, NC: SAS Institute Inc.
- Sullivan, A. M., Hayes, M. M., Beltran, C. P., Cohen, A. P., Soffler, M., Cooper, S., & Schwartzstein, R. M. (2024). Do we teach critical thinking? A mixed methods study of faculty and student perceptions of teaching and learning critical thinking at three professional schools. *Medical Teacher*, 1–8. <https://doi.org/10.1080/0142159X.2024.2316862>
- Sun, W., & Zou, B. (2022). A study of pre-service EFL teachers' acceptance of online teaching and the influencing factors. *Language Teacher Education and Technology Forum*, 26(2), 38–49. <https://doi.org/10125/73476>
- Susilo, A., Mufanti, R., & Fitriani, A. (2021). Promoting EFL students' critical thinking and self-voicing through CIRC technique in Academic Writing courses. *Studies in English Language and Education*, 8(3), 917–934. <https://doi.org/10.24815/siele.v8i3.21149>
- Susnjak, T., & McIntosh, T. R. (2024). ChatGPT: The end of online exam integrity. *Education Sciences*, 14(6), 656. (<https://www.mdpi.com/2227-7102/14/6/656>).
- T Aldulaijan, A., & M Almalky, S. (2025). The impact of generative AI tools on postgraduate students' learning experiences: New insights into usage patterns. *Journal of Information Technology Education: Research*, 24. <https://doi.org/10.28945/5428>
- Taranakanti, V., & Davidson, C. J. (2023). Metacognition through an Iterative Anatomy AI Chatbot: An innovative playing field for educating the future generation of medical students. *Anatomia*, 2(3), 271–281. <https://doi.org/10.3390/anatomia2030025>
- Thomas, D. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*. *American Journal of Evaluation*, 27, 237–246. <https://doi.org/10.1177/1098214005283748>
- Tsai, C.-Y., Lin, Y.-T., & Brown, I. K. (2024). Impacts of ChatGPT-assisted writing for EFL English majors: Feasibility and challenges. *Education and Information Technologies*, 29 (17), 22427–22445. <https://doi.org/10.1007/s10639-024-12722-y>

- Vargas-Murillo, A. R., de la Asuncion, I. N. M., & de JesusGuevara-Soto, F. (2023). Challenges and opportunities of AI-assisted Learning: A systematic literature review on the impact of ChatGPT Usage in Higher Education. *International Journal of Learning, Teaching and Educational Research*, 22(7), 122–135. <https://doi.org/10.26803/ijler.22.7.7>
- Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 36(1), 157–178. (<https://ssrn.com/abstract=2002388>)
- Vygotsky, L.S. (1986). Thought and language. MIT Press.
- Wade, C. (1995). Using writing to develop and assess critical thinking. *Teaching of Psychology*, 22(1), 24–28. <https://doi.org/10.1207/s15328023top2201>
- Wale, B. D., & Bishaw, K. S. (2020). Effects of using inquiry-based learning on EFL students' critical thinking skills. *Asian-Pacific Journal of Second and Foreign Language Education*, 5(1). <https://doi.org/10.1186/s40862-020-00090-2>
- Wang, Q. (2025). EFL learners' motivation and acceptance of using large language models in English academic writing: An extension of the UTAUT model [Original Research]. *Frontiers in Psychology*, 15. <https://doi.org/10.3389/fpsyg.2024.1514545>
- Welhaf, M. S., Phillips, N. E., Smeekens, B. A., Miyake, A., & Kane, M. J. (2022). Interpolated testing and content pretesting as interventions to reduce task-unrelated thoughts during a video lecture. *Cogn Res Princ Implicit*, 7(1), 26. <https://doi.org/10.1186/s41235-022-00372-y>
- Wen, Q., Sun, M., & Zhang, L. (2018). The development of critical thinking skills of foreign language majors: A longitudinal study (外语专业大学生思辨技能发展趋势跟踪研究). *Foreign Language. World*, 6, 12–19.
- Yan, D. (2023a). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, 28, 13943–13967. <https://doi.org/10.1007/s10639-023-11742-4>
- Yan, D. (2023b). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, 28, 13943–13967. <https://doi.org/10.1007/s10639-023-11742-4>
- Yang, Y. T. C., & Gamble, J. (2013). Effective and practical critical thinking-enhanced EFL instruction. *ELT Journal*, 67(4), 398–412. <https://doi.org/10.1093/elt/cct038>
- Yang, Y. T. C., Gamble, J. H., Hung, Y.-W., & Lin, T.-Y. (2013). An online adaptive learning environment for critical-thinking-infused English literacy instruction. *British Journal of Educational Technology*, 45(4), 723–747. <https://doi.org/10.1111/bjet.12080>
- Yih Mah, B., Rahim, S. A., Marimuthu, R., & Chone Liaw, S. (2021). The Effectiveness of WeCWI-enabled web-based instructional tool to improve writing performance and critical thinking level among undergraduate students. *IBIMA Business Review*, 1–12. <https://doi.org/10.5171/2021.932485>
- Yin, S., Guo, K., & Li, D. (2024). To assign a topic or not: examining critical thinking in English as a foreign language speaking. *RELC Journal*. <https://doi.org/10.1177/00336882231222651>
- Yin, X. (2024). A Critical Thinking-oriented Approach in English as a Foreign Language Writing: Using ChatGPT as an Ideal Peer Australian Journal of Applied Linguistics, Waseda University, Tokyo, Japan.
- Yin, X., Bin Mohd Saad, M. R., & Binti Abdul Halim, H. (2024). Technology-enhanced social learning (TSL) to foster critical thinking dispositions and thinking in writing. *Cogent Education*, 11(1), Article 2341584. <https://doi.org/10.1080/2331186x.2024.2341584>
- Yin, X., & Saad, M.R.B.M. (2023). Coupling Coordination Degree Model between the Two Connected Systems: Critical Thinking and EFL Writing. iGREduc 2023: Critical Thinking and Problem-Solving: Empowering the Next Generation of Learners, Malaysia.
- 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, 1–17. <https://doi.org/10.1016/j.tsc.2023.101363>
- Zou, D., Xie, H., & Wang, F. L. (2022). Effects of technology enhanced peer, teacher and self-feedback on students' collaborative writing, critical thinking tendency and engagement in learning. *Journal of Computing in Higher Education*, 35(1), 166–185. <https://doi.org/10.1007/s12528-022-09337-y>
- Zou, M., & Huang, L. (2023). To use or not to use? Understanding doctoral students' acceptance of ChatGPT in writing through technology acceptance model. *Frontiers in Psychology*, 14, 1–9. <https://doi.org/10.3389/fpsyg.2023.1259531>
- Zuckerman, M., Flood, R., Tan, R. J. B., Kelp, N., Ecker, D. J., Menke, J., & Lockspeiser, T. (2023). ChatGPT for assessment writing. *Medical Teacher*, 45(11), 1224–1227. <https://doi.org/10.1080/0142159X.2023.2249239>