Computers and Education: Artificial Intelligence 9 (2025) 100489

Contents lists available at ScienceDirect

Computers and Education: Artificial Intelligence

journal homepage: www.sciencedirect.com/journal/computers-and-education-artificial-intelligence

Engagements with GPT responses and learner prompts in ChatGPT-based learning of English argumentative writing logic and their impacts

Ruofei Zhang [a], Di Zou [b,*], Haoran Xie [c], Fu Lee Wang [d]

[a] Centre for English and Additional Languages, Lingnan University, Hong Kong SAR China
[b] Department of English and Communication, The Hong Kong Polytechnic University, Hong Kong SAR China
[c] School of Data Science, Lingnan University, Hong Kong SAR China
[d] School of Science and Technology, Hong Kong Metropolitan University, Hong Kong SAR China

ARTICLE INFO

Keywords:
AI
Artificial intelligence
Chatbot
ChatGPT
English argumentative writing
Logic learning
Second language writing

ABSTRACT

ChatGPT can be defined as a chatbot powered by OpenAI's GPT language models, which has shown promise in improving English-as-a-foreign-language (EFL) writing knowledge and skills. However, its application to developing EFL argumentative writing logic remains largely unexplored, despite the importance of this area. Moreover, existing studies have highlighted deep learner engagement with ChatGPT-based learning but have not examined how engagement varies between two key components of this learning method: GPT responses (GPT's messages to learners) and learner prompts (learners' messages to GPT). To better understand the mechanisms and efficacy of ChatGPT-based learning for EFL argumentative writing, we developed a discipline-specific GPT-4-powered chatbot for learning English argumentative writing logic. Forty-two Chinese university students used the tool for 45–75 min. Learner engagement in GPT responses and learner prompts was assessed via eye movements on corresponding interface areas of ChatGPT recorded by a Tobii eye-tracker. Their learning outcomes were assessed via pre-post-delayed tests and pre-post writing tasks. Semi-structured interviews were also administered. Our findings revealed that learners engaged with GPT responses frequently but for short durations, and with learner prompts infrequently but for longer durations. Engagement in GPT responses appears to facilitate logic knowledge development, whereas engagement in learner prompts may be associated with challenges in developing writing logic. Based on the results, we explored the factors influencing the patterns and impacts of learner engagement with ChatGPT-based learning of English argumentative writing logic and offered implications for future implementation of this learning method.

1. Introduction

ChatGPT is a chatbot powered by OpenAI's GPT models, widely used in education to facilitate skill and knowledge development (Baidoo-Anu & Ansah, 2023). In ChatGPT-based learning, learners engage in natural-language interactions with ChatGPT by submitting questions or requests (learner prompts) and receiving replies or feedback (GPT responses) (Kohnke et al., 2023). This learning approach has been increasingly studied and applied across disciplines, including language education (Kasneci et al., 2023). By examining this tool from the language pedagogical perspective, researchers have identified ChatGPT affordances for generating unlimited, diverse instructional materials, providing immediate personalised feedback, and facilitating authentic language use, arguing for its potential in assisting language development, including English-as-a-foreign-language (EFL) writing, despite ongoing concerns regarding its reliability, creativity, etc. (Baidoo-Anu & Ansah, 2023; Kohnke et al., 2023). Supporting this potential, Teng (2024a) reviewed 20 studies on ChatGPT-based EFL writing education and reported predominantly positive outcomes.

The effectiveness of ChatGPT-based learning of EFL writing may be partly attributed to its capacity to foster learner engagement. Learner engagement refers to students' proactive involvement in instructional materials and pedagogical components (Ge & Ifenthaler, 2018). In ChatGPT-based contexts, learners engage with two essential interactive components: GPT responses and learner prompts, which manifests across multiple dimensions (Ge & Ifenthaler, 2018). Behaviourally, learner engagement reflects on observable actions (Bygate & Samuda, 2009), such as attentively reading GPT responses and crafting

* Corresponding author. HHB914, 11 Yuk Choi Road, Hung Hom, Kowloon, Hong Kong SAR China.
E-mail addresses: ruofeizh@gmail.com (R. Zhang), dizoudaisy@gmail.com (D. Zou), hrxie@ln.edu.hk (H. Xie), pwang@hkmu.edu.hk (F.L. Wang).

https://doi.org/10.1016/j.caeai.2025.100489
Received 19 September 2025; Received in revised form 17 October 2025; Accepted 17 October 2025
Available online 22 October 2025
2666-920X/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

well-considered prompts. Cognitively, it manifests in focused attention and strategic use of learning strategies (Philp & Duchesne, 2016), including internalising new information from GPT responses and elaborating knowledge points when formulating prompts (Lestari, 2024). Emotionally, it encompasses learners' emotional reactions to various aspects of learning tasks (Hiver et al., 2021), including emotional connections with ChatGPT and enjoyment in the learning environment. Engagement across these dimensions with both interactive components constitutes the dynamic dialogue between learners and ChatGPT, forming the core of the learning process (Kohnke et al., 2023). A synthesis study by Kasneci et al. (2023) argued that ChatGPT-based learning could foster deep learner engagement through varied instructional materials and personalised learning experiences. Teng (2024a) also supported the notion that ChatGPT effectively engages learners in EFL writing tasks. Research showed that deeper learner engagement typically leads to better language learning outcomes (Hiver et al., 2021; Philp & Duchesne, 2016). By the logic, learners may effectively develop EFL writing knowledge and skills through deep engagement in GPT responses and learner prompts in ChatGPT-based learning contexts.

For recent years, a growing body of research has examined ChatGPT-based learning of general EFL writing; however, studies specifically focusing on its application to developing English argumentative writing logic remain scarce. English argumentative writing logic refers to inferencing and reasoning that construct and support claims, organise ideas, and address counterarguments in argumentative writings, which is critical for the strength, coherence, cohesion, and overall quality of English argumentative writing (Murray, 2012). Research has indicated that EFL learners, even at upper-intermediate to advanced proficiency levels, frequently struggle with logic in English argumentative writing due to insufficient logical knowledge and skills. For example, Lismay (2020) examined English argumentative essays by 28 EFL university students and found that over 76 % exhibited reasoning problems and poor logical quality. Similarly, El Khoiri and Widiati (2017), analysing English argumentative essays from 40 EFL university students, found that most students faced significant challenges in logic knowledge and committed frequent reasoning errors that substantially undermined their writing quality. Given these findings, scholars have advocated for greater attention to developing English argumentative writing logic among EFL learners (El Khoiri & Widiati, 2017; Murray, 2012). Considering the demonstrated effectiveness of ChatGPT-based learning in improving EFL writing (Teng, 2024a), this learning method appears promising for supporting EFL students' development of English argumentative writing logic.

Furthermore, most previous studies have examined learner engagement with ChatGPT-based language learning holistically (e.g., Woo et al., 2024; Zhang et al., 2025a), overlooking potential variations in engagement across its interactive components. Jiang and Hyland (2024) investigated linguistic differences between GPT responses and learner prompts, suggesting that GPT responses tend to feature abstract descriptions, while learner prompts exhibit higher levels of persuasive argumentation. Lestari (2024) further noted that learners engage with GPT responses primarily to process and internalise new knowledge and tasks, whereas they generate learner prompts to elaborate personal opinions, ask questions, and request further instructional materials or tasks. The differences in linguistic characteristics and learner utilisation between GPT responses and learner prompts may lead to variations in learner engagement with these components. Given that a comprehensive understanding of learner engagement is crucial for effectively conceptualising, designing, and implementing pedagogical tasks and materials (Philp & Duchesne, 2016), examining learner engagement separately in GPT responses and learner prompts can offer valuable insights into the mechanisms underlying ChatGPT-based learning. Such research can also help identify effective ChatGPT-based learning designs and learner behaviours, thus informing future implementations of this method.

Additionally, previous research on learner engagement with ChatGPT-based learning has primarily relied on self-report measures, such as questionnaires and interviews (e.g., Teng, 2024b). However, learner engagement can be objectively measured through observable behaviours, notably eye-tracking metrics like total fixation count and average fixation duration (Latif, 2019): Total fixation count is the total number of relatively stable eye movements, and average fixation duration is the average length of these stable gazes (Obaidellah et al., 2018). Researchers suggest that higher total fixation counts and longer average fixation durations on specific learning materials or pedagogical components typically reflect deeper learner engagement (Que et al., 2023; Wang et al., 2021). Therefore, learner engagement in ChatGPT-based learning can be objectively measured—separately for GPT responses and learner prompts—by collecting and analysing total fixation count and average fixation duration within the respective areas. Considering the multidimensional nature of learner engagement (Philp & Duchesne, 2016), integrating these objective metrics with self-report data holds promise for yielding a more comprehensive and robust understanding of learner engagement in ChatGPT-based learning.

To address these research gaps, we developed a GPT-4-powered chatbot specifically to support students' development of English argumentative writing logic knowledge and skills, and required 42 EFL learners to use it for approximately 45–75 min. Using a triangulation mixed-method approach, we collected and analysed data from learners' total fixation counts, average fixation durations, pre- and post-test results, pre- and post-intervention English argumentative writing tasks, and semi-structured interviews, so as to explore EFL learner engagement in GPT responses and learner prompts, and examine how this engagement influenced their learning outcomes in terms of logic knowledge and argumentative writing logic quality. Through this comprehensive exploration, our study aims to deepen understanding of the mechanisms underlying ChatGPT-based learning of English argumentative writing logic and generate insights to inform future research and practical applications across educational domains. The study was guided by two research questions:

RQ1: How do EFL learners engage with GPT responses and learner prompts in ChatGPT-based learning of English argumentative writing logic? Why?
RQ2: How do EFL learners' engagement in GPT responses and learner prompts in ChatGPT-based learning of English argumentative writing logic influence their logic knowledge and logic in English argumentative writing? Why?

2. Literature review

2.1. Chatbot-based learning and ChatGPT-based learning

Chatbot-based learning refers to developing knowledge and skills through natural-language interactions between humans and computers (Pérez et al., 2020). Since the 1960s, this approach has evolved alongside increasingly sophisticated conversational systems—from ELIZA to Anthropic Claude—demonstrating consistent effectiveness (Chang et al., 2023). Learners engage in exercises, instructional dialogues, role-plays, independent writing, and open-ended debates via interacting with chatbots, which systematic reviews and meta-analyses confirm as beneficial across multiple fields, including language education (Deng & Yu, 2023; Zhang et al., 2023c).

ChatGPT represents a significant breakthrough in chatbot-based learning. Since 2018, OpenAI's GPT models—such as GPT-3 (175 billion parameters; Brown, 2020) and GPT-4 (~1.8 trillion parameters; Schreiner, 2023)—have showcased exceptional natural language understanding and generation. Chatbots powered by these models (collectively termed ChatGPT) have demonstrated considerable educational potential (Kohnke et al., 2023). Baidoo-Anu and Ansah (2023) synthesised research highlighting ChatGPT's support for personalised and interactive learning, formative assessments, and ongoing feedback.

In another synthesis study, Kasneci et al. (2023) identified ChatGPT affordances in creating varying educational content, eliciting deep engagement, managing rich human-computer interaction, and personalising learning experiences, which may engage learners and facilitate knowledge development.

Particularly, ChatGPT-based learning has demonstrated great usefulness in EFL education. For example, Song and Song (2023) engaged 50 Chinese undergraduates over 12 weeks, comparing conventional teacher-centred instruction with ChatGPT-based, where the chatbot answered questions, analysed writings, and provided real-time feedback and personalised suggestions. Their analysis of pre- and post-writing samples demonstrated ChatGPT's superiority in enhancing EFL writing skills. Similarly, Ghafouri et al. (2024) involved 24 university students over 10 weeks, with one group receiving ChatGPT-based instruction and the other traditional lessons. Pre-, post-, and delayed writing assessments revealed significant gains favouring the ChatGPT group. Kucuk (2024) engaged 30 undergraduates over seven weeks using ChatGPT for EFL grammar learning; the chatbot offered instruction, answered queries, analysed texts, and corrected errors. Pre- and post-tests showed notable grammar improvements, and focus group feedback indicated high learner satisfaction with this approach.

Several theoretical frameworks offer valuable insights into the mechanisms underlying the effectiveness of ChatGPT-based learning. Among these, sociocultural theory (Vygotsky & Cole, 1978) is particularly salient. This theory emphasises the critical role of social interaction with and scaffolded support provided by "more capable peers" in the development of knowledge, skills, and the successful completion of learning tasks (p. 86). Within the context of ChatGPT-based learning, the chatbot can be conceptualised as a virtual "more capable peer," delivering personalised instructional content and feedback tailored to learners' individual knowledge gaps and needs (Elbably & Nemt-allah, 2024). This aligns with the notion that learning is mediated through interaction with knowledgeable others, thereby facilitating cognitive growth.

In language learning specifically, ChatGPT's role as a "more capable peer" is further substantiated by Krashen's (1992) Input Hypothesis (Yang & Li, 2024). This hypothesis posits that language acquisition occurs most effectively when learners receive input that is slightly beyond their current proficiency level—commonly referred to as comprehensible input. ChatGPT-based learning corresponds closely to this principle by providing comprehensible and contextually relevant linguistic input that adapts dynamically to the learner's competence, thus optimising the conditions for language development.

Self-regulated learning theory (Zimmerman, 2013) also provides explanations for the efficacy of ChatGPT-based learning, which highlights learners' proactive role in managing their learning processes through self-regulated learning strategies, such as goal-setting, planning, elaborating, rehearsal, and self-reflection. ChatGPT supports self-regulated learning strategies by affording learners substantial autonomy, facilitating personalised goal-setting and planning, enabling flexible pacing, and encouraging self-reflection. Additionally, the platform offers emotional support, which collectively fosters learners' motivation and persistence, thereby promoting the development of knowledge and skills (Baidoo-Anu & Ansah, 2023; Yıldız, 2023; Zhu, 2025).

Another pertinent theoretical lens is cognitive load theory (Paas & Sweller, 2014, pp. 27–42; Sweller, 2011), which explains learning efficacy in terms of managing cognitive demands. According to this theory, optimal learning is achieved when intrinsic cognitive load (the inherent complexity of the material) is appropriately matched to the learner's proficiency, extraneous load (irrelevant or distracting cognitive demands) is minimised, and germane load (the cognitive resources devoted to processing and understanding) is maximised. ChatGPT may help regulate these cognitive loads by assessing learner proficiency, delivering clear and concise instructional materials, and prompting active cognitive engagement through meaningful interactions and discussions with learners (Kohnke et al., 2023; Woo et al., 2024), which may help learners allocate mental resources efficiently and thereby enhance learning outcomes.

Finally, Krashen's (1982) affective filter theory elucidates another dimension of ChatGPT's effectiveness. This theory emphasises that a low affective filter—characterised by heightened motivation, comfortableness, and self-efficacy alongside reduced anxiety—enhances learners' capacity to process comprehensible input, thereby improving language acquisition. Empirical evidence suggests that ChatGPT-based learning contributes positively in this regard by increasing learner motivation, comfortableness, and self-efficacy while alleviating anxiety (Yıldız, 2023). This is achieved through immediate, personalised feedback and the freedom for learners to control the pace of their learning in ChatGPT-based learning (Annamalai & Nasor, 2025). Moreover, ChatGPT's nonjudgmental, friendly, honest, and calm demeanour creates a comfortable and supportive learning environment, further enhancing learners' sense of control and emotional engagement during interactions (Siregar et al., 2025).

In sum, research highlights ChatGPT's considerable usefulness and effectiveness in educational contexts, especially in EFL learning, as explained by multiple theoretical frameworks. By generating instructional materials, providing immediate feedback, analysing learner outputs, and facilitating personalised interaction, ChatGPT-based learning can enhance language proficiency. However, its underlying mechanisms remain insufficiently explored (Kohnke et al., 2023).

2.2. Developing English argumentative writing logic

English argumentative writing is a form of English writing for expressing viewpoints and presenting arguments to persuade readers. For EFL learners, it is both essential and challenging (Hirvela, 2013). Studies show that EFL learners often struggle to address logical fallacies and establish strong logical connections, which diminishes the quality of their argumentative writings (El Khoiri & Widiati, 2017; Saidi, 2020). Therefore, it is crucial to help EFL students enhance English argumentative writing proficiency by developing logic knowledge and skills (Murray, 2012).

Unlike general EFL writing instruction that emphasises linguistic components (grammar, vocabulary, sentence structure), learning argumentative writing logic focuses on rhetorical knowledge—principles of argument construction and fallacy identification—and higher-order cognitive skills such as reasoning and persuasion within authentic contexts (El Khoiri & Widiati, 2017; Murray, 2012). Effective strategies identified include: (a) detailed instructions on logical concepts (Saidi, 2020); (b) exercises with feedback (El Khoiri & Widiati, 2017); (c) analysis of authentic English arguments from a logical perspective (Hundleby, 2010); and (d) discussions of logical concepts and connections within authentic contexts (Guo & Lee, 2023). Empirical studies support these approaches. For instance, Nejmaoui (2019) provided a semester-long course on English argumentative writing logic, reporting significant improvements in students' writing samples. Song and Sparks (2019) engaged middle school students in digital game-based exercises to identify reasoning errors in English arguments, which proved effective in enhancing logic knowledge. Additionally, Zhang et al. (2023b, 2024) developed a tree-model chatbot to provide logic instructions and exercises, yielding positive pre- and post-test results in students' logic knowledge and skills.

Overall, the literature underscores the importance of argumentative logic learning and effective strategies for EFL learners. Yet, empirical studies remain limited.

2.3. Developing English writing and English argumentative writing logic via ChatGPT

ChatGPT has shown promise in enhancing English writing proficiency (Teng, 2024a). Teng's review of 20 studies identified multiple

ChatGPT applications in writing classes, including topic and outline suggestion, material recommendation and summarisation, proofreading, question answering, and generating templates or sentence starters. Su et al. (2023) synthesised research on ChatGPT-based learning for English argumentative writing, noting its support for brainstorming, providing feedback on outlines and drafts, and facilitating post-writing reflection.

So far, empirical research specifically on ChatGPT-based learning of argumentative writing logic remains scarce but suggests potential benefits. For example, Guo and Lee (2023) required students to develop critical thinking skills for English science writing via academic discussions with ChatGPT and reported positive results. Lee et al. (2023) utilised ChatGPT to assist debate skill development and identified its usefulness in engaging learners, initiating tasks, offering suggestions, analysing learner output, and providing feedback from the logical perspective. They emphasise prompt engineering—training GPT-based chatbots with curated examples and detailed instructions—to improve response depth, accuracy, and contextual relevance, which is vital for personalised interactions and discipline-specific skill development. Khampusaen (2024) engaged 30 undergraduates in a 16-week ChatGPT-assisted EFL argumentative writing programme. Pre- and post-writing demonstrated the significant effectiveness of the programme in enhancing EFL argument construction, evidence integration, and academic voice development. Darmawansah et al. (2025) conducted a ChatGPT-based activity to enhance EFL speaking and argumentative skills through group brainstorming, information gathering, scripting, and debating, with ChatGPT providing support. Thirty-four Chinese undergraduates in the ChatGPT group outperformed a control group in speaking skills and critical thinking awareness. Annamalai et al. (2025) interviewed 20 students using ChatGPT for argumentative writing, revealing positive perception of this tool regarding vocabulary support, feedback, and emotional benefits, alongside strong willingness to continue using the tool.

However, Su et al. (2023) raised concerns about the ethical issues and inconsistent quality of GPT responses in ChatGPT-based learning. Teng (2024a) also warns of potential learner over-reliance on ChatGPT, which may impair critical thinking development.

In conclusion, while previous research has acknowledged ChatGPT-based learning's significant potential for developing English argumentative writing logic, concerns about response quality and learner over-reliance persist. Further empirical investigation is warranted to explore and optimise this instructional approach.

2.4. Learner engagement with ChatGPT-based learning of English writing

Learner engagement in language learning tasks manifests behaviourally through observable actions (Gobert et al., 2015), cognitively through mental effort and strategies employed to achieve objectives (Plass et al., 2015), and emotionally through affective connections and intrinsic motivation (Hiver et al., 2021). These three dimensions of learner engagement are closely intertwined and interrelated (Hiver et al., 2021). Given this complexity, a combination of objective and subjective measures is essential to capture the full scope of learner engagement (Philp & Duchesne, 2016). Subjectively, introspective interviews are commonly used to gauge learner engagement (Philp & Duchesne, 2016). Objectively, eye-tracking methodologies that analyse eye movements, such as fixation count and fixation duration, in relevant areas are effective indicators of engagement (Latif, 2019).

A thorough understanding of learner engagement is crucial for fostering effective learning outcomes (Hiver et al., 2021), particularly in writing education, which is explainable via the noticing hypothesis (Schmitt and Schmitt, 2020). The hypothesis is based on three concepts: input, referring to all instructional materials learners are exposed to; intake, referring to the knowledge points learners internalise from the input and incorporate into their long-term memory; output, referring to what language learners produce and reflects their current knowledge. This dynamic cycle—input leading to intake through noticing, which informs output, and subsequent feedback influencing further input—facilitates effective language acquisition – in the process, learner engagement acts as a facilitator of the noticing (Schmitt and Schmitt, 2020). Deeply engaged learners excel at quickly noticing and processing information (Stroud, 2014), filtering out irrelevant messages (Yu & Tsuei, 2023), and constructing mental models efficiently (Ellis, 2019), all of which correlate positively with learning outcomes (Hiver et al., 2021). In the context of EFL writing education, empirical studies demonstrate that deep engagement with learning materials and tasks is especially crucial (Hirvela, 2013), as it significantly enhances writing proficiency (Tian & Zhou, 2020).

ChatGPT-based learning has been shown to foster deep learner engagement. In a meta-analysis of 17 studies, Heung and Chiu (2025) found a medium effect size on overall student engagement with ChatGPT-based learning compared to non-ChatGPT methods, attributing this to ChatGPT's ability to personalise learning materials and feedback, generate high-quality instructional content, and facilitate collaboration. Wu and Ho (2025) examined 687 Chinese university students in a ChatGPT-based programming course, where learners perceived GPT responses as relevant and useful, leading to increased confidence and satisfaction, and enhanced engagement and perceived academic achievement.

The effectiveness of ChatGPT-based learning in promoting learner engagement is also evident in EFL writing education. For example, Woo et al. (2024) required 21 Chinese students to do a writing task through 45-min discussion with ChatGPT. The researchers reported the usefulness of ChatGPT in engaging students by tailoring responses to learner prompts and managing natural conversations. Similarly, Teng (2024b) required 45 Chinese to edit and proofread their writings under ChatGPT's assistance. The questionnaire and interview results revealed the usefulness of ChatGPT in engaging learners by providing detailed and structured feedback.

In sum, existing studies emphasise the critical role of learner engagement in developing EFL writing proficiency and the overall deep learner engagement with ChatGPT-based learning. However, most studies treat engagement in ChatGPT-based learning as a single construct, neglecting its distinct components. Additionally, previous studies have predominantly relied on qualitative self-reported data, such as questionnaires and interviews (e.g., Teng, 2024b). While these methods provide valuable insights, incorporating objective data could further enrich the understanding of learner engagement in ChatGPT-based learning.

2.5. Analysis of the existing literature and the research rationales

Based on the literature review, we argue that ChatGPT can effectively support the learning of English argumentative writing logic. Proven strategies for developing logical knowledge and skills in English argumentative writing include reading explicit instructional materials (Nejmaoui, 2019), completing targeted logical exercises (Song & Sparks, 2019), analysing model texts for logical structure (Zhang et al. (2023, 2024) examining authentic argumentative texts (El Khoiri & Widiati, 2017). These approaches align well with ChatGPT's capabilities, which include delivering personalised instruction and examples, offering exercises with immediate feedback, facilitating open-ended academic discussions, and providing comprehensive text analysis (Baidoo-Anu & Ansah, 2023; Kasneci et al., 2023; Kohnke et al., 2023).

Within ChatGPT-based learning, learners interact with two key components—GPT responses and learner prompts—engaging simultaneously on behavioural, cognitive, and emotional dimensions. When engaging with GPT responses, learners attentively read instructions and feedback, processing new knowledge through analysis, comprehension, memorisation, and internalisation (Lestari, 2024; Zhang et al., 2025b), often accompanied by increased motivation and enjoyment (Oktawirawan, 2022). When engaging with learner prompts, learners

type prompts, completing exercises while elaborating on, analysing, and synthesising logical concepts (Lestari, 2024; Zhang et al., 2025b), accompanied by feelings of comfort, confidence, and control.

Reflecting on these two components, ChatGPT has the potential to elicit deep engagement in logic learning across multiple dimensions. From a sociocultural perspective (Elbably & Nemt-allah, 2024; Vygotsky & Cole, 1978), ChatGPT acts as a virtual "more capable peer," engaging learners in academically relevant, naturalistic dialogue that motivates active reading, analysis, and prompt generation (Woo et al., 2024). Furthermore, ChatGPT supports learner autonomy by allowing self-paced study and providing personalised materials and tasks (Annamalai & Nasor, 2025). This environment encourages the deployment of self-regulated learning strategies, such as resource seeking, questioning, and planning (Zimmerman, 2013), thereby fostering behavioural, cognitive, and emotional engagement (Zimmerman, 2013). Moreover, ChatGPT's nonjudgmental stance (Skjuve et al., 2021), immediate, detailed feedback (Annamalai & Nasor, 2025; Teng, 2024b), and diversified, high-quality personalised content (Heung and Chiu, 2025) can elicit emotional engagement by reducing anxiety and enhancing motivation, satification, and self-efficacy (Wu & Ho, 2025; Yıldız, 2023).

Such multifaceted engagement in GPT responses and learner prompts may facilitate English argumentative writing logic learning. From the perspective of cognitive load theory (Paas & Sweller, 2014, pp. 27–42; Sweller, 2011), this engagement substantially increases learners' germane cognitive load, thereby promoting deeper processing and comprehension of logical knowledge and skills. From the perspective of the noticing hypothesis (Schmitt and Schmitt, 2020), deep engagement in GPT responses and learner prompts promotes greater noticing of knowledge intake, facilitating more efficient comprehension and internalisation of English argumentative writing logic. As learners' logical knowledge deepens, they become more engaged with academic prompts, perpetuating a virtuous cycle of noticing and input generation through triggering GPT responses. Collectively, current research strongly suggests that ChatGPT-based learning holds considerable potential to elicit sustained learner engagement in GPT responses and learner prompts and develop logical knowledge and skills in English argumentative writing.

Two key research gaps emerge from the analysis of existing literature. Firstly, while studies have established the effectiveness of ChatGPT-based learning in enhancing general EFL writing proficiency (Ghafouri et al., 2024; Kucuk, 2024; Song & Song, 2023), critical thinking skills (Guo & Lee, 2023), and argumentative speaking skills (Darmawansah et al., 2025; Lee et al., 2023), few have specifically examined the development of logic knowledge and skills within English argumentative writing. The results of these existing studies may also be inapplicable in English argumentative writing logic learning, considering its distinct differences from general critical thinking education and EFL writing and speaking learning in pedagogical foci and learning methods (El Khoiri & Widiati, 2017; Murray, 2012). Given the significance of logical competence in EFL argumentative writing (Murray, 2012) and the promising capabilities of ChatGPT in this area (Zhang et al., 2024), it is essential to fully utilise ChatGPT in developing English argumentative writing logic. Addressing this gap will enhance our understanding of ChatGPT's potential as a language learning and logic learning tool and yield valuable insights for future education of English argumentative writing.

Secondly, existing studies have examined ChatGPT-based learning as a whole and reported overall deep learner engagement (Heung and Chiu, 2025; Wu & Ho, 2025), including within the context of EFL writing education (Teng, 2024b; Woo et al., 2024). However, these studies have largely overlooked the distinct learner engagement associated with the interactive components of this learning method—learner prompts and GPT responses (Lestari, 2024). Given the differences in linguistic characteristics (Jiang & Hyland, 2024) and learner utilisation patterns (Lestari, 2024) between these two components, learner prompts and GPT responses may elicit varying types and degrees of engagement, which may be reflected in distinct eye movement patterns when attending to the respective areas and influence learning outcomes in divergent ways. For instance, as learners primarily engage with prompts through typing and with GPT responses through reading, fixation durations may be longer during prompt formulation than when reading responses. Moreover, engagement in learner prompts involves higher-order cognitive processes to a greater extent than engagement in GPT responses (Lestari, 2024), potentially contributing more substantially to learning outcomes in ChatGPT-based environments. Given the critical role of understanding learner engagement in enhancing language learning efficacy (Philp & Duchesne, 2016), empirical research that separately investigates how learners engage with learner prompts and GPT responses within ChatGPT-based learning is warranted. By combining subjective measures of perceived engagement in learner prompts and GPT responses with objective eye-tracking data focused on these distinct components, a more nuanced and comprehensive understanding of learner engagement with ChatGPT can be achieved, offering valuable practical and theoretical insights for future educational applications.

To address these gaps, we developed a GPT-4-powered chatbot, *LogicalHamster*, for developing logic knowledge and skills in English argumentative writing. Using this tool, we investigated how EFL learners engage with the two key components of ChatGPT-based learning—GPT responses and learner prompts—and examined the potential effects of such engagement on their learning outcomes. By undertaking this research, we may better understand the mechanisms and effects of learner engagement with ChatGPT-based learning, providing insights relevant not only to developing logic skills and English argumentative writing proficiency but also potentially to other educational domains employing ChatGPT technology.

3. Method

This study employed a concurrent triangulation mixed-method approach, integrating the collection, analysis, and interpretation of quantitative data (total fixation counts and average fixation duration) and qualitative data (interview transcripts) to provide a comprehensive understanding of learner engagement in GPT responses and learner prompts in ChatGPT-based learning of English argumentative writing logic. Additionally, pre- and post-tests and writing tasks were administered to assess learning outcomes in relation to learners' initial proficiency levels. This approach helped to ensure the credibility and validity of our research findings.

3.1. Participants

We recruited 42 Chinese undergraduate and postgraduate students (34 female, eight male), aged between 18 and 28 ($M = 22.90$, $SD = 3.27$), from a university in Hong Kong using voluntary response sampling. According to self-reported biographical data, all participants had recently obtained IELTS scores ranging from 6 to 7.5, indicating their homogenous, upper-intermediate to advanced level of English proficiency, categorised as "competent" to "good" English users (British Council; https://takeielts.britishcouncil.org/teach-ielts/test-information/ielts-scores-explained). All participants were enrolled in English-as-a-medium courses and required to complete argumentative writing assignments in English, reflecting a strong need to enhance their English argumentative writing abilities. Furthermore, all students reported familiarity with using ChatGPT and expressed no technophobia, making them suitable candidates for ChatGPT -based learning interventions. None had previous experience with ChatGPT -based instructional activities, logic training, or prompt engineering. Consequently, all respondents were included in the final sample.

Before the intervention, each participant was provided with a consent form, which they signed after being assured that (a) the experiment

would cause no harm to them, (b) their information would remain anonymous, (c) their performance in the study would have no consequences, and (d) they could withdraw from the study at any time.

3.2. GPT-4-powered chatbot for developing knowledge and skills of logic in English argumentative writing

Based on GPT-4, we developed a POE chatbot named *LogicalHamster* (https://poe.com/LogicalHamster), designed to teach learners about seven common reasoning errors in English argumentative writing as well as strategies to address these errors (El Khoiri & Widiati, 2017; Murray, 2012). Similar to many ChatGPT-based learning tools, *LogicalHamster* delivers educational content and supports skill development via interactive, natural language dialogues involving GPT responses and learner-generated prompts. The tool's graphical user interface (see Fig. 1) is organised into two primary, user-adjustable sections: the GPT response area, which presents instructional content and feedback generated by *LogicalHamster*, and the learner prompt area, where learners actively compose and submit text-based questions or responses. Both interface areas can be resized by learners to accommodate individual preferences and optimise their interaction with the system.

Following recommendations by Lee et al. (2023), we employed prompt engineering techniques to enhance *LogicalHamster*'s effectiveness in engaging learners and teaching English argumentative writing logic. Specifically, *LogicalHamster* was designed to teach English argumentative writing logic through five approaches recommended by previous research: (a) delivering explicit instructions on the forms, meanings, and structures of target logical concepts (Murray, 2012); (b) presenting illustrative examples with detailed explanations (Saidi, 2020); (c) providing practice exercises with immediate feedback (Song & Sparks, 2019); (d) facilitating discussions on logical concepts through open-ended questions (Guo & Lee, 2023); and (e) analysing logical issues in learners' submitted arguments to offer personalised feedback (Hundleby, 2010).

To foster learner engagement, *LogicalHamster* adopted four additional strategies suggested by prior studies: (a) tailoring responses according to learners' identified needs and preferences expressed in prompts (Kasneci et al., 2023); (b) embodying an engaging persona (Ruan et al., 2019)—a humorous and lively hamster character; (c) providing immediate and personalised feedback on learner inputs (Teng, 2024b); and (d) conducting interactions in a natural, humorous conversational tone (Zhang et al., 2023c).

In a pilot study, five students with backgrounds similar to our target participants independently interacted with *LogicalHamster* at their own pace and preference. Semi-structured interview feedback from these students indicated that *LogicalHamster* was user-friendly, engaging, and helpful for improving their English argumentative writing logic. They recommended an optimal learning duration of between 45 and 75 min per session. Additionally, evaluations of GPT responses during these learner interactions identified no biases or inaccuracies. Therefore, LogicalHamster can be considered a valid and reliable tool for ChatGPT-based learning of English argumentative writing logic.

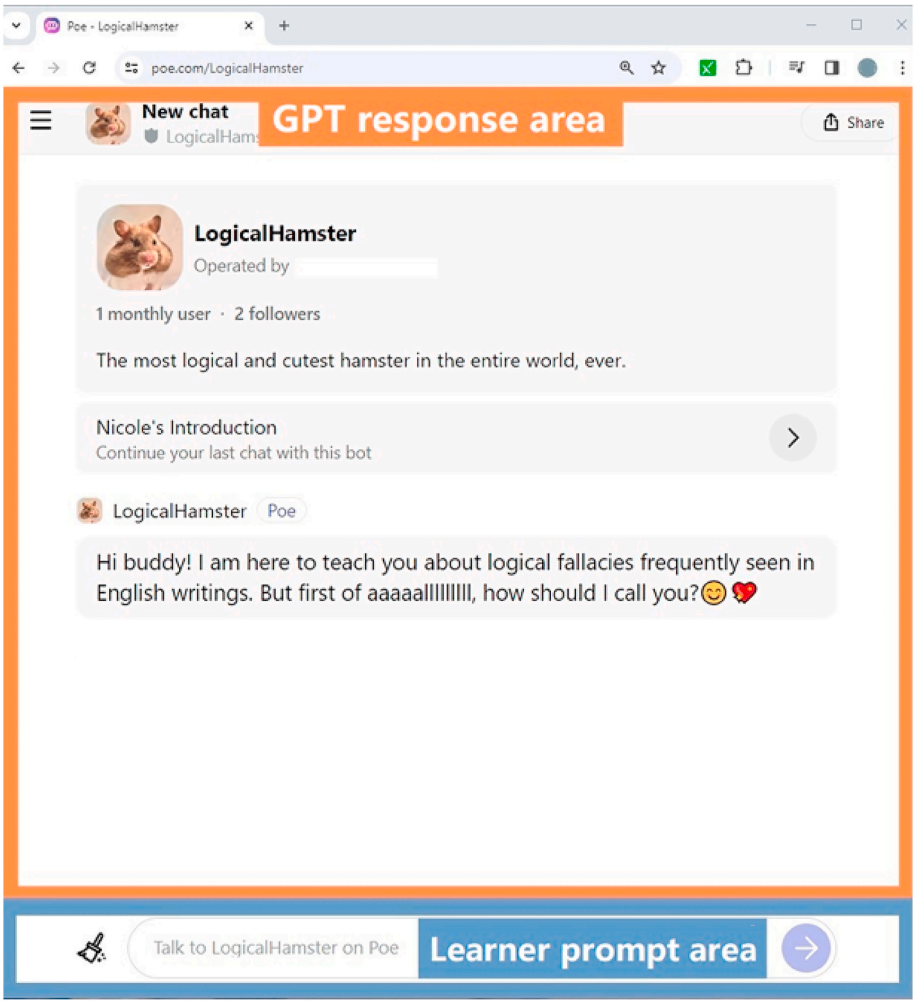

Fig. 1. *LogicalHamster*'s graphical user interface

3.3. Data collection instruments

To thoroughly explore the research questions, we employed both objective measures (eye-tracking data, test scores, and writing scores) and self-report data (semi-structured interview transcripts) to assess learner engagement in GPT responses and learner prompts in ChatGPT-based learning of English argumentative writing logic and potential changes in students' logical knowledge and their English argumentative writing logic as a result of this learning experience.

3.3.1. Eye-tracking data

Consistent with prior research (Latif, 2019; Obaidellah et al., 2018), we objectively measured learner engagement in GPT responses and learner prompts using total fixation counts and average fixation durations (in seconds). During the ChatGPT-based learning, students worked individually on desktop computers while their eye movements were recorded using a Tobii X2-30 screen-based eye tracker, operating at a sampling rate of 30 Hz. The recordings were analysed using Tobii Studio Enterprise, a tool noted for its high accuracy and precision (Wang et al., 2021), to calculate total fixation counts and average fixation durations specifically within the GPT response and learner prompt areas of the graphical user interface. Importantly, as participants could dynamically resize both the GPT response and learner prompt areas according to their preferences during the learning, the boundaries of these areas were continuously adjusted in the eye-tracking metric collection, aiming to ensure the validity and reliability of our engagement indicators despite interface variability.

3.3.2. Pre, post, and delayed-tests on logic knowledge

To assess possible improvements in students' logic knowledge, we administered three knowledge tests on English argumentative writing logic at three stages: prior to the ChatGPT-based learning, immediately after, and one week later. Following Zhang et al.'s (2024) framework, the test required students to identify and explain reasoning errors in English arguments and to associate these errors with appropriate terminology (see Appendix A). The three tests have the same content and format, ensured the internal validity.

The tests were blind-scored according to Zhang et al.'s (2024) scoring system (Appendix B), with total scores ranging from 0 to 21. The research team independently scored the tests, achieving high inter-rater reliability (Cohen's kappa = .96), with any discrepancies resolved through discussion.

3.3.3. English essay writing and revision tasks

To evaluate the potential development of students' English argumentative writing logic, we assigned participants pre-treatment essay writing and post-treatment revision tasks. Prior to the ChatGPT-based learning, participants wrote a 250-word essay within 40 min, adhering to IELTS academic writing guidelines (https://www.chinaielts.org/sites/all/themes/newyasi/images/txjs/%E5%AD%A6%E6%9C%AF%E7%B1%BB%E5%86%99%E4%BD%9C_Academic_writing_task_2A.pdf), which are commonly used to assess proficiency in this domain (e.g., Soodmand Afshar et al., 2017). The essay prompt, "Young people who commit crimes should be treated in the same way as adults. To what extent do you agree or disagree?" was sourced from IELTS academic writing topic banks. Following the learning intervention, students were given an additional 30 min to revise their essays' logic. Participants completed both writing tasks individually using MS Word on a desktop computer, without access to external resources or the Internet.

Original and revised essays were blind-scored based on logical criteria derived from Finken and Ennis's Illinois Critical Thinking Essay Scoring Rubric and IELTS's academic writing assessment criteria (see Appendix C). Scores ranged from 0 to 100. Scoring was conducted independently by two authors and one colleague, all experienced in teaching and assessing English essays. The scoring exhibited good validity (Cohen's kappa = .85), with differences reconciled through discussion.

3.3.4. Semi-structured interview questions

We conducted semi-structured interviews, allowing interviewees the flexibility to reflect on their learning experiences, provide explanations, share examples, and elaborate on details. The interviews were based on two questions aligned with the research questions:

Concerning RQ1:

- In the previous learning, what percentage of your attentional resources were devoted to GPT responses/learner prompts? Why?

Concerning RQ2:

- Did your engagement in GPT responses/learner prompts have positive/negative/neutral/mixed influences on your logic knowledge and English argumentative writing logic? Why?

The instruments were reviewed and validated based on expert recommendations. In the pilot study, five students with backgrounds similar to our target participants performed the ChatGPT-based learning of English argumentative writing logic, on whom we tested and ensured the reliability and validity of the instruments.

3.4. Experimental procedures

The experiment spanned three weeks (Fig. 2). In Week One, we collected and analysed students' biographical reports to select participants, who then signed consent forms. A 20-min orientation on English argumentative writing logic and ChatGPT-based learning was also provided.

In Week Two, participants wrote a 250-word English essay on MS Word within 40 min, followed by a 30-min pre-test assessing their logic knowledge. The knowledge test was scheduled after the writing task to minimise any potential influence on writing performance.

After the pre-test, participants engaged in ChatGPT-based learning of English argumentative writing logic using *LogicalHamster* on desktop computers for 45–75 min, based on feedback from a pilot study. Participants' eye movements were recorded using a Tobii X2-30 screen-

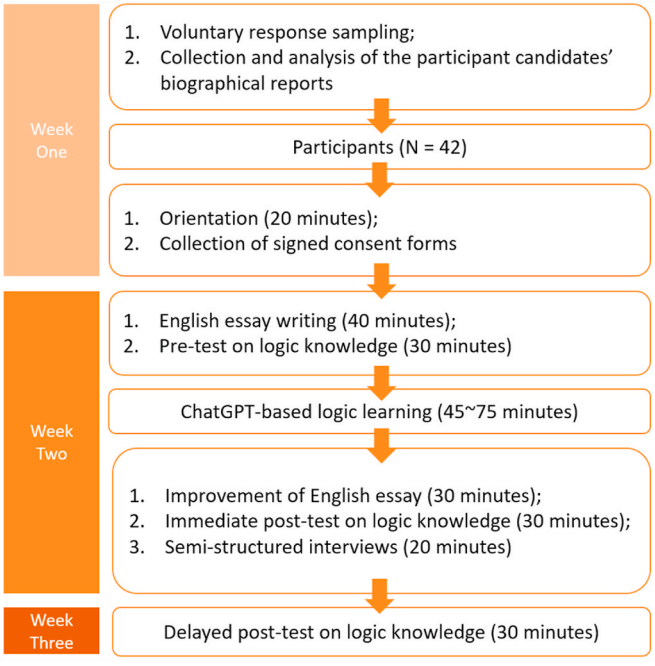

Fig. 2. Experimental procedures

based eye tracker, with no interventions apart from necessary technical assistance.

Following the learning session, participants revised their pre-learning essays from a logical perspective within 30 min and then completed an immediate post-test on logic knowledge, also within 30 min. The post-test was sequenced after the essay revision to further minimise its impact on writing performance.

Additionally, 20 participants were randomly selected for semi-structured interviews. Before the main interviews, we explained the concepts of learner engagement related to GPT responses and learner prompts, emphasising their distinction from engagement with *LogicalHamster*'s interface. Concept-checking questions (e.g., "Is engagement in learner prompts and GPT responses the same as engagement with *LogicalHamster*'s interface?") were used to ensure participants understood these concepts. The main interviews lasted approximately 20 min and were conducted in conversational Chinese, the participants' native language.

In Week Three, participants completed a delayed post-test on the target logic knowledge within 30 min. The pre-, post-, and delayed tests were identical to ensure internal validity.

3.5. Data analysis

3.5.1. Qualitative analyses

We conducted a thematic analysis of the semi-structured interview data, following six steps. First, we recorded and transcribed the semi-structured interviews and familiarised ourselves with the content by repeatedly reviewing the transcripts. Next, we generated initial codes to categorise the interviewees' primary viewpoints concerning:

1. The extent of their engagement in GPT responses and learner prompts, and why, to address RQ1;
2. Their perceived impacts of engaging with GPT responses and learner prompts, and why, to address RQ2.

In the third step, three authors collaboratively analysed three transcripts to establish a consensus on the coding method. The fourth step involved a deductive analysis of the remaining transcripts individually, using the established codes. Any new learner opinions identified were added to the coding scheme. Simultaneously, we calculate the frequency of learner viewpoints and selected representative texts for quoting. Fifth, we compared individual coding outcomes to ensure internal validity (Cohen's *kappa* = .90), resolving differences through discussion. Lastly, we jointly translated the analysis results and quotations into English.

3.5.2. Quantitative analyses

To address RQ1, descriptive analysis focusing on means and standard deviations of total fixation counts and average fixation durations on GPT response area and learner prompts was conducted. Additionally, we conducted paired-sample t-tests in SPSS to explore differences in fixation counts and durations between GPT responses and learner prompts.

To address RQ2, we conducted Partial Least Squares Structural Equation Modeling (PLS-SEM) using the SEMinR package in R (R core team, 2019), to predictively examine the relationships between learner engagement in GPT responses and learner prompts and learning outcomes while controlling for learner factors. The exogenous constructs were: (a) "Engagement in GPT responses", measured by learners' total fixation counts and average fixation durations on the GPT responses area; (b) "Engagement in learner prompts", measured by learners' total fixation counts and average fixation durations on the learner prompts area; (c) "Learners' prior logic knowledge", measured by learners' pre-test scores of logic knowledge; (d) "Learners' prior English argumentative writing logic", measured by learners' pre-writing scores. The endogenous constructs were: (a) "ChatGPT-based learning of English argumentative writing logic outcomes concerning logic knowledge", measured by learners' immediate and delayed post-test scores of logic knowledge; (b) "ChatGPT-based learning of English argumentative writing logic outcomes concerning English argumentative writing logic", measured by learners' post-writing scores. Path coefficient significance was tested at a 5 % level through bootstrapping with 10,000 subsamples.

To further address RQ2, we conducted multiple linear regressions in SPSS based on the PLS-SEM findings, focusing on the predictive powers of learner engagement on learning outcomes.

Individual data analyses demonstrated high internal validity (Cohen's *kappa* = .96), with discrepancies resolved through discussion.

3.5.3. Reliability and validity

Our data demonstrate linearity and homoscedasticity, with points evenly dispersed across scatterplots and aligning along the diagonal in P-P plots (Field, 2018). They also demonstrate normality with skewness from −2 to +2 and kurtosis from −7 to +7 (Byrne, 2010). Hence, our data distributions are suitable for paired-sample t-tests and multiple linear regressions (Field, 2018). They are also suitable for PLS-SEM, which impose no data distribution requirements (Hair et al., 2021).

According to widely accepted rules of thumb in quantitative research, our sample size ($N = 42$) is adequate for the statistical analyses conducted. Specifically, for paired-sample t-tests, a minimum of 30 observations is generally recommended to ensure sufficient statistical power and approximate normality of difference scores (Field, 2018), which our sample surpasses. For multiple linear regression, the standard guideline is at least 10 observations per predictor to prevent overfitting and ensure stable estimates (Field, 2018; Hair et al., 2019). Our regression models include three predictors ("Pre-test scores of logic knowledge", "Total fixation counts on GPT response area", and "Average fixation durations on GPT response area"), thus satisfying this criterion. For PLS-SEM, the frequently cited "10-times rule" requires a minimum sample size equal to ten times the largest number of indicators for any latent variable (Hair et al., 2019). As the largest number of indicators in our model is two, our sample size comfortably exceeds this requirement.

In multiple linear regressions, multicollinearity wasn't a concern (Pearson's r for all indicator pairs <.90).

As shown in Appendix D, our measurement models of PLS-SEM showed satisfactory indicator reliability, with most variables showing loadings above .708, high internal consistency (Cronbach's *alpha* > .700), and strong convergent validity (AVE >.50). Collinearity issues were not present (VIFs <3). Notably, "Average fixation durations on GPT response area" had slightly lower but acceptable loadings (below .708 but above .40). Given the model's reliability, validity, and theoretical basis, we retained these variables per Hair et al.'s recommendation (2021).

4. Results

4.1. RQ1. How do EFL learners engage with GPT responses and learner prompts in ChatGPT-based learning of English argumentative writing logic? Why?

4.1.1. Quantitative analysis results

In ChatGPT-based learning of English argumentative writing logic, our participants had a total of 3903.90 fixations ($SD = 1800.86$) on the GPT response area, and 456.05 fixations ($SD = 481.77$) on the learner prompt area (Fig. 2). Their average fixation duration was .21 s ($SD = .05$) on the GPT response area and .26 s ($SD = .11$) on the learner prompt area (Fig. 3).

The paired-sample t-tests (Table 1) revealed significant differences between GPT responses and learner prompts in learners' total fixation count ($t(41) = 13.11$, $p < .001$) and average fixation duration ($t(41) = -4.08$, $p < .001$). The analysis showed that learners had more fixations of shorter duration on GPT responses than on learner prompts. This suggests that learners engaged differently with ChatGPT-based learning of English argumentative writing logic, with frequent short fixations on

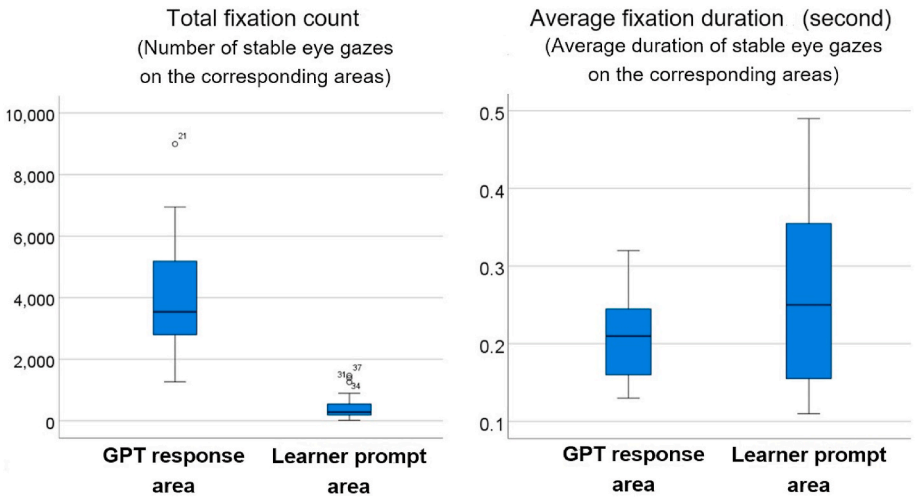

Fig. 3. The participants' total fixation counts and average fixation durations (second) on GPT responses and learner prompts

Table 1
Paired-sample *t*-tests examining the differences in learners' total fixation counts and average fixation duration between GPT response area and learner prompt area.

	M	*SD*	95 % CI	*t* (41)	Sig.
Total fixation counts	3447.84	1642.977	(2915.255, 3980.438)	13.105	<.000**
Average fixation duration	−.05385	.08251	(-.08059, −.02710)	−4.075	<.000**

**p < .001.

GPT responses and fewer long fixations on learner prompts.

4.1.2. Qualitative analysis results

Our semi-structured interview findings corroborated the quantitative data regarding students' varying engagement in GPT responses and learner prompts in learning English argumentative writing logic. As illustrated in Fig. 4, the majority of interviewees (95 %) reported deep engagement in GPT responses, allocating between 61 % and 100 % of their attentional resources to this component. Students particularly emphasised the importance of GPT responses in enhancing their logic knowledge and supporting their learning tasks. For example, one interviewee noted, "[GPT responses] are the only source of new knowledge, so I devoted most of my time and effort to them." Many interviewees (75 %) also appreciated *LogicalHamster*'s emojis and friendly and humorous communication style, indicating that it "made the learning process less boring" and further facilitated their engagement in GPT responses. Nevertheless, a few interviewees (10 %) acknowledged selectively engaging with different types of GPT responses according to their perceived relevance and usefulness. For instance, one interviewee stated: "I think the examples were very helpful in my understanding of [the target logic knowledge], so I read them repeatedly and think most about them … while some *LogicalHamster*'s jokes and encouragement are not that useful, so I tended to ignore them."

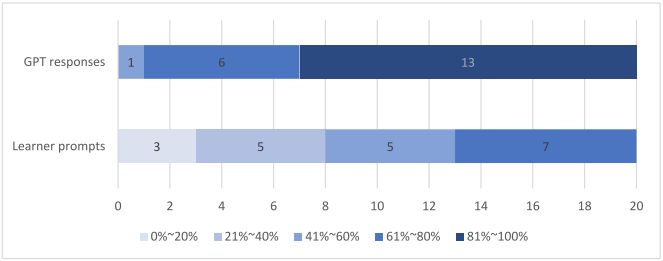

Fig. 4. Numbers of interviewees reporting different perceived levels of engagement in GPT responses and learner prompts, using a scale of 0 %–100 % of their total attentional resources

Learners demonstrated comparatively lower engagement with prompt production, with 40 % of the interviewees reporting less than 40 % of their attentional resources dedicated to learner prompts. These learners explained that "learning was mainly about receiving new information," and that "close attention to [their] own output was unnecessary." A few interviewees (15 %) perceived *LogicalHamster* as incapable of fully understanding complex prompts or handling in-depth discussions, thereby viewing deep engagement with prompt generation as "meaningless." Others (10 %) found producing prompts "tiring" and deliberately reduced their engagement. However, some interviewees (35 %) reported high-level engagement with prompt production, actively formulating exploratory questions and engaging in discussions with *LogicalHamster*, which they considered beneficial for deepening their understanding of logical concepts and reasoning errors. Additionally, several interviewees (25 %) commented positively on *LogicalHamster*'s friendly and patient communication style, which encouraged them to produce more prompts. As one interviewee remarked, "It felt like talking with a knowledgeable, good friend … it motivated me to talk more."

4.2. RQ2. How do EFL learners' engagement in GPT responses and learner prompts in ChatGPT-based learning of English argumentative writing logic influence their logic knowledge and logic in English argumentative writing? Why?

4.2.1. Quantitative analysis results

Fig. 5 summarizes the PLS-SEM findings on the relationships between learner engagement in GPT responses, learner prompts, and learning outcomes in ChatGPT-based learning of English argumentative writing logic. Controlling for learner factors, "Engagement in learner prompts" significantly influenced "ChatGPT-based learning of English argumentative writing logic outcomes," explaining 17.7 % of the variance ($\beta = -.30$, $|t| > 1.96$, 95 % BCCI = (−.529, .082)). "Learner engagement in GPT responses" showed potential significant effects on "ChatGPT-based learning of English argumentative writing logic outcomes regarding logic knowledge" with a coefficient nearing statistical significance, explaining 11.0 % of the variance ($\beta = .35$, $t = 1.882$, BCCI = (−.129, .636), $f^2 = .110$).

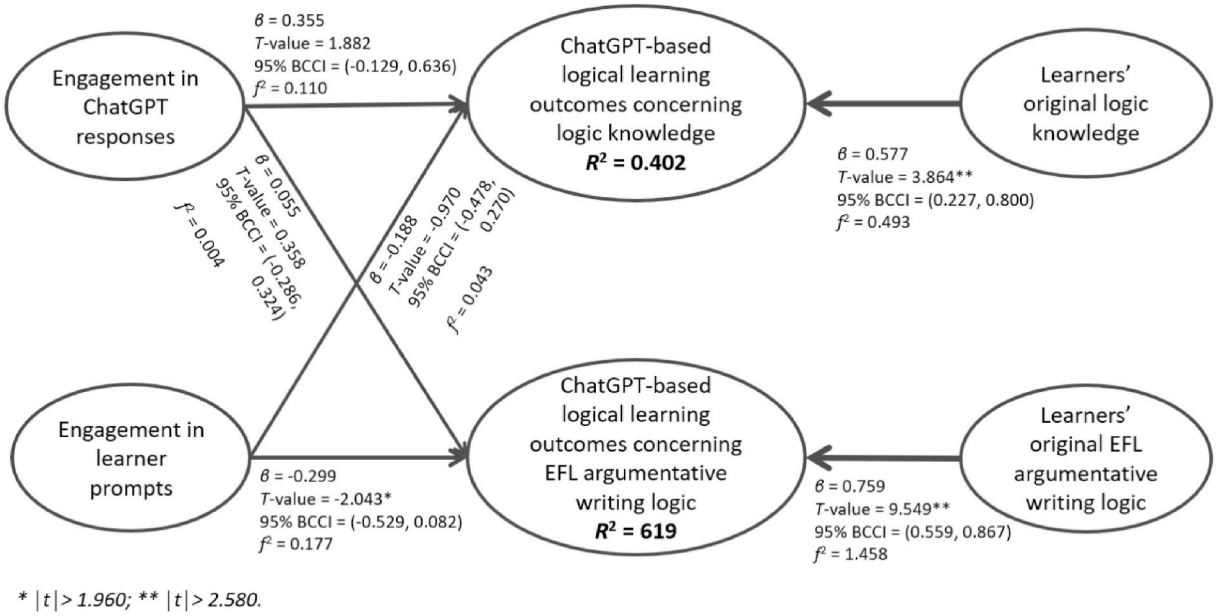

Fig. 5. Summary of the results of the PLS-SEM analysis

Using multiple regression analyses, we further explored the possible relationships suggested by the PLS-SEM analysis between learner engagement and academic outcomes of ChatGPT-based learning of English argumentative writing logic. As for learner engagement in GPT responses (Table 2), our regression results reveal its significant predictive power on learners' long-term retention of logic knowledge after controlling the learner factor, $R^2 = .14$, $F\ (3,\ 35) = 3.43$, $p < .05$. Calculated by G*Power, the estimated power of the R^2 tests is very strong (.99). The independence among predictors was acceptable (Durbin-Watson values between 1 and 3).

The parameter estimates (Table 3) showed that total fixation counts on GPT response area significantly enhanced the long-term retention of logic knowledge ($b = .001$, $t = 2.37$, $p = .02$), even after controlling the learner factor. The statistical powers estimated by G* Power were strong (.98). On average, one more fixation on GPT responses leads to .001-point higher delayed post-test scores of logic knowledge.

Our multiple regression analysis results also showed significant predictive power of learner engagement in learner prompts on learners' English argumentative writing logic, $R^2 = .10$, $F\ (3,\ 35) = 5.27$, $p = .01$ (Table 4). Calculated by G*Power, the estimated power of the R^2 tests is strong (.98). The independence among predictors was acceptable (Durbin-Watson value between 1 and 3).

The parameter estimates (Table 5) showed that learners' fixation durations on prompts significantly reduced their development of English argumentative writing logic after controlling the learner factor, $b = -22.07$, $t = -2.83$, $p = .01$. The statistical powers estimated by G* Power were strong (.99). On average, .1 more seconds of fixation duration on learner prompts leads to 2.21-point lower post-writing scores.

In sum, our quantitative analysis showed that, after controlling the learner factor, more engagement in learner prompts reduced the development of EFL argumentative writing logic, while more engagement in GPT responses develop logic knowledge better, especially long-

Table 3
Parameter estimates of the effects of learner engagement in GPT responses on logic knowledge.

	B	*SE (B)*	*β*	*t*	Sig.
(Constant)	8.315	2.708		3.071	.004*
Pre-test scores of logic knowledge	.580	.190	.441	3.047	.004*
Total fixation counts on GPT response area	.001	.000	.359	2.374	.023*
Average fixation durations on GPT response area	3.695	10.784	.052	.343	.734

*p < .05.

Table 2
Model estimates of the effects of learner engagement in GPT responses on logic knowledge.

	Predictors	Outcomes	R^2	R^2_{change}	F_{change}	Sig. F_{change}	Durbin-Watson
Model 1	(Constant), Pre-test scores of logic knowledge	Immediate post-test scores of logic knowledge	.204	.204	9.487	.004*	n/a
Model 2	(Constant), Pre-test scores of logic knowledge, Total fixation counts on GPT response area, Average fixation durations on GPT response area		.301	.097	2.438	.102	1.758
Model 1	(Constant), Pre-test scores of logic knowledge	Delayed post-test scores of logic knowledge	.146	.146	6.332	.016*	n/a
Model 2	(Constant), Pre-test scores of logic knowledge, Total fixation counts on GPT response area, Average fixation durations on GPT response area		.286	.140	3.430	.044*	1.884

*p < .05.

Table 4
Model estimates of the effects of learner engagement in learner prompts on English argumentative writing logic.

	Predictors	Outcomes	R^2	R^2_{change}	F_{change}	Sig. F_{change}	Durbin-Watson
Model 1	(Constant), Pre-writing scores	Post-writing scores	.555	.555	46.078	<.001**	n/a
Model 2	(Constant), Pre-writing scores, Total fixation counts on learner prompt area, Average fixation durations on learner prompt area		.658	.103	5.266	.010*	2.260

*p < .05; **p < .001.

Table 5
Parameter estimates of the effects of learner engagement in learner prompts on English argumentative writing logic.

	B	*SE (B)*	*β*	*t*	Sig.
(Constant)	30.151	7.099		4.247	<.001**
Pre-writing scores	.754	.098	.770	7.724	<.001**
Total fixation counts on learner prompts	.000	.002	−.028	−.258	.798
Average fixation durations on learner prompt area	−22.072	7.802	−.309	−2.829	.008*

*p < .05.

term retention.

4.2.2. Qualitative analysis results

The semi-structured interview findings aligned closely with our quantitative data, revealing diverse student perspectives on the impacts of engaging with GPT responses and learner prompts, as illustrated in Fig. 6. Regarding the development of logic knowledge, all interviewees recognised the value of GPT responses, highlighting in particular the clarity, comprehensibility, and precision of GPT-generated instructions. One interviewee commented, "The instructions and examples [in GPT responses] are clear, specific, and surprisingly easy-to-understand." Another interviewee said, "[The instructions in GPT responses] are really well organised with highted key words … Although logical concepts are pretty abstract, I found myself understanding the instruction very quickly." Nevertheless, two interviewees (10 %) considered *LogicalHamster*'s humour and encouragement to be "annoying" and "distracting," perceiving these features as detrimental to their learning efficiency. In terms of improving English argumentative writing logic, most interviewees (60 %) acknowledged the beneficial effects of GPT responses, especially the examples provided. However, some participants (40 %) felt these responses were "too simple and superficial" compared to real-world logical errors. "The authentic logical problems in my writing are far more subtle and complex than the examples." Quoted from one interviewee. Additionally, some interviewees (35 %) noted that while GPT responses were helpful in identifying logical errors, they lacked sufficient contextual guidance on addressing these errors in authentic writing contexts. For example, one interviewee admitted, "I understand the logical concepts from [GPT responses], while I don't know how to apply the knowledge in my authentic argumentative writings."

Eleven interviewees (55 %) considered their learner prompts to have a neutral impact on their improvement in English argumentative writing logic, arguing that "no new knowledge can be gained from [their] own output." Three participants (15 %) believed that generating prompts diverted attention and cognitive effort away from acquiring new knowledge. For example, one interviewee commented, "Producing prompts will exhaust my attentional effort that I can use to think more about the instructions and logic knowledge." Nonetheless, some interviewees acknowledged the benefits of prompt production for both developing logic knowledge (30 %) and improving English argumentative writing logic (45 %) by guiding them to comprehend and internalise the new knowledge and reflect their knowledge gaps. For example, one interviewee commented, "I understood the logical concepts better by

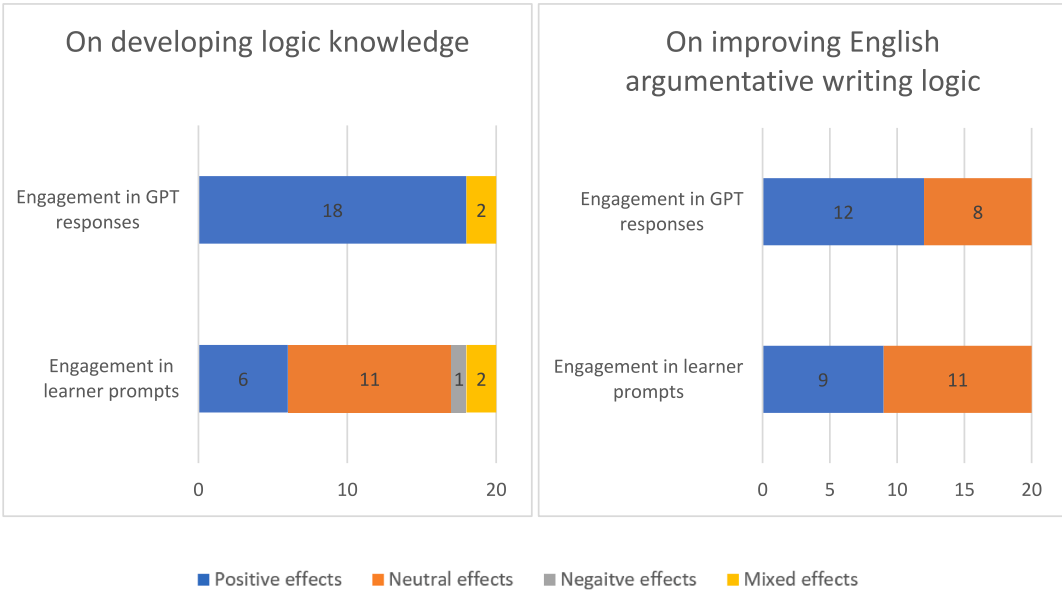

Fig. 6. Numbers of interviewees having different perceived effects of engaging in GPT responses and learner prompts

elaborating them [in prompt generation].” Another interviewee said, “I noticed some of my reasoning errors and logical misconceptions when typing down and thinking about my prompts.” These learners reported that engaging in discussions through prompt generation deepened their understanding by placing logic knowledge into meaningful and contextualised scenarios.

5. Discussion

5.1. RQ1. How do EFL learners engage with GPT responses and learner prompts in ChatGPT-based learning of English argumentative writing logic? Why?

Our quantitative and qualitative data converged to reveal varying levels of learner engagement in GPT responses and learner prompts in ChatGPT-based learning of English argumentative writing logic. Through thorough analysis, we identified several reasons for this diversity in engagement.

Firstly, learners exhibited high engagement with ChatGPT-based learning due to its perceived utility. Semi-structured interview results indicated that learners were particularly engaged with GPT responses, especially the instructions and examples, which they deemed beneficial for grasping and retaining new knowledge (“[GPT responses] are the only source of new knowledge, so I devoted most of my time and effort to them.” “I think the examples were very helpful in my understanding of [the target logic knowledge], so I read them repeatedly and focused on them.”). In contrast, many students found the jokes and encouragement within GPT responses (“… some *LogicalHamster*’s jokes and encouragement are not that useful, so I tended to ignore them.”) and learner prompts (“Close attention to my own output was unnecessary.”) less useful for enhancing their logic learning, leading to a deliberate reduction in their engagement with these parts. These findings align with sociocultural theory (Vygotsky & Cole, 1978; Elbably & Nemt-allah, 2024), which conceptualises learning as a socially mediated process facilitated by interaction with more capable peers or tools. In this instance, ChatGPT functions as a virtual “more capable peer,” delivering high-quality, discipline-specific instruction and feedback that foster sustained engagement across behavioural, cognitive, and emotional dimensions. Moreover, the results correspond with expectancy-value theory (Eccles & Wigfield, 2002), which posits that learners’ motivation is closely linked to their perception of the value and utility of learning materials. The prioritisation of instructional content in GPT responses over peripheral elements reflects learners’ strategic allocation of effort towards components they deem instrumental for success. These observations further corroborate recent findings by Kasneci et al. (2023) and Heung and Chiu (2025), who emphasise the critical role of high-quality GPT responses in enhancing learner engagement and facilitating effective language learning. Together, these theoretical perspectives underscore the importance of delivering targeted, meaningful input through AI-mediated instruction to maximise learner engagement.

Secondly, learners engaged deeply with ChatGPT-based learning due to its entertaining and supportive communication style. *LogicalHamster*’s use of humour, praise, encouragement, and emojis was found to elicit positive emotions such as enjoyment and comfort, as expressed by our interviewees: “*[LogicalHamster*’s*]* funny emojis and jokes made the learning process less boring.” Such positive emotions will enhance learner engagement. Similar effects have been documented in prior studies on ChatGPT-based learning, which highlight learners’ increased comfort and enjoyment attributed to ChatGPT’s nonjudgmental, friendly, and calm conversational style (Siregar et al., 2025; Skjuve et al., 2021). This underscores the critical role of effective communication skills in the design of educational chatbots (Pérez et al., 2020; Ruan et al., 2019). Given that students often experience anxiety and lack motivation in logic learning (Zhang et al., 2024), ChatGPT’s ability to foster positive emotions and motivation is particularly valuable.

Thirdly, some learners displayed low engagement with ChatGPT-based learning due to established learning habits. As reported in the semi-structured interviews, many students preferred passive lecture-based approaches, finding prompt production “tiring” and expressing reluctance to engage deeply with learner prompts. Similar trends have been observed in research on chatbot-based learning (Deng & Yu, 2023), highlighting the significant impact of established habits on learner engagement. Research on the relationship between learners’ established learning habits and engagement with ChatGPT-based learning remains scarce, suggesting a promising area for future investigation.

Additionally, negative attitudes towards ChatGPT may contribute to low engagement levels. In the semi-structured interviews, many students expressed doubts about ChatGPT’s ability to comprehend complex prompts or engage in in-depth conversations, leading them to undervalue careful communication with the chatbot. While the influence of learner attitudes on engagement is well-documented in technology-enhanced learning (e.g., Binhomran & Altalhab, 2021), this aspect is less explored in ChatGPT-based learning research, indicating a potential area for further exploration.

It is noteworthy that learners engaged with GPT responses and prompts in different ways, as evidenced by the eye-tracking data: participants exhibited frequent short fixations on GPT responses, while fixations on learner prompts were infrequent and of longer duration. This discrepancy can be attributed to the different actions associated with each component: typing for prompts and reading for GPT responses. Typing generally prolongs fixation duration compared to reading, resulting in longer average fixation durations on prompts than on GPT responses. Another explanation for the eye-tracking data, from a cognitive load theory perspective (Sweller, 2011), suggests that engaging with learner prompts imposes more extraneous load than GPT responses. This is supported by semi-structured interview transcripts, with one participant noting, “[The instructions in GPT responses] are well-organised with highlighted keywords … Although logical concepts are abstract, I found myself understanding the instruction quickly.” Another participant remarked, “Producing prompts will exhaust my attentional effort that I can use to think more about the instructions and logic knowledge.” These findings highlight the need to investigate learner engagement in GPT responses and learner prompts separately and underscore the importance of utilising mixed methods, such as eye-tracking data and semi-structured interview transcripts in our case, to comprehensively gauge learner engagement with ChatGPT-based learning.

5.2. RQ2. How do EFL learners’ engagement in GPT responses and learner prompts in ChatGPT-based learning of English argumentative writing logic influence their logic knowledge and logic in English argumentative writing? Why?

Our data, both quantitative and qualitative, converged to reveal the varied impacts of engagement in GPT responses and learner prompts within ChatGPT-based learning of English argumentative writing logic. Regarding GPT response content, results from PLS-SEM and multiple regression analyses revealed its significant effectiveness in enhancing logic knowledge, particularly for long-term retention. Semi-structured interview findings further corroborated these results, with all interviewees acknowledging the positive impact of GPT responses on their logic knowledge, emphasising the clarity, comprehensibleness, and precision of the instructional content (“The instructions and examples [in GPT responses] are clear, specific, and surprisingly easy to understand.”). Our findings resonate with the noticing hypothesis (Schmitt and Schmitt, 2020), which emphasises that deep engagement with input is essential for facilitating the noticing and development of language knowledge. These results also align with sociocultural theory (Vygotsky & Cole, 1978), which posits that engaging with a “more knowledgeable peer”—in this case, ChatGPT within ChatGPT-based learning environments—provides personalised input and scaffolding that facilitate

learners' cognitive development. Additionally, Krashen's (1992) Input Hypothesis underpins the importance of comprehensible input slightly beyond the learner's current competence. The well-structured and clear GPT responses serve as such input, facilitating the integration of new logical knowledge into existing cognitive schemas. This synergistic theoretical framework underscores how GPT responses, by providing optimally challenging and scaffolded input, promotes effective learning of English argumentative writing logic.

Regarding the presentation of instructional content in GPT responses, our semi-structured interviews revealed that GPT responses effectively presented logic knowledge systematically and concisely, promoting efficient knowledge processing ("[The instructions in GPT responses] are well-organised with highlighted keywords … Although logical concepts are abstract, I found myself understanding the instruction quickly."). Eye-tracking data echoed these findings, showing that learners frequently fixated on GPT responses for shorter durations, indicating efficient information processing. Interpreted through the lens of cognitive load theory (Sweller, 2011), these findings indicate that ChatGPT's instructional design effectively reduces extraneous cognitive load—the unnecessary mental effort imposed by poorly organised or redundant information. By minimising this extraneous load, learners can allocate greater working memory resources to germane cognitive load, which involves the meaningful processing, understanding, and application of logical concepts. Consequently, ChatGPT's concise and well-structured responses appear to support deeper learning by optimising the cognitive resources available for mastering argumentative writing logic.

Given the high quality and comprehensible presentation of GPT content, ChatGPT emerges as a valuable tool for developing logic knowledge, alleviating concerns regarding its capability to teach domain-specific knowledge (Baidoo-Anu & Ansah, 2023). Therefore, we recommend that practitioners leverage this technology in developing logic knowledge in English argumentative writing.

However, deep engagement in GPT responses demonstrated limited effectiveness in enhancing English argumentative writing logic, as indicated by our PLS-SEM results. Semi-structured interviews revealed that the learning materials from GPT responses were often perceived as "too simple" and lacking depth and context ("The authentic logical problems in my writing are far more subtle and complex than the examples."). This finding suggests that more in-depth, contextualised materials and professional guidance are essential for in-depth understanding of logic of authentic English argumentative writing, echoing many researchers' concerns about the limitations of GPT responses in developing specialised skills for professional requirements (Baidoo-Anu & Ansah, 2023). To improve ChatGPT-based learning in this direction, it may be necessary to provide comprehensive examples closely aligned with real-world writing and specific guidance on applying logic knowledge in practical contexts. Future research efforts should address these issues to enhance the effectiveness of ChatGPT-based learning of English argumentative writing logic.

In terms of engagement in learner prompts, our findings revealed mixed impacts on the development of logic knowledge and English argumentative writing logic. PLS-SEM and multiple regression analyses indicated that deep engagement in learner prompts did not significantly enhance logic knowledge and may, in fact, hinder the development of English argumentative writing logic. Supporting this, a majority of interviewees (70 %) reported neutral or negative impacts of deep engagement in learner prompts on their logic knowledge, while 55 % noted similar effects on their English argumentative writing logic. Participants described prompt engagement as effort-demanding and attention-consuming, diverting focus away from instructional materials and exhausting their cognitive resources for processing and constructing new knowledge and skills ("Producing prompts will exhaust my attentional effort that I can use to think more about the instructions and logic knowledge."). The infrequent long-duration fixations on prompts further indicate relatively low engagement efficiency. From the standpoint of cognitive load theory (Sweller, 2011), these findings imply that the process of generating and considering prompts may impose excessive extraneous cognitive load, thereby reducing overall learning efficiency. This interpretation aligns with Barcroft's (2006) Resource Depletion for Output hypothesis, which argues that the cognitive effort required to produce output—such as prompt generation—can exhaust limited cognitive capacity, constraining the learner's ability to assimilate new information effectively. Therefore, while learner prompts have the potential to foster deeper understanding, excessive or poorly structured engagement with prompt generation may inadvertently hinder learning by exhausting learners' cognitive resources. This underscores the need for carefully balancing the cognitive demands of productive activities within ChatGPT-based learning environments to optimise the development of both logic knowledge and English argumentative writing skills.

Nevertheless, engaging with learner prompts can yield positive outcomes, as indicated by our semi-structured interviews. Some learners found that considering and generating prompts effectively facilitated their understanding and internalisation of logic knowledge ("I understood the logical concepts better by elaborating them [in prompt generation]."). From the perspective of cognitive load theory (Sweller, 2011), this process likely induces germane cognitive load—the mental effort devoted to schema construction and automation—thereby enhancing the integration of new knowledge into long-term memory. Moreover, several learners described how engaging with prompts prompted metacognitive monitoring and reflection on their own reasoning processes and knowledge gaps ("I noticed some of my reasoning errors and logical misconceptions when typing down and thinking about my prompts."). This aligns with frameworks of self-regulated learning (Zhang & Zou, 2024; Zimmerman, 2013), wherein active reflection, goal-setting, and self-monitoring are central to effective learning. By encouraging learners to articulate and evaluate their thought processes through prompt generation, ChatGPT-based interactions appear to scaffold key aspects of self-regulation. Together, these theoretical perspectives suggest that engagement in learner prompts not only elicits essential germane cognitive load but also fosters metacognitive awareness and regulatory skills, both of which are critical for deep and sustained learning.

It is noteworthy that engagement with neither learner prompts nor GPT responses demonstrated statistically significant positive impacts on the development of English argumentative writing logic, as shown in the PLS-SEM results. This may be explained by Bloom's taxonomy of learning objectives (Bloom & Krathwohl, 2020), which indicates that applying knowledge is more cognitively complex than mere comprehension and memorisation, requiring greater effort. Our findings suggest that enhancing English argumentative writing logic is inherently challenging, and that simply recalling and understanding logical concepts may not suffice for effectively applying that knowledge to improve the logic quality of authentic English argumentative writing. In developing difficult language skills like writing, null and negative results are common in technology-enhanced learning (Zhang et al., 2023a). To gain a more comprehensive understanding of the effects of engagement in learner prompts and GPT responses in ChatGPT-based learning on English argumentative writing logic, prolonged intervention may be necessary, following the recommendations of Zhang et al. (2023a).

6. Implications

The study demonstrated the potential of ChatGPT in fostering deep engagement and effectively developing logic knowledge and argumentative writing logic among upper-intermediate to advanced EFL learners. Given the recognised importance of logical knowledge and skills in argumentative writing (Murray, 2012) and the frequent challenges EFL learners face in this area (El Khoiri & Widiati, 2017), our proposed model of ChatGPT-based learning holds considerable promise in EFL education. Practitioners may thus benefit from employing our

model to enhance upper-intermediate to advanced EFL learners' argumentative writing proficiency.

For future practitioners of ChatGPT-based learning of English argumentative writing logic, we have further implications. Firstly, we recommend a tailored approach, taking into account learners' prior learning experiences and attitudes towards ChatGPT. Our study indicated that students with passive learning habits and negative attitudes towards chatbots were less engaged with ChatGPT-based learning, leading to reduced learning outcomes. Therefore, conducting pre-intervention surveys to identify learners' experiences and attitudes, as suggested by Zhang et al. (2023a), is recommended. Based on survey results, educators can provide targeted training and orientation sessions to improve students' understanding and acceptance of ChatGPT-based learning, so as to enhance their understanding and acceptance of this learning method.

Secondly, long-term engagement with ChatGPT-based learning may be necessary to achieve statistically significant improvements in the logical quality of English argumentative writing. Although our evidence supports its effectiveness in developing logical knowledge, our statistical analyses indicated that deep engagement with neither learner prompts nor GPT responses yielded significant gains in learners' logical writing quality after a single session. Interview data suggest that applying logical knowledge to authentic English argumentative writing is highly cognitively demanding and requires sustained guidance and scaffolding. Therefore, to achieve meaningful and statistically significant improvements in students' argumentative writing logic, we recommend that future practitioners implement ChatGPT-based learning interventions over an extended timeframe. Initial sessions might focus on delivering targeted instruction and exercises to help novices develop and internalise foundational logical concepts and structures. Subsequently, ChatGPT could be configured to assign tasks with more cognitive complexity, such as analysis, evaluation, and construction of English argumentative texts from a logical perspective, guiding learners to apply their logical knowledge in authentic contexts.

Thirdly, our findings emphasise the importance of prompt engineering. Although ChatGPT has been recognised as useful in educational contexts, concerns remain regarding its efficacy in delivering discipline-specific knowledge and developing professional skills (Baidoo-Anu & Ansah, 2023). To address these concerns, prompt engineering, involving curated examples, knowledge bases, predefined responses, and detailed instructions, can significantly enhance the quality, depth, and contextual relevance of ChatGPT-generated learning materials (Lee et al., 2023). Moreover, assigning engaging conversational personas and humorous tones can elicit positive emotions, further boosting learner engagement. In this study, we applied prompt engineering in designing *LogicalHamster*, a GPT-4-powered chatbot tasked with analysing learners' logical arguments and providing personalised feedback. To enhance engagement, *LogicalHamster* adopted a humorous conversational style with a friendly and approachable "hamster" persona. According to our semi-structured interview findings, such tailored communication evoked positive emotions, such as enjoyment and comfort, which, in turn, enhance learner engagement.

Fourthly, learners should strategically allocate their engagement with prompts and GPT response during ChatGPT-based learning of English argumentative writing logic, particularly during the initial stages. Developing proficiency in English argumentative writing logic is highly cognitively demanding, requiring intensive critical and abstract thinking (El Khoiri & Widiati, 2017). At the early stages, it is essential for learners to focus on understanding and internalising fundamental logical concepts (Oaksford & Chater, 2020). Our findings indicate that deep engagement in learner prompts can introduce excessive extraneous cognitive load, thereby depleting the cognitive resources necessary for processing new logical knowledge. Therefore, we recommend that practitioners assess cognitive load and strategically plan how to allocate attention across different components of ChatGPT-based learning when implementing this approach. For novice learners or those with relatively limited cognitive capacity, we suggest prioritising attention to GPT responses for understanding and internalising logical knowledge rather than focusing too heavily on prompt production.

The theoretical implications of this study are also substantial. Firstly, our findings reinforce both sociocultural theory (Vygotsky & Cole, 1978) and the Input Hypothesis (Krashen, 1992) within ChatGPT-based learning environments. The significant positive effects of engagement in GPT responses on learning outcomes, particularly in developing logic knowledge, indicate that learners benefit from interactive, supportive communication with ChatGPT in understanding and internalising complex logical concepts. This illustrates that ChatGPT can serve as a "more capable peer" (Vygotsky & Cole, 1978, p. 86), facilitating natural and constructive dialogues with EFL learners in developing English argumentative writing. Our study not only supports previous applications of these theories in ChatGPT-based learning (Yang & Li, 2024) but also extends them into the realms of logic education and argumentative writing. Future research may focus on designing, implementing, and analysing ChatGPT-based learning for logic and English argumentative writing through the lenses of sociocultural theory and the Input Hypothesis.

Secondly, our study contributes to cognitive load theory by revealing the different impacts of GPT responses and learner prompts on cognitive load and learning outcomes within ChatGPT-based learning environments. As evidenced by our eye-tracking data, test results, and semi-structured interviews, GPT responses present logical learning materials in an easily digestible manner, facilitating the development of logic knowledge while minimising extraneous cognitive load. Meanwhile, engagement in learner prompts is often perceived as "tiring" but effective in prompting the construction and integration of new knowledge. This finding suggests that learner prompts can impose significant extraneous and germane cognitive demands. The distinct impacts of GPT responses and learner prompts underscore the necessity of investigating these components separately in ChatGPT-based learning. We recommend that educators design ChatGPT-based tasks that effectively balance cognitive loads with reference to our findings.

Thirdly, our study reinforces self-regulated learning theory (Zimmerman, 2013) in the context of ChatGPT-based learning for English argumentative writing logic. Semi-structured interview results indicate that engagement with both GPT responses and learner prompts can stimulate self-regulated learning processes: learners reported feeling more motivated to seek additional materials and ask questions when receiving encouragement and praise from GPT responses, and they tended to elaborate on their knowledge and reflect on gaps when generating prompts. These findings not only echo previous research on the relevance of self-regulated learning theory in ChatGPT-based environments (e.g., Zhu, 2025) but also enrich the framework of ChatGPT-based, self-regulated learning by highlighting the different roles of GPT responses and learner prompts. Future researchers are encouraged to foster self-regulation in ChatGPT-based learning through strategic engagement with both components.

7. Conclusion and limitations

This paper presents a mixed-method study into learner engagement with the two main components of ChatGPT-based learning of English argumentative writing logic—GPT responses and learner prompts—and its consequent effects. Findings indicate that learners engaged deeply with both components—particularly GPT responses—attributed to the perceived usefulness of the content and ChatGPT's engaging and supportive communication style. Deep engagement in GPT responses positively influenced the development of logical knowledge, especially regarding long-term retention, owing to the high-quality input and the systematic, concise presentation of information. Conversely, extensive engagement in learner prompts may impose excessive extraneous cognitive load, potentially hindering learners' progress in English argumentative writing logic. Neither deep engagement in GPT responses

nor learner prompts yielded statistically significant improvements in logical writing skills, likely due to the limited learning duration and the high cognitive demands of applying logical knowledge to authentic writing tasks. Drawing on sociocultural theory, cognitive load theory, self-regulated learning theory, the Input Hypothesis, the Noticing Hypothesis, and the Resource Depletion for Output Hypothesis, we examine the factors underlying the dynamics of learner engagement in ChatGPT-based learning for English argumentative writing logic.

While this study provides valuable insights into learner engagement with ChatGPT-based learning, future researchers and practitioners should pay particular attention to the scope within which our findings can be appropriately interpreted and applied. Firstly, *LogicalHamster*, the GPT-4-powered chatbot utilised in this study, was specifically developed to teach EFL argumentative writing logic and engage learners. It achieves the goals effectively by delivering targeted logical knowledge, analysing learners' submitted arguments, providing personalised feedback, and engaging learners through a humorous and conversational interaction style. As *LogicalHamster* incorporates fundamental interaction components common to most ChatGPT-based learning tools (see Lestari, 2024), our study specifically focused on learner engagement with ChatGPT-based learning at a micro-level, rather than examining *LogicalHamster* itself. Our results hold implications for other ChatGPT-based learning contexts, despite variations in specific GPT training methods or interface designs. Nonetheless, we admit ChatGPT-based instructional tools may differ across disciplines, GPT model versions, and prompt-engineering approaches (Lee et al., 2023). Consequently, caution must be exercised when applying our findings to other disciplines or chatbot designs. Practitioners should remain mindful of potential differences between ChatGPT versions and operations when referring to our findings for implementing ChatGPT-based learning of English argumentative writing logic. Further research exploring learner engagement with ChatGPT-based learning across diverse academic, GPT model versions, and instructional contexts is therefore encouraged.

Secondly, our participant sample comprised Chinese undergraduate and postgraduate learners with upper-intermediate to advanced EFL proficiency. Considering the potential impact of learner factor in language learning contexts (Arslanyilmaz, 2012), our findings may not be directly generalisable to learners from other educational settings, proficiency levels, or cultural backgrounds. Moreover, the notable gender imbalance in our sample, with significantly more female than male participants, may have affected learner engagement patterns and outcomes. Practitioners and future researchers should thus exercise caution when transferring our conclusions to different learner populations. Further studies involving more gender-balanced samples, and encompassing a broader range of proficiency levels and educational backgrounds, are recommended to enhance the robustness and generalisability of research on ChatGPT-based learning.

Thirdly, although our sample size ($N = 42$) meets commonly cited heuristic thresholds—specifically, at least 30 observations for paired-sample t-tests, a minimum of 10 observations per predictor for multiple linear regression, and ten times the largest number of indicators per latent construct for PLS-SEM—it remains relatively modest. Hence, concerns remain regarding the statistical power and robustness of our findings. Additionally, we quantified learner engagement in ChatGPT-based learning through eye-tracking metrics and assessed learning outcomes via test and writing scores, analysing their interrelations statistically. However, eye-tracking primarily captures observable behavioural engagement (Latif, 2019), whereas learner engagement is a multidimensional construct that may not be fully externalised through visible behaviours (Philp & Duchesne, 2016). For instance, deep emotional and cognitive engagement can remain unobservable and thus may not have been adequately reflected in our models. Consequently, our statistical findings are likely to represent behavioural engagement more strongly than other dimensions. Researchers and practitioners should therefore exercise caution when generalising our statistical analysis results to broader or more complex contexts. To obtain a more comprehensive and reliable understanding of learner engagement in ChatGPT-based learning of English argumentative writing logic, future studies should employ larger samples and integrate quantitative measures of unobservable engagement—such as emotional and cognitive dimensions assessed via surveys—alongside behavioural indicators within their analytical frameworks.

Additional limitations should also be acknowledged. Firstly, we collected eye-movement data using a Tobii X2-30 eye tracker. Although it is a top-ranked device, the accuracy and precision of this data collection method could be further enhanced. Moreover, our research focused on a single ChatGPT-based learning session. Given that logic knowledge and argumentative writing proficiency typically require sustained practice, future studies should investigate long-term interventions to gain a fuller understanding of the effectiveness of ChatGPT-based learning. Thirdly, our findings highlight the critical roles of task complexity and learner cognitive load in influencing the dynamics and effectiveness of ChatGPT-based learning of English argumentative writing logic, echoing previous argument (e.g., El Khoiri & Widiati, 2017). However, this factor was not incorporated into our models. Future research should explicitly measure cognitive load and examine its interaction with learner engagement and learning outcomes within this context, aiming to obtain a more comprehensive understanding of this learning method.

Finally, as ChatGPT gains prominence across various educational domains, our findings on its application in developing logic knowledge and skills in English argumentative writings offer valuable insights into its process, mechanism, and effects. Given the limited existing research on the mechanisms underlying ChatGPT-based learning, further studies examining learner behaviours and outcomes across diverse contexts and disciplines are highly encouraged.

CRediT authorship contribution statement

Ruofei Zhang: Writing – review & editing, Writing – original draft, Methodology, Investigation, Data curation, Conceptualization. **Di Zou:** Writing – review & editing, Supervision, Resources, Methodology, Investigation, Conceptualization. **Haoran Xie:** Writing – review & editing, Validation, Supervision, Resources, Conceptualization. **Fu Lee Wang:** Writing – review & editing, Visualization, Validation, Supervision, Resources, Conceptualization.

Data availability statement

The participants of this study did not give written consent for their data to be shared publicly, so due to the sensitive nature of the research supporting data is not available.

Ethics declarations

I confirm that all the research meets ethical guidelines and adheres to the legal requirements of the study country.

Funding

The work described in this paper was partially supported by the Teaching Development and Language Enhancement Grant of University Grant Committee, The Hong Kong Polytechnic University (Project No. PolyU/TDLEG25-28/IICA/P/02) and the Start-up Fund for New Recruits of The Hong Kong Polytechnic University (Project No. P0056518).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Logic knowledge test (Zhang et al., 2024)

Seven common logical fallacies in English argumentative writings		
A. Begging the Question	B. Red Herring	C. False Alternatives
D. Post hoc fallacy	E. Hasty Generalization	F. Slippery Slope
G. Faulty Analogy		

In *Column A* of the following table, there list seven sentences from English argumentative writings, each committing one of the above seven types of logical fallacies. Please complete the following table by:

1. Explain the meanings and structures of the fallacies of each sentence in *Column B*.
2. Match the terminology of the fallacies of each sentence in *Column C*.

[A] Example sentences	[B] Meanings and structures of the fallacy in the sentence	[C] Terminology of the fallacy
1 All the world's top technical companies are working on the development of AI technology because AI technology is the hottest topic in this field.		
2 Either we allow cars in the city centre, or the local businesses will suffer.		
3 Students' learning efficiency significantly declined after they had online learning. Hence, online learning has negative effects on students' learning efficiency.		
4 If we do not solve the problem of air pollution today, the problem may continue being delayed day after day. In the end, this problem will never be solved.		
5 To say humans are immortal is like saying a car can run forever.		
6 According to our semi-structured interviews of 200 students in Hong Kong, Asian students are overall reluctant to ask their teachers for help.		
7 It is true that the quality of second language education is not good in our country. However, the native language education is a bigger problem and in more need of our attention.		

Appendix B. Scoring system of the tests on logic knowledge (Zhang et al., 2024)

Categories		The score for each type of logical fallacy
Fallacy meaning	The student does not provide any correct information concerning the structure and meaning of the logical fallacy in the example sentence.	0
	The student provides some correct but incomplete, inaccurate information concerning the structure and meaning of the logical fallacy in the example sentence.	1
	The student provides correct, complete, clear explanations for the structure and meaning of the logical fallacy in the example sentence.	2
Fallacy terms	The student matches the meaning and term of the fallacy correctly.	0
	The student fails to match the meaning and term of the fallacy.	1
Total		3

Appendix C. Scoring system of the logic in English essays (adapted from Finken and Ennis's Illinois Critical Thinking Essay Scoring Rubric and IELTS's academic writing assessment criteria)

Categories		Scores
Supporting Reasons	The accuracy, relevance and sufficiency of the evidence students used to support their arguments	0–25
Reasoning	The clarity of arguments and convincingness of reasoning	0–25
The commitment of reasoning errors	The frequencies and depths of reasoning errors	0–25
Coherence	The linking of ideas through logical sequencing throughout the students' writings	0–25
Total		0–100

Appendix D. Indicators of the validity and reliability of the PLS-SEM analysis

Table 1
Indicators for the reliability and validity of the PLS-SEM measurement models

Variables	Indicators	Factor loadings	rhoC	AVE
Engagement in GPT responses	Total fixation counts on GPT response area	.986	.722	.594
	Average fixation durations on GPT response area	.466		
Engagement in learner prompts	Total fixation counts on learner prompts	.839	.832	.712
	Average fixation durations on learner prompt area	.848		
Learners' prior logic knowledge	Pre-test scores of logic knowledge	1.000	1.000	1.000
Learners' prior English argumentative writing logic	Pre-writing scores	1.000	1.000	1.000
ChatGPT-based learning of English argumentative writing logic outcomes concerning logic knowledge	Immediate post-test scores of logic knowledge	.943	.936	.880
	Delayed post-test scores of logic knowledge	.934		
ChatGPT-based learning of English argumentative writing logic outcomes concerning English argumentative writing logic	Post-writing scores	1.000	1.000	1.000

Table 2
Indicators for the collinearity of the PLS-SEM structural model

Endogenous constructs	Exogenous constructs	
	ChatGPT-based learning of English argumentative writing logic outcomes concerning logic knowledge	ChatGPT-based learning of English argumentative writing logic outcomes concerning English argumentative writing logic
Engagement in GPT responses	1.924	1.943
Engagement in learner prompts	1.403	1.332
Learners' prior logic knowledge	1.118	n/a
Learners' prior English argumentative writing logic	n/a	1.037

References

Annamalai, N., & Nasor, M. (2025). Exploring ChatGPT in education: Unveiling learners' experiences through the lens of self-determination theory. *Smart Learning Environments, 12*(1), 59. https://doi.org/10.1186/s40561-025-00393-2

Annamalai, N., Uthayakumaran, A., Bervell, B., & Kumar, R. (2025). Examining the use of ChatGPT in argumentative writing: A mixed methods study. *Journal of Advanced Academics.* , Article 1932202X251333859. https://doi.org/10.1177/1932202X251333859

Arslanyilmaz, A. (2012). An online task-based language learning environment: Is it better for advanced-or intermediate-level second language learners? *Turkish Online Journal of Educational Technology-TOJET, 11*(1), 20–35. https://eric.ed.gov/?id=EJ976564.

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI, 7*(1), 52–62. Retrieved from https://dergipark.org.tr/en/pub/jai/issue/77844/1337500.

Barcroft, J. (2006). Can writing a new word detract from learning it? More negative effects of forced output during vocabulary learning. *Second Language Research, 22*(4), 487–497. https://doi.org/10.1191/0267658306sr276oa

Binhomran, K., & Altalhab, S. (2021). The impact of implementing augmented reality to enhance the vocabulary of young EFL learners. *JALT CALL Journal, 17*(1), 23–44. https://eric.ed.gov/?id=EJ1313360.

Bloom, B. S., & Krathwohl, D. R. (2020). *Taxonomy of educational objectives: The classification of educational goals. Book 1, cognitive domain.* Longman.

British Council. *Understanding and explaining IELTS scores*. IELTS. Retrieved October 24, 2025, from https://takeielts.britishcouncil.org/teach-ielts/test-information/ielts-scores-explained.

Brown, T. B. (2020). Language models are few-shot learners. *arXiv preprint arXiv: 2005.14165*. https://doi.org/10.48550/arXiv.2005.14165.

Bygate, M., & Samuda, V. (2009). Creatin gpressure in task pedagogy: The joint roles of field, purpose, and engagement within the interaction approach. In A. Mackey, & C. Polio (Eds.), *Multiple perspectives on interaction: Second language research in honour of Susan M. Gass* (pp. 90–116). New York, NY: Taylor and Francis/Routledge.

Byrne, B. M. (2010). *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming.* Taylor and Francis Group Publication.

Chang, D. H., Lin, M. P. C., Hajian, S., & Wang, Q. Q. (2023). Educational design principles of using AI chatbot that supports self-regulated learning in education: Goal setting, feedback, and personalization. *Sustainability, 15*(17), Article 12921. https://doi.org/10.3390/su151712921

Darmawansah, D., Rachman, D., Febiyani, F., & Hwang, G. J. (2025). ChatGPT-supported collaborative argumentation: Integrating collaboration scripts and argument mapping to enhance EFL students' argumentation skills. *Education and Information Technologies, 30*(3), 3803–3827. https://doi.org/10.1007/s10639-024-12986-4

Deng, X., & Yu, Z. (2023). A meta-analysis and systematic review of the effect of chatbot technology use in sustainable education. *Sustainability, 15*(4), 2940. https://doi.org/10.3390/su15042940

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*(1), 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

El Khoiri, N., & Widiati, U. (2017). Reasoning errors in Indonesian EFL learners' argumentative writing: Students' perspectives. *Dinamika Ilmu, 17*(1), 71–81. https://doi-org.ezproxy.eduhk.hk/10.21093/di.v17i1.638.

Elbably, Y., & Nemt-allah, M. (2024). Grand challenges for ChatGPT usage in education: Psychological theories, perspectives and opportunities. *Psychology Research on Education and Social Sciences, 5*(2), 31–36. https://doi.org/10.5281/zenodo.12601568

Ellis, N. C. (2019). Essentials of a theory of language cognition. *The Modern Language Journal, 103*, 39–60. https://doi.org/10.1111/modl.12532

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed). London: SAGE Publications Ltd.

Ge, X., & Ifenthaler, D. (2018). Designing engaging educational games and assessing engagement with game-based learning. In I. Management Association (Ed.), *Gamification in education: Breakthroughs in research and practice* (pp. 1–19). IGI Global.

Ghafouri, M., Hassaskhah, J., & Mahdavi-Zafarghandi, A. (2024). From virtual assistant to writing mentor: Exploring the impact of a ChatGPT-based writing instruction protocol on EFL teachers' self-efficacy and learners' writing skill. *Language Teaching Research.* , Article 13621688241239764. https://doi.org/10.1177/13621688241239764

Gobert, J. D., Baker, R. S., & Wixon, M. B. (2015). Operationalizing and detecting disengagement within online science microworlds. *Educational Psychologist, 50*(1), 43–57. https://doi.org/10.1080/00461520.2014.999919

Guo, Y., & Lee, D. (2023). Leveraging ChatGPT for enhancing critical thinking skills. *Journal of Chemical Education, 12*, 4876–4883. https://doi.org/10.1021/acs.jchemed.3c00505

Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). *Partial least squares structural equation modeling (PLS-SEM) using R: A workbook* (p. 197). Springer Nature.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review, 31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203.

Heung, Y. M. E., & Chiu, T. K. (2025). How ChatGPT impacts student engagement from a systematic review and meta-analysis study. *Computers and Education: Artificial Intelligence, 8*, Article 100361. https://doi.org/10.1016/j.caeai.2025.100361

Hirvela, A. (2013). Preparing English language learners for argumentative writing. In L. C. de Oliveira, & T. Silva (Eds.), *L2 writing in secondary classrooms* (pp. 67–86). New York: Routledge.

Hiver, P., Al-Hoorie, A. H., Vitta, J. P., & Wu, J. (2021). Engagement with language learning: A systematic review of 20 years of research methods and definitions. *Language Teaching Research.* , Article 13621688211001289. https://doi.org/10.1177/13621688211001289

Hundleby, C. (2010). The authority of the fallacies approach to argument evaluation. *Informal Logic, 20*(3), 279–308. https://scholar.uwindsor.ca/cgi/viewcontent.cgi?article=1007&context=philosophypub.

Jiang, F., & Hyland, K. (2024). Does ChatGPT argue like students? Bundles in argumentative essays. *Applied Linguistics.* , Article amae052. https://doi.org/10.1093/applin/amae052

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Khampusaen, D. (2024). The impact of ChatGPT on academic writing skills and knowledge: An investigation of its use in argumentative essays. *LEARN Journal: Language Education and Acquisition Research Network, 18*(1), 963–988. https://doi.org/10.70730/PGCQ9242

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal.* , Article 00336882231162868. https://doi.org/10.1177/00336882231162868

Krashen, S. (1982). *Principles and practice in second language acquisition*. New York: Prentice-Hall International.

Krashen, S. D. (1992). The input hypothesis: An update. In J. E. Alatis (Ed.), *Linguistics and language pedagogy: The state of the art* (pp. 409–431). Georgetown University Press.

Kucuk, T. (2024). ChatGPT integrated grammar teaching and learning in EFL classes: A study on Tishk international university students in Erbil, Iraq. *Arab World English Journal*, 100–111. http://eprints.tiu.edu.iq/id/eprint/1507.

Latif, M. M. A. (2019). Eye-tracking in recent L2 learner process research: A review of areas, issues, and methodological approaches. *System, 83*, 25–35. https://doi.org/10.1016/j.system.2019.02.008

Lee, E. Y., Il, N. G. D., An, G. H., Lee, S., & Lim, K. (2023). ChatGPT-Based debate game application utilizing prompt engineering. In *Proceedings of the 2023 international conference on research in adaptive and convergent systems* (pp. 1–6). https://doi.org/10.1145/3599957.3606244

Lestari, S. (2024). The implementation of ChatGPT-based learning for higher education in Indonesia: Systematic literature review. *ELS-JISH, 7*(2). https://doi.org/10.34050/elsjish.v7i2.35265

Lismay, L. (2020). Logical fallacies on students' English argumentative writing. *Journal of English Language Pedagogy), 5*(2), 21–27. https://doi.org/10.36665/elp.v5i2.321

Murray, N. (2012). *Writing essays in English language and linguistics: Principles, tips and strategies for undergraduates*. Cambridge University Press.

Nejmaoui, N. (2019). Improving EFL learners' critical thinking skills in argumentative writing. *English Language Teaching, 12*(1), 98–109. https://eric.ed.gov/?id=EJ1200049.

Oaksford, M., & Chater, N. (2020). New paradigms in the psychology of reasoning. *Annual Review of Psychology, 71*(1), 305–330. https://doi.org/10.1146/annurev-psych-010419-051132

Obaidellah, U., Al Haek, M., & Cheng, P. C. H. (2018). A survey on the usage of eye-tracking in computer programming. *ACM Computing Surveys, 51*(1), 1–58. https://doi.org/10.1145/3145904

Oktawirawan, D. H. (2022). Chatbot is my buddy: The role of chatbot as human friend in daily life. In *Proceedings of international conference on psychological studies (ICPSYCHE), 3* pp. 35–41). https://proceeding.internationaljournallabs.com/index.php/picis/index.

Paas, F., & Sweller, J. (2014). *Implications of cognitive load theory for multimedia learning. The Cambridge handbook of multimedia learning* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9781139547369, 4.

Pérez, J. Q., Daradoumis, T., & Puig, J. M. M. (2020). Rediscovering the use of chatbots in education: A systematic literature review. *Computer Applications in Engineering Education, 28*(6), 1549–1565. https://doi-org.ezproxy.eduhk.hk/10.1002/cae.22326.

Philp, J., & Duchesne, S. (2016). Exploring engagement with tasks in the language classroom. *Annual Review of Applied Linguistics, 36*, 50–72. https://doi.org/10.1017/S0267190515000094

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist, 50*(4), 258–283. https://doi.org/10.1080/00461520.2015.1122533

Que, Y., Zheng, Y., Hsiao, J. H., & Hu, X. (2023). Studying the effect of self-selected background music on reading task with eye movements. *Scientific Reports, 13*(1), 1704. https://doi.org/10.1038/s41598-023-28426-1

Ruan, S., Jiang, L., Xu, J., Tham, B. J. K., Qiu, Z., Zhu, Y.Landay, J. A., … (2019). Quizbot: A dialogue-based adaptive learning system for factual knowledge. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1–13). https://doi.org/10.1145/3290605.3300587. Glasgow, Scotland, UK.

Saidi, M. (2020). The relationship between Iranian EFL learners' linguistic and logical intelligences and the frequency of fallacies and evidence in their argumentative writing: A gender-based study. *Journal of English Language Pedagogy and Practice, 12*(25), 151–169. https://doi.org/10.30495/JAL.2020.675547

Schmitt, N., & Schmitt, D. (2020). *Vocabulary in language teaching*. Cambridge University Press.

Schreiner, M. (2023). GPT-4 architecture, datasets, costs and more leaked. *The Decoder, 11*. https://the-decoder.com/gpt-4-architecture-datasets-costs-and-more-leaked/.

Siregar, J. L., Sitorus, S. T., & Hartati, R. (2025). Matching communication with ChatGPT: An innovative solution to maintain harmonious relationships between close friends. *International Journal Of Humanities, Social Sciences And Business (INJOSS), 4*(3), 343–356. https://injoqast.net/index.php/JOSSS/article/view/274.

Skjuve, M., Følstad, A., Fostervold, K. I., & Brandtzaeg, P. B. (2021). My chatbot companion– a study of human-chatbot relationships. *International Journal of Human-Computer Studies, 149*, 102601. https://doi.org/10.1016/j.ijhcs.2021.102601.

Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology, 14*, Article 1260843. https://doi.org/10.3389/fpsyg.2023.1260843

Song, Y., & Sparks, J. R. (2019). Measuring argumentation skills through a game-enhanced scenario-based assessment. *Journal of Educational Computing Research, 56*(8), 1324–1344. https://doi.org/10.1177/0735633117740605

Soodmand Afshar, H., Movassagh, H., & Radi Arbabi, H. (2017). The interrelationship among critical thinking, writing an argumentative essay in an L2 and their subskills. *Language Learning Journal, 45*(4), 419–433. https://doi.org/10.1080/09571736.2017.1320420

Stroud, R. (2014). Student engagement with learning vocabulary with CALL. In S. Jager, L. Bradley, E. J. Meima, & S. Thouësny (Eds.), *Proceedings of the 2014 EUROCALL conferenceCALL design: Principles and practice* (pp. 340–344). https://doi.org/10.14705/rpnet.2014.000242. Groningen, The Netherlands.

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing, 57*, Article 100752. https://doi.org/10.1016/j.asw.2023.100752

Sweller, J. (2011). Cognitive load theory. *Psychology of Learning and Motivation, 55*, 37–76. Academic Press.

Teng, M. F. (2024a). A systematic review of ChatGPT for English as a foreign language writing: Opportunities, challenges, and recommendations. *International Journal of TESOL Studies, 6*(3), 36–57. https://doi.org/10.58304/ijts.20240304

Teng, M. F. (2024b). "ChatGPT is the companion, not enemies": EFL learners' perceptions and experiences in using ChatGPT for feedback in writing. *Computers and Education: Artificial Intelligence, 7*, Article 100270. https://doi.org/10.1016/j.caeai.2024.100270

Tian, L., & Zhou, Y. (2020). Learner engagement with automated feedback, peer feedback and teacher feedback in an online EFL writing context. *System, 91*, Article 102247. https://doi.org/10.1016/j.system.2020.102247

Vygotsky, L. S., & Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard university press.

Wang, Y., Lu, S., & Harter, D. (2021). Multi-sensor eye-tracking systems and tools for capturing student attention and understanding engagement with learning: A review. *IEEE Sensors Journal, 21*(20), 22402–22413. https://doi.org/10.1109/JSEN.2021.3105706

Woo, D. J., Wang, D., Guo, K., & Susanto, H. (2024). Teaching EFL students to write with ChatGPT: Students' motivation to learn, cognitive load, and satisfaction with the learning process. *Education and Information Technologies*, 1–28. https://doi.org/10.1007/s10639-024-12819-4

Wu, C. H., & Ho, V. T. (2025). Critical factors for why ChatGPT enhances learning engagement and outcomes. *Education and Information Technologies*, 1–32. https://doi.org/10.1007/s10639-025-13346-6

Yang, L., & Li, R. (2024). ChatGPT for L2 learning: Current status and implications. *System, 124*, Article 103351. https://doi.org/10.1016/j.system.2024.103351

Yıldız, T. A. (2023). The impact of ChatGPT on language learners' motivation. *Journal of Teacher Education and Lifelong Learning, 5*(2), 582–597. https://doi.org/10.51535/tell.1314355

Yu, Y. T., & Tsuei, M. (2023). The effects of digital game-based learning on children's Chinese language learning, attention and self-efficacy. *Interactive Learning Environments, 31*(10), 6113–6132. https://doi.org/10.1080/10494820.2022.2028855.

Zhang, R., & Zou, D. (2024). Self-regulated second language learning: A review of types and benefits of strategies, modes of teacher support, and pedagogical implications. *Computer Assisted Language Learning, 37*(4), 720–765. https://doi.org/10.1080/09588221.2022.2055081.

Zhang, R., Zou, D., & Cheng, G. (2024). Chatbot-based learning of logical fallacies in EFL writing: Perceived effectiveness in improving target knowledge and learner motivation. *Interactive Learning Environments, 32*(9), 5552–5569.

Zhang, R., Zou, D., & Cheng, G. (2023a). Technology-enhanced language learning with null and negative results since 2000: A systematic review based on the activity

theory. *Education and Information Technologies*, 1–61. https://doi.org/10.1007/s10639-023-11993-1

Zhang, R., Zou, D., & Cheng, G. (2023b). Chatbot-based training on logical fallacy in English argumentative writing. *Innovation in Language Learning and Teaching*, 1–14. https://doi.org/10.1080/17501229.2023.2197417

Zhang, R., Zou, D., & Cheng, G. (2023c). A review of chatbot-based learning: Pedagogical approaches, implementations, factors leading to effectiveness, theories, and future directions. *Interactive Learning Environments*, 1–29. https://doi.org/10.1080/10494820.2023.2202704

Zhang, R., Zou, D., Cheng, G., & Xie, H. (2025a). Flow in ChatGPT-based logic learning and its influences on logic and self-efficacy in English argumentative writing. *Computers in Human Behavior, 162*, 108457. https://doi.org/10.1016/j.chb.2024.108457.

Zhang, R., Zou, D., & Cheng, G. (2025b). ChatGPT affordance for logic learning strategies and its usefulness for developing knowledge and quality of logic in English argumentative writing. *System, 128*, 103561. https://doi.org/10.1016/j.system.2024.103561.

Zhu, M. (2025). Leveraging ChatGPT to support self-regulated learning in online courses. *TechTrends, 1–11*. https://doi.org/10.1007/s11528-025-01075-z

Zimmerman, B. J. (2013). Theories of self-regulated learning and academic achievement: An overview and analysis. In *Self-regulated learning and academic achievement* (pp. 10–45). https://doi.org/10.4324/9781410601032-5

Ruofei Zhang is a lecturer at Centre for English and Additional Languages, Lingnan University. Her research interests include second language acquisition, technology-enhanced language learning, game-based learning, chatbot-based learning, and self-regulated learning. She has published more than 40 research papers in international journals, including *Computer Assisted Language Learning, Computers in Human Behavior*, and *System*.

Prof. Di Zou is an Associate Professor at Department of English and Communication, The Hong Kong Polytechnic University. Her research interests include second-language acquisition, technology-enhanced language learning, game-based language learning and AI in language education. She has published more than 150 research papers in international journals, including *Computers & Education, Computer Assisted Language Learning, Language Teaching Research*, and *British Journal of Educational Technology*. Her totally Google Scholar citation counts is 10,285, h-index is 47, and i-10 index is 99 at Google Scholar. She has been recognised as one of the World's Top 2 % Scientists by Stanford University for four consecutive years (2021–2024).

Prof. XIE Haoran is a Professor and the Person-in-Charge of Division of Artificial Intelligence, the Acting Associate Dean of School of Data Science, and the Director of LEO Dr. David P. Chan Institute of Data Science at Lingnan University, Hong Kong. His research interests include Artificial Intelligence, Big Data, Educational Technologies. He has totally published 414 research publications including 239 journal articles and 5 patents. His totally Google Scholar citation counts is 18,647, h-index is 55, and i-10 index is 180 at Google Scholar. He has been included in the list of The World's Top 2 % Scientists by Stanford University, is a senior member of ACM and IEEE, and has been honored with 24 research awards.

Prof. Philips Fu Lee Wang is a Professor and the Dean of Science and Technology at Hong Kong Metropolitan University. His research interests include Applied Algorithm, Educational Technology, Information Retrieval, and Machine Learning. He has served as the principal investigator/project leader for over 30 grants, including six grants awarded by the Research Grants Council of HKSAR. He has published over two hundred articles in refereed journals and conference proceedings, including *Decision Support Systems*, *IEEE Transactions on Cybernetics*, *IEEE Intelligent Systems*, *IEEE Multimedia*, *IEEE Transactions on Affective Computing*, *Information & Management*, *Information Processing & Management, Information Sciences*, *Neurocomputing*, and *Optics and Lasers in Engineering* among others. He was also the Chair of ACM Hong Kong Chapter and IEEE Hong Kong Section Computer Society.