

ChatGPT affordance for logic learning strategies and its usefulness for developing knowledge and quality of logic in English argumentative writing

Ruofei Zhang ^a, Di Zou ^{b,*}, Gary Cheng ^c

^a Centre for English and Additional Languages, Lingnan University, Hong Kong Special Administrative Region

^b Department of English and Communication, The Hong Kong Polytechnic University, Hong Kong Special Administrative Region

^c Department of Mathematics and Information Technology, The Education University of Hong Kong, Hong Kong Special Administrative Region

ARTICLE INFO

Keywords:

Chatbot
ChatGPT
English argumentative writing
Logical learning
Second language writing

ABSTRACT

Logic learning is essential for developing knowledge and quality of logic in English argumentative writing. Its main strategies include gathering logic instructions (Gathering), understanding logical concepts (Understanding), exercising knowledge of logic (Exercising), analysing logic in authentic writings (Analysing), and revising and creating logical links under guidance (Crafting) – which ChatGPT may afford. However, there has been limited exploration of ChatGPT-based logic learning. To address this gap, we developed a GPT-4-based logic learning bot and engaged 40 EFL university students, exploring ChatGPT's affordance for logic learning strategies and its usefulness for developing logical knowledge and quality of logic in English argumentative writing. We measured ChatGPT affordance by analysing learner-bot conversations and evaluated the usefulness of ChatGPT affordance through pre-post knowledge tests and essay writing tasks. Semi-structured interviews were also conducted. Our findings revealed that ChatGPT afforded 12 sub-strategies of logic learning, with Gathering and Exercising strategies being the most frequently used. ChatGPT-based logic learning strategies, especially Gathering and Exercising strategies, significantly developed knowledge and quality of logic in English argumentative writing. Through analysing the results, we identified ChatGPT features influencing the frequency and effectiveness of logical learning strategies and offered implications for implementing ChatGPT-based logic learning.

1. Introduction

Logic in English argumentative writing refers to the use of inference and reasoning to support theses and arguments, demonstrated through the accuracy, sufficiency, and relevance of evidence, clarity of arguments, and convincingness of idea links within the writings (Hale, 2002; Murray, 2012). English argumentative writings with higher logic quality tend to have stronger arguments, better coherence, and greater effectiveness in persuading learners (de Swart, 2018). Thus, logic quality has been widely accepted as a crucial criterion for evaluating English argumentative writing (de Swart, 2018). Closely associated with logic quality is knowledge of logic, which compasses the understanding of principles and mechanisms of diverse logical concepts and the skills of applying them in specific

* Corresponding author. Department of English and Communication, The Hong Kong Polytechnic University, Hong Kong Special Administrative Region.

E-mail addresses: ruofeizh@gmail.com (R. Zhang), dizoudaisy@gmail.com (D. Zou), chengks@edu.hk (G. Cheng).

contexts (Oaksford & Chater, 2020). Researchers argued that learners equipped with sufficient knowledge of logic could avoid and address most reasoning errors in English argumentative writing and thereby achieve a high quality of logic in the writing (Oaksford & Chater, 2020). However, in general, English-as-a-foreign-language (EFL) learners have insufficient knowledge of logic and face great challenge in creating high quality of logic in English argumentative writing (Murray, 2012; Saidi, 2020). For example, El Khoiri and Widiati (2017) required 40 EFL university students to compose English argumentative essays and found that most of their students displayed surprisingly low logic quality in argumentative writing. The researchers also found that most EFL learners lacked knowledge of logic in English argumentative writing and perceived the knowledge as abstract, complex, and incomprehensible (El Khoiri & Widiati, 2017).

Researchers and educators have proposed strategies for EFL learners to develop knowledge and quality of logic in English argumentative writing. First, learners should receive explicit instructions for logical concepts' forms, meanings, and structures so they can recognise the concepts in arguments (Murray, 2012). Second, learners should comprehend the underlying meanings and structures of logical concepts to develop critical thinking abilities and understand logical links (El Khoiri & Widiati, 2017). Third, learners should do exercises to reinforce knowledge of logic (Song & Sparks, 2019). Fourth, learners should analyse logical links in English writings so that they can understand the logic in different authentic contexts (Selpia & Setyarini, 2020). Finally, learners should be guided to modify and create logical links in argumentative writings so that they can solve real logical problems and enhance the logical quality of real argumentative writings (Guo & Lee, 2023).

As a large language model with excelling natural language understanding and generation capabilities, ChatGPT is a promising tool for logic education that can respond to learners' intents to apply various logical learning strategies (Zhang et al., 2023a). ChatGPT may respond to learners' requirements for logic instruction by creating and presenting personalised instructions (Kasneci et al., 2023). It may respond to learners' attempts to understand logical concepts by asking and answering exploratory questions (Herft, 2023) and providing associated examples (Kasneci et al., 2023). It may respond to learners' needs for logic exercising by providing varying exercises with immediate feedback (Baidoo-Anu & Ansah, 2023). It may support learners' actions of analysing authentic writing logic by breaking down the writings from the required perspectives and providing relevant comments (Kohnke et al., 2023). It may also respond to learners' efforts in crafting logical links by offering personalised comments and suggestions on learner output (Kasneci et al., 2023). This way, ChatGPT may afford various logic learning strategies and promote the development of knowledge and quality of logic in English argumentative writing.

Despite its great potential, ChatGPT-based logic learning has yet to be implemented or investigated. Given the considerable significance of logic learning for EFL learners (El Khoiri & Widiati, 2017), harnessing the potential of ChatGPT to enhance logic learning outcomes holds exceptional value. By investigating ChatGPT's possible affordances for logic learning strategies and its influence on logic learning outcomes, we may better understand the underlying mechanisms of ChatGPT-based learning and provide insights for logic and English argumentative writing pedagogies.

Moreover, research remains limited regarding learning strategies in ChatGPT-based settings (Kasneci et al., 2023), both within and beyond the scope of logic education, which is an investigation-worthy topic. The quantity of learning strategy applications is positively correlated to learning outcomes (Glogger et al., 2012). To properly leverage educational technology and maximise technology-enhanced learning outcomes, an in-depth understanding is essential to the frequency and impacts of learning strategies in specific learning settings (Mayes & de Freitas, 2013; Zhang & Zou, 2022). Hence, to apply ChatGPT efficiently for maximised logic learning outcomes, it is meaningful to explore the frequency and impacts of logic learning strategies in ChatGPT-based settings, the results of which may be applicable in other domains and disciplines.

To fill these gaps, we developed a GPT-4-based bot for logic learning, *LogicalHamster*, and conducted an experiment involving 40 Chinese university students. By analysing learner prompts to *LogicalHamster* and learner performances in knowledge tests and writing tasks, we aimed to understand ChatGPT affordances for various logic learning strategies and its impacts on developing knowledge and quality of logic in English argumentative writing. By undertaking this study, we may deepen the understanding of ChatGPT-based learning mechanism and provide implications for future design and implementation of this approach across disciplines. Two questions guided this study.

RQ1: To what extent does ChatGPT support EFL learners' learning strategies for developing knowledge and quality of logic in English argumentative writing? Why?

RQ2: Do ChatGPT-based logic learning strategies influence EFL learners' development of knowledge and quality of logic in English argumentative writing? How and why?

2. Literature review

2.1. Developing EFL students' knowledge and quality of logic in English argumentative writing

Logic is about valid reasoning and inferences, entailing a series of statements or premises in a coherent manner to justify a conclusion (DeLancey, 2017). In English argumentative writing, it reflects writers' ideas and thought processes and determines the extent to which the writing achieves its primary goal of persuading the reader of the merits of a certain viewpoint or action (de Swart, 2018). Despite its crucial importance, high logic quality is not easy to achieve, especially for EFL learners (El Khoiri & Widiati, 2017). Even with advanced English language skills and a firm grasp of logic in their native language writing, EFL learners often encounter significant challenges in creating sound logic in English argumentative writing (El Khoiri & Widiati, 2017), due to the distinctive features and expectations of English writing in reasoning and presentation from other languages (Murray, 2012).

To enhance the quality of logic in English argumentative writing, it is crucial to develop a specific understanding of logic within this context (Murray, 2012). One essential aspect of the knowledge of logic for improving English argumentative writing is about reasoning errors (Saidi, 2020). Researchers found that EFL learners generally produced many reasoning errors in English argumentative writing, most of which are relatively simple and straightforward (Oaksford & Chater, 2020). EFL learners can address most of these errors and largely enhance the quality of logic in their argumentative writing once they have acquired the knowledge about reasoning errors (Oaksford & Chater, 2020). Reasoning errors commonly-seen in English argumentative writings included: (a) Begging the Question, a fallacy where the conclusion is assumed in the premises; (b) Red Herrings, a fallacy involving the diversion of attention from the relevant issue by introducing an irrelevant topic; (c) Hasty Generalization, a fallacy that draws a conclusion from a small or unrepresentative sample; (d) Faulty Analogy, a fallacy that relies on misleading or superficial similarities between two things; (e) Post Hoc, a fallacy that assumes a causal relationship based solely on temporal succession; (f) False Alternatives, a fallacy that restricts considerations to only two options when more possibilities exist; (g) Slippery Slope, a fallacy that suggests a small initial step will lead to a chain of related events culminating in a significant outcome (El Khoiri & Widiati, 2017; Murray, 2012). By acquiring a comprehensive understanding of these logical concepts, EFL learners can identify and rectify most reasoning errors in their English argumentative writings and largely elevate the quality of their writing logic (Murray, 2012).

Researchers and educators have proposed and implemented many strategies and approaches to developing knowledge and quality of logic in argumentative writing. For example, Nejmaoui (2019) provided EFL learners with instruction on logical concepts in English argumentative writings. Student performance in pre-post essay writing tasks indicated the significant usefulness of the instruction for developing quality of logic in English argumentative writing. Song and Sparks (2019) required middle school students to do exercises on knowledge of logic in an educational digital game. The test and survey results indicated the usefulness of the exercises for enhancing skills in identifying reasoning errors. Selpia and Setyarini (2020) investigated high-school EFL students' knowledge of logic and identified their overall low proficiency levels. The researchers argued that teachers should guide students to analyse logic in authentic English writings while asking and answering students' questions about logic.

2.2. Learning strategies

Learning strategies are purposeful actions that directly influence learning outcomes by aiding in the completion of academic tasks and the acquisition and retention of information (Nisbet & Shucksmith, 2017). Bloom's revised taxonomy classifies learning strategies based on six cognitive levels: Remembering, Understanding, Applying, Analysing, Evaluating, and Creating (Bloom & Krathwohl, 2020). Another influential model is Kolb's Experiential Learning Theory, which organises strategies into four cognitive stages: Concrete Experience, Reflective Observation, Abstract Conceptualization, and Active Experimentation (Kolb et al., 2014). Merrill's Principles of Instruction categorise learning strategies based on five cognitive processes: Problem-Centred Learning, Activation, Demonstration, Application, and Integration (Merrill, 2002). While these models have received criticism, such as Bloom's taxonomy being criticised for rigidity in real-life situations (Tutkun et al., 2012), Kolb's Experiential Learning Theory for oversimplifying cognitive styles (Schenck & Cruickshank, 2015), and Merrill's Principles of Instruction for not being suitable for beginners (Learning Everest, 2024, March 1), they have nevertheless made significant contributions to the field and are widely used for analysing learning strategies across diverse settings and domains.

Researchers have proposed approaches to promote learning strategy application. One is providing guiding questions and suggestions to enhance learner awareness of learning strategy application and to remind students to employ strategies in relevant situations (Wong et al., 2019; Zhang & Zou, 2024). Another approach is fostering high learner autonomy, allowing learners to freely and flexibly apply strategies according to individual needs (Bai & Wang, 2023). Empirical studies support the usefulness of these approaches. For example, Chu et al. (2023) developed a digital vocabulary-learning game that encouraged strategy application by granting learners a high level of autonomy and offering strategy-related suggestions based on their learning progress. The results demonstrated the significant effectiveness of the game in facilitating learning strategy application.

In sum, the literature has provided different models to categorise learning strategies. Offering high learner autonomy and guiding questions and suggestions on strategies could promote learning strategy application and thus improve learning outcomes.

2.3. Chatbot-based learning

Chatbot-based learning is acquiring knowledge and skills through natural, synchronous conversations with conversational systems, i.e., chatbots like Claude and ChatGPT (Chang et al., 2023). In this context, students generate prompts, such as questions or requests, to interact with the chatbot and apply learning strategies (Adamopoulou & Moussiades, 2020).

Chatbot-based learning has been increasingly implemented in various domains and disciplines in recent years. Hwang and Chang (2021) reviewed 29 SSCI journal articles on chatbot-based learning, identifying rapid research growth in this field. However, they noted a need for more design for learning activities in this setting. Kuhail et al. (2023) systematically reviewed 36 studies and discovered that educational chatbots could recommend tutorials, foster academic discussions, answer questions, encourage emotional expression, and provide emotional support. However, most conventional chatbot-based learning programs were chatbot-driven and limited in prompt production flexibility and learner autonomy in applying learning strategies. Smutny and Schreiberova (2020) systematically examined 47 educational chatbots and identified four aspects to evaluate the quality of educational chatbots: (a) Teaching, concerning chatbots' capability to recommend learning materials, provide feedback, and monitor learning progress; (b) Humanity, concerning chatbots' capability to manage themed discussions and answer specific questions; (c) Affect, concerning chatbots' capability to provide greetings, show pleasant personality, and entertain learners; (d) Accessibility, concerning chatbots'

capability to detect learner intents and provide contextually-appropriate responses. They also described the educational chatbots that can analyse individual learning requests and offer individualised comments as the top level.

In sum, chatbots are helpful learning tools overall. However, conventional chatbot-based learning is mainly passive, rigid, English-centric, and limited in prompt production flexibility and learner autonomy in applying learning strategies. Research on learning strategies was also scarce in this field.

2.4. AI and generative AI

AI is a broad term encompassing various technologies that can perform tasks traditionally associated with human intelligence, such as reasoning, learning, and language understanding. As a subset of AI technologies, Generative AI specialises in generating fresh, meaningful content, such as text, images, and music. Its unique capability to personalise and adapt content to individual user needs makes it highly valuable across diverse fields, including education and evaluation (Feuerriegel et al., 2024).

In recent years, AI and Generative AI has been increasingly used in writing education and demonstrated great potential. For example, Nazari et al. (2021) explored the impacts of a 12-week AI-enhanced EFL academic writing course on adult learners. In the course, 60 students engaged in group discussions and completed weekly writing tasks with the aid of an AI-powered writing assistance tool, which generated scores, corrections, and suggestions; the other 60 students completed the same tasks without the aid of the tool. Analyses of the pre-post tests revealed the significant positive impacts of the AI-powered writing assistance tool on learners' behaviours, emotions, and cognition in the writing course. AI also improved students' self-efficacy and emotions in English writing. Hsiao and Chang (2023) conducted an 18-week online EFL course integrated with AI-powered tools for 43 secondary school students. In the course, AI scored students' English writings, provided suggestions on grammar and spelling, and presented information on the usage and frequencies of lexis in authentic writing. The results showed that AI-enhanced learning significantly enhanced academic performance, regardless of students' grade levels, genders, and school locations. Alharbi (2023, pp. 1–15) systematically reviewed 104 studies on AI-powered writing assistance tools, highlighting the growing application of these tools by students and their overall usefulness in writing education, particularly in the field of L2 education. Four main affordances of AI-powered writing assistance tools in writing education were revealed: (a) evaluating students' writing outcomes, (b) providing writing corrective feedback, (c) translating students' writing, (d) drafting texts for the students.

Despite their great potential in educational contexts, incorporating AI and Generative AI into writing classrooms remains controversial. Some educators were concerned that the application of AI might diminish learner engagement in learning activities and weaken their critical thinking and analysis abilities (Baidoo-Anu & Ansah, 2023). Others argued that AI would threaten academic integrity by offering unethical help to students (Alharbi, 2023).

2.5. ChatGPT-based learning

GPT, an exemplary example of generative AI, has demonstrated remarkable utility in chatbot-based learning (Baidoo-Anu & Ansah, 2023). Kohnke et al. (2023) conducted a synthesis study on ChatGPT-based language learning. They highlighted its usefulness in delivering learning materials of various genres, analysing writings from different perspectives, and providing personalised suggestions. However, research on the underlying mechanism and effectiveness of ChatGPT-based learning was still limited. Baidoo-Anu and Ansah (2023) conducted an exploratory synthesis study on ChatGPT-based learning. They identified its advantages over most conventional chatbot-based learning in tailoring learning materials to students' needs, promoting interaction, and providing formative assessments with ongoing feedback. However, the researchers expressed concerns about the effectiveness of ChatGPT in teaching specialised domain-specific knowledge or specific professional skills, owing to the unlabelled training data of ChatGPT. In another synthesis study, Kasneci et al. (2023) identified the usefulness of ChatGPT in understanding learner prompts, generating varying instructions and exercises, scoring academic performance, improving student engagement and interaction, and personalising learning experiences.

So far, the application of ChatGPT in writing education remains limited, though it has demonstrated great potential. AlAfnan et al. (2023) examined ChatGPT-based learning of communication and writing. They provided ChatGPT 30 questions that learners may ask and analysed ChatGPT's responses based on similarity and task accomplishment. The results showed that ChatGPT could provide accurate and reliable instructional contents, encourage students to explore theory-based questions, and support their application of the target knowledge in communication and writing.

Moreover, we found no literature on using ChatGPT for logic learning, though empirical studies suggested the great potential of this approach. For example, Guo and Lee (2023) asked university students to develop critical thinking skills by discussing and writing an essay with ChatGPT. The results revealed the significant effectiveness of ChatGPT-based learning for enhancing perceived critical thinking skills. Zhang et al. (2023a) employed a chatbot to deliver instructions, examples, and exercises on logic to Chinese EFL learners. The results of pre-post tests and writing tasks showed that the chatbot-based learning enhanced knowledge and quality of logic in English argumentative writing. However, the chatbot developed by Zhang et al. was built upon a conventional conversational system. Recognising that GPT surpasses most conventional conversational systems in natural language understanding and generation, they proposed the immense potential of employing ChatGPT in logic education.

In sum, the literature review revealed different logic learning strategies and measures of logic learning outcomes. ChatGPT shows promise in supporting writing education and logic education, but there has been no empirical research in this direction.

2.6. Conceptual framework

Through the literature review, we have developed a conceptual framework of ChatGPT-based logic learning. This framework integrates key insights from seminal studies in logic learning (e.g., Selpia & Setyarini, 2020; Song & Sparks, 2019), chatbot-based learning (e.g., Zhang et al., 2023a, Zhang et al., 2023b), and learning strategy theories (e.g., Bloom & Krathwohl, 2020; Merrill, 2002) (Fig. 1).

Our framework posits that ChatGPT facilitates the learning of logic in English argumentative writing through affording five main logic learning strategies: (a) Gathering, concerning acquiring information about logic (Murray, 2012) from ChatGPT (Kasneci et al., 2023); (b) Understanding, focusing on comprehending the meanings and structures of logical concepts (El Khoiri & Widiati, 2017) from interacting with ChatGPT (Herft, 2023); (c) Exercising, concerning asking ChatGPT for logic exercises and doing the exercises (Song & Sparks, 2019) with ChatGPT's feedback (Baidoo-Anu & Ansah, 2023); (d) Analysing, concerning examining the logic in authentic English arguments or argumentative writings (Selpia & Setyarini, 2020) under ChatGPT's guidance (Kohnke et al., 2023); and (e) Crafting, concerning forming and modifying logical links in English argumentative writings under ChatGPT's guidance (Guo & Lee, 2023). These strategies are not only integral to the learning process but also serve as indicators of the effectiveness of ChatGPT's affordances. Learner prompts can reflect learner application of different logic learning strategies, thus serving as the indicator of ChatGPT affordance for logic learning strategies. For instance, frequent learner prompts requesting logic exercises or explanations may signify learners' active application of Gathering or Understanding strategies and ChatGPT's strong affordances for these strategies.

The frequent application of these strategies is hypothesised to enhance ChatGPT-based logic learning outcomes, which falls into two categories. One is the development of knowledge of logic in English argumentative writing, measurable via immediate and delayed post-tests assessing the learners' understanding and retention of logical concepts (Selpia & Setyarini, 2020). The other is the development of quality of logic in English argumentative writing, evaluated through post-learning revisions of learners' argumentative texts from a logical perspective (Nejmaoui, 2019).

The outcomes of ChatGPT-based logic learning are also influenced by learner factors. One is learners' prior knowledge of logic in English argumentative writing. Assessed through pre-tests, this factor helps gauge the baseline understanding of each learner, which may affect their progression (Song & Sparks, 2019). The other factor is learners' prior quality of logic in English argumentative writing. Measured via argumentative writing tasks completed prior to the learning intervention, this indicates the learner's starting capability in applying logical reasoning in their argumentative writing (Nejmaoui, 2019).

EFL learners in general need development in knowledge and quality of logic in English argumentative writing. The literature review revealed the high value and significant potential of ChatGPT in supporting various logic learning strategies and thereby promoting

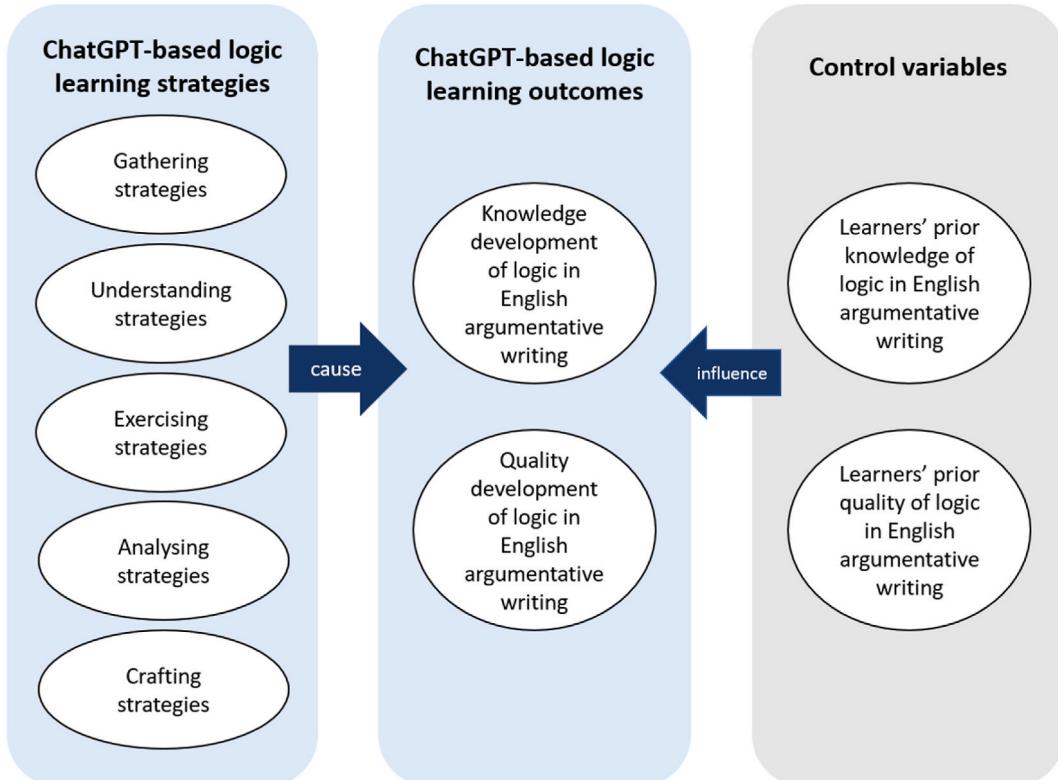


Fig. 1. Conceptual framework.

logic learning outcomes, while research on ChatGPT-based logic learning strategies remains limited. To bridge this gap, this study aims to explore ChatGPT-based logic learning based on our proposed framework, focusing on how ChatGPT may afford logic learning strategies and how its affordance may influence logic learning outcomes.

3. Method

In this study, we developed a logic learning chatbot based on GPT-4 and engaged 42 Chinese EFL university students, aiming to explore the affordances of ChatGPT for various logic learning strategies and the effectiveness of ChatGPT-based logic learning strategies in enhancing knowledge and quality of logic in English argumentative writing. To ensure a comprehensive understanding of the affordances and effectiveness, we employed a concurrent triangulation mixed-method approach (Riazi & Farsani, 2023), collecting and analysing quantitative data (test and writing scores) and qualitative data (interview transcripts) simultaneously. By doing so, we quantitatively measured the participants' changes in knowledge and quality of logic in English argumentative writing, while qualitatively exploring their perceptions and experiences with the ChatGPT-based logic learning. The concurrent analysis facilitated triangulation and complementarity, validating our findings and providing deeper insights (see Riazi & Farsani, 2023). Moreover, our study employed pre-post tests and writing tasks. This method established baselines for each participant's initial knowledge and quality of logic in English argumentative writing and directly measured their changes in knowledge and quality of logic attributable to the ChatGPT-based logic learning. Further details of our study methodology are provided below.

3.1. Participants

We recruited Chinese EFL university students from a university in Hong Kong using voluntary response sampling. We used voluntary response sampling for convenience and accessibility, so as to easily and quickly gather data from a broad audience with minimal logistical effort. To control voluntary response bias, we supplemented the sampling by incentives in the form of gift cards to encourage a diverse range of participants. Fifty-two students signed up for the participation in our study and completed a pre-treatment report via an online survey platform, Wenjuanxing.com. This report collected information on several criteria relevant to our study, concerning (a) biographical data, (b) needs for completing EFL argumentative writings, (c) technophobia, and (d) experiences with technology-enhanced learning, logic learning, ChatGPT-based learning, and prompt engineering training.

Based on the responses, we carefully selected our participants to ensure a sample representative of our study's requirements. We excluded nine students who did not frequently need to complete EFL argumentative writings and three students with prior experience in logic learning. This refined selection process resulted in a cohort of 40 students (32 females and eight males), aged 18 to 27 ($M = 22.85$, $SD = 3.05$). These participants were all enrolled in English-as-a-medium courses, frequently completed EFL argumentative writings, had some experience with technology-enhanced learning, no technophobia, and had no prior experience in logic learning, ChatGPT-based learning, or prompt engineering training. Appendix A presents the participant information.

Before the intervention, each participant signed a consent form. To address potential voluntary response bias and protect participants' rights, we ensured anonymity of data, clarified that their performance in the study would not have any consequences, and affirmed their right to withdraw from the study at any time.

3.2. Instruments

3.2.1. ChatGPT-4-based logic learning bot

Based on ChatGPT-4, we created a POE bot, *LogicalHamster* (<https://poe.com/LogicalHamster>). It focuses on teaching knowledge and skills about seven commonly seen reasoning errors, which are essential for improving quality of logic in English argumentative writing (Murray, 2012): Begging the Question, Red Herrings, Hasty Generalisation, Faulty Analogy, Post Hoc, False Alternatives, and Slippery Slope.

LogicalHamster can support all five categories of logic learning strategies (Appendix B). To support Gathering strategies, our bot could create and deliver varying instructions about the target logic concepts in whatever difficulty levels and presentation manners catering to learners' proficiency levels and individual needs. To facilitate Understanding strategies, *LogicalHamster* could ask and answer exploratory questions, comment on learners' paraphrasing of logical concepts, and provide examples of logical concepts with explanations. To support Exercising strategies, *LogicalHamster* could create and provide exercises about logic concepts, evaluate learner performance, and provide immediate feedback. These exercises cover diverse formats, such as multiple-choice questions, true/false questions, and error-spotting tasks. Additionally, learners could customise the exercise formats based on their individual requirements and preferences by clearly communicating their desired formats to *LogicalHamster* (e.g., "Please provide me with an argumentative paragraph that includes a reasoning error for me to identify"). For Analysing strategies, *LogicalHamster* would provide suggestions for analysing the quality of logic in argumentative writing and make examples of writing analysis. To aid in Crafting strategies, *LogicalHamster* could provide references in forming and modifying logical links, evaluate students' logical links, and offer personalised comments.

To encourage students' application of learning strategies, *LogicalHamster* provides guiding questions and suggestions about logic learning strategies based on learner progress, following Wong et al.'s (2019) recommendations. Students are free to generate prompts to *LogicalHamster* through typing, thus having high autonomy in applying learning strategies (Bai & Wang, 2023). Furthermore, *LogicalHamster* employs a humorous and supportive tone and consistently offers praise and encouragement for strategy application.

To evaluate the quality of *LogicalHamster*, we conducted a pilot study involving five students with similar backgrounds to our

intended participants. These students used *LogicalHamster* however they liked for over 45 min and subsequently participated in unstructured interviews lasting 15 min. During these interviews, they shared their perceptions of the logic learning tool and provided suggestions for further enhancements. The student feedback indicated that interacting with *LogicalHamster* was easy, supportive of applying learning strategies, and conducive to developing logical knowledge and quality of logic in English argumentative writing. Based on their experiences, the students recommended a duration of 45–75 min for ChatGPT-based logic learning using *LogicalHamster*. Additionally, two authors thoroughly analysed the interactions between *LogicalHamster* and the five students, identifying the consistently friendly, positive, and unbiased conversational styles exhibited by *LogicalHamster*. We also examined the logic instructional content in *LogicalHamster*'s responses, cross-referencing it with Bassham et al. (2010), El Khoiri and Widiati (2017), Murray (2012), and the website of the Department of Philosophy, Texas State University (<https://www.txstate.edu/philosophy/resources/fallacy-definitions/>). Our examination revealed no instances of bias or inaccuracy. These findings suggested that *LogicalHamster* was a valid and reliable tool for ChatGPT-based logic learning.

3.2.2. Pre-post-delayed tests on knowledge of logic

Following Song and Sparks (2019), we measured students' knowledge of logic by asking them to complete a 30-min paper-based knowledge test before the learning, immediately after the learning, and one week after the learning. The three tests were identical to ensure internal validity. The test, adapted from Zhang et al. (2023a), Zhang et al. (2023b) asked students to identify reasoning errors in English arguments, explain their meanings and structures, and match them with corresponding terms (Appendix C). Since the tests focuses on knowledge of logic in the context of English argumentative writing, the students were required to complete the tests using English.

The results were blind-scored by the research team using Zhang et al.'s (2023a) scoring system, with total scores ranging from 0 to 21 (Appendix D). The scoring showed high validity (Pearson's $r = 0.96$), with discrepancies resolved through discussion.

3.2.3. English essay writing and revision tasks

Following Nejmaoui (2019), we measured students' English argumentative logic via English essay writing and revision tasks. Before the learning, we asked the participants to write a 250-word English argumentative essay within 40 min about "Young people who commit crimes should be treated in the same way as adults. To what extent do you agree or disagree?" The requirements and topic were adopted from IELTS academic writing, following previous studies (Soodmand Afshar et al., 2017). After the learning, we asked the participants to revise the logic of their essays within 30 min. The students did the writing and revision tasks using MS Word without any reference.

We collected and analysed both the initial writing and the post-treatment revision by the participants. The quality of logic in the initial and the revised writings were separately blind-scored by two authors and a colleague with experience in teaching and assessing English essays. Using a 0–100 scale, the scoring criteria were adapted from Finken and Ennis's Illinois Critical Thinking Essay Scoring Rubric and IELTS's academic writing criteria (Appendix E). The scorings of the initial writing (Pearson's $r = 0.87$) and the revision (Pearson's $r = 0.85$) demonstrated acceptable validity, with discrepancies resolved through discussion.

3.2.4. Semi-structured interviews

To triangulate the quantitative results, we conducted semi-structured interviews with our participants immediately after the logic learning and the revision task. For each interviewee, we first explained the concepts of logic learning strategies for about 5 min and verified their understanding of the concepts using concept-checking questions before proceeding to the main interview. Based on their performance in the later interviews, all the interviewees had understood the concepts of the five logical learning strategies accurately and comprehensively.

The interviews were conducted in conversational Chinese, our participants' native language, to facilitate clear communication. During the 20-min main interviews, the participants were asked to respond to two questions while reviewing their chat histories with *LogicalHamster*. We developed and refined the interview questions based on expert recommendations to ensure the content validity. In a pilot study involving five students with backgrounds similar to our target participants, we administered these questions after the students had independently used *LogicalHamster* for over 45 min. Two authors then independently conducted thematic analyses of their responses. The comparative results of these analyses showed that the topics discussed by the students were consistent across the group (Cohen's $\kappa = 0.89$), and closely aligned with our research questions (Cohen's $\kappa = 0.93$). These the pilot study results highlighted the high inter-rater reliability. The questions are as follows:

Concerning RQ1.

- How frequently did you apply different logic learning strategies in the ChatGPT-based learning, using a scale from never to very frequently? Why?

Concerning RQ2.

- In the ChatGPT-based learning, do you perceive any positive/neutral/negative/mixed impact of logic learning strategies on your development of knowledge and quality of logic in English argumentative writing? Why?

3.3. Experimental procedures

The experiment lasted for three weeks (Fig. 2).

In Week One, we collected and analysed students' biographical reports and selected participants. For the selected participants, we collected their signed consent forms and provided a 20-min orientation on knowledge and quality of logic in English argumentative writing, ChatGPT, and ChatGPT-based logic learning. Then, we collected a digital copy of prior, recently completed English argumentative essays from each participant, which would be used in their ChatGPT-based logic learning in Week Two.

In Week Two, we asked the participants first to complete an English essay writing task and then complete the pre-test of knowledge of logic, aiming to control the influence of the knowledge test on learners' writing performance.

Subsequently, the participants conducted ChatGPT-based logic learning on a desktop computer using *LogicalHamster*. Among the five logic learning strategies, the Analysing and Crafting strategies must be applied to authentic argumentative writing, ideally using students' original essays (e.g., Guo & Lee, 2023). The other strategies do not require this. Therefore, we returned the digital copies of the original essays to the participants so they could use these essays to conduct the Analysing and Crafting strategies. The learning was self-paced, aligning with the inherent flexibility of ChatGPT-based education (Baidoo-Anu & Ansah, 2023). However, to control the possible impact of learning duration on learner behaviours and outcomes, we confined each learning session to a range of 45–75 min based on recommendations from students in the pilot study. During this session, the first author was monitoring the participants' learning to provide timely technical support when specifically requested. Other than that, the first author did not intervene in the learning and maintained a physical distance from the participants, aiming to minimise the potential impact of observer bias caused by researcher observation.

After the learning, we asked the participants first to revise the logic of the essays they composed just before the learning and then do an immediate post-test of knowledge of logic. We arranged the task sequence this way to control the influence of the knowledge test on learners' revising performance. Additionally, we used an online random picker (<https://minicool.com/random-picker/>) to randomly select 20 participants for semi-structured interviews (Participants 1, 3, 5, 7, 10, 11, 14, 16, 17, 19).

In Week Three, the participants completed a delayed post-test of knowledge of logic.

3.4. Data analysis

3.4.1. Prompt analysis

We retrieved the learner prompts to *LogicalHamster* from the POE system, read them repeatedly to familiarise ourselves with the content, and then coded learner prompts based on our proposed framework of ChatGPT-based logic learning strategies: (a) Gathering, (b) Understanding, (c) Exercising, (d) Analysing, and (e) Crafting. Table 1 presents our coding scheme.

During the coding, we inductively established 12 sub-strategies. Concerning Gathering, we found learners (a) asking *LogicalHamster* for instructions (e.g., "What is False Analogy?") and (b) asking for further elaboration of instructions (e.g., "Please further explain the

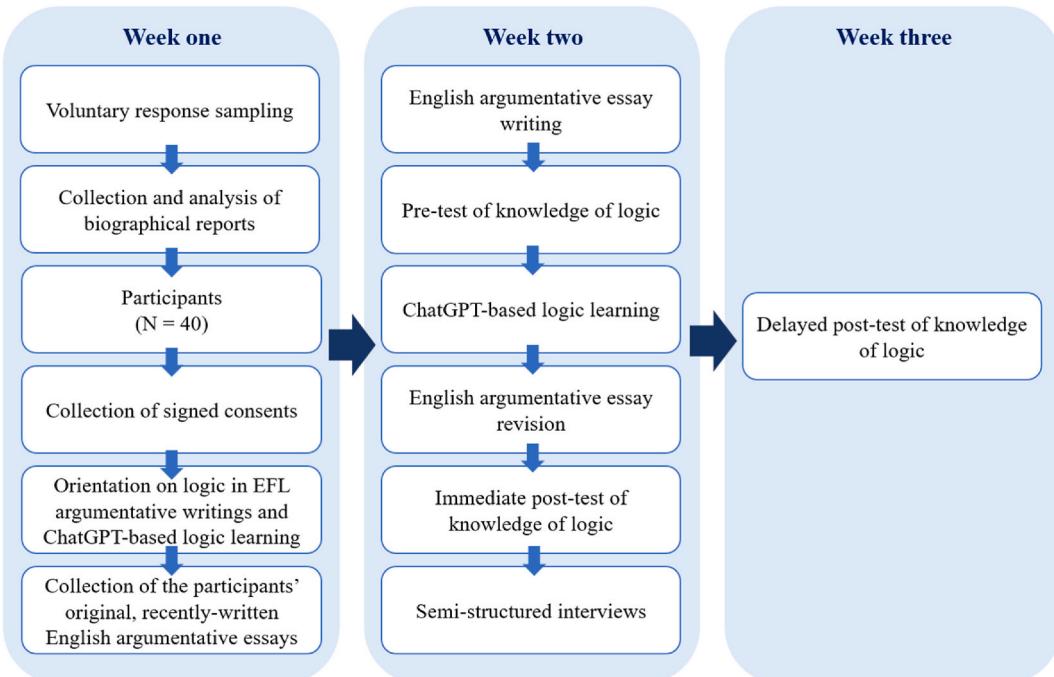


Fig. 2. Experimental procedures.

Table 1

Coding scheme of learner prompts for applying different logic learning strategies.

Logic learning strategies	Sub-codes	Example prompts
Gathering	Asking for instructions	“What is False Analogy?”
	Asking for further elaboration of instructions	“Please further explain the meaning of False Analogy.”
Understanding	Paraphrasing logical concepts	“So, False Analogy is about drawing comparisons between incomparable subjects and leading to an invalid conclusion. Am I right?”
	Asking exploratory questions Answering exploratory questions	“Why is False Analogy a fallacy, while True Analogy is not?” “False Analogy is a fallacy because it draws comparisons between incomparable subjects and leads to an invalid conclusion. However, True Analogy draws comparisons between comparable subjects, so it leads to a valid conclusion.”
Exercising	Asking for examples Asking for exercises	“Please give me some examples about False Analogy.” “Please give me some T/F questions about False Analogy.”
	Asking for adjustment of logic exercises in format and content Doing exercises	“Could you give me some more challenging exercises?” “True. The given sentence commits False Analogy.”
Analysing Crafting	Asking for advice for writing analysis Asking for reference to form and modify logical links	“From which perspectives can I analyse the following essay from the logical perspective?” “Please give me some references to modify this logical link.”
	Asking for comments on newly-crafted logical links	“I have modified/created the logical link as follows. What do you think about it?”

meaning of False Analogy”).

Regarding Understanding, we found learners (a) paraphrasing logical concepts for *LogicalHamster* to check their understanding (e.g., “So, False Analogy is about drawing comparisons between incomparable subjects and leading to an invalid conclusion. Am I right?”), (b) asking exploratory questions about logical concepts to *LogicalHamster* (e.g., “Why is False Analogy a fallacy, while True Analogy is not?”), (c) answering *LogicalHamster*’s exploratory questions about logical concepts (e.g., “False Analogy is a fallacy because it draws comparisons between incomparable subjects and leads to an invalid conclusion. However, True Analogy draws comparisons between comparable subjects, so it leads to a valid conclusion”), and (d) asking *LogicalHamster* for examples to understand logical concepts better (e.g., “Please give me some examples about False Analogy”).

Regarding Exercising, we found learners (a) asking *LogicalHamster* for logic exercises (e.g., “Please give me some T/F questions about False Analogy”), (b) asking *LogicalHamster* to adjust the format and content of logic exercises (e.g., “Could you give me some more challenging exercises?”), and (c) doing logic exercises (e.g., “True. The given sentence commits False Analogy”).

Regarding Analysing, we found learners submitting their original writings to *LogicalHamster* and asking for its advice for writing analysis (e.g., “From which perspectives can I analyse the following essay from the logical perspective?”).

Regarding Crafting, we found learners (a) asking *LogicalHamster* for reference to form and modify logical links in their original essays (e.g., “Please give me some references to modify this logical link”) and (b) asking *LogicalHamster* for comments on their newly-crafted logical links (e.g., “I have modified/created the logical link as follows. What do you think about it?”).

Two authors and a colleague individually coded the prompts and established sub-codes. The comparison showed that our coding outcomes had acceptable internal validity (Cohen’s $\kappa = 0.83$). The differences were solved via discussion.

3.4.2. Qualitative analyses

We audio-recorded, transcribed, and analysed the interviews following a meticulous process adapted from the standard approaches of qualitative content analysis outlined by Mayring (2004). Firstly, we went through the transcripts multiple times to familiarise ourselves with the content. Then, we created initial codes to categorise the interviewees’ main opinions according to the RQs: (a) The frequency at which learners applied different logic learning strategies (i.e., very frequently, somewhat frequently, sometimes, seldom, or never) and why (RQ1); (b) Their perceived impacts of different logic learning strategies on learning outcomes (i.e., positive, negative, neutral, or mixed) and why (RQ2). Appendix F presents the coding scheme with a pseudo example. Based on the scheme, two authors collectively analysed three transcripts and discuss on the coding method, ensuring consistency in the application of the coding scheme. After having agreed on a coding method, two authors independently analysed the remaining transcripts based on the established coding scheme. Throughout this process, frequencies were calculated.

We have included representative quotations from the transcripts to delve deeper into learner perceptions of various ChatGPT-based logic learning strategies. These quotes were crucial in illustrating the nuanced, unique experiences of individual learners. For example, Participant 19 noted, “*LogicalHamster* provided me unlimited, interesting examples in whatever contexts I required!” Such insight not only supports our findings concerning the positive influence of Understanding on the outcomes of ChatGPT-based logic learning but also enhances our understanding of how the effectiveness of the strategies may result from high-quality GPT-generated examples. In this way, we grounded our analysis of the application and impacts of ChatGPT-based logic learning strategies based on authentic participant voices.

To assess the internal reliability of this coding process, we compared the coding results and calculated Cohen’s κ coefficient. This resulting high κ value (0.84) indicates a strong agreement among the two coders. Finally, we collectively translated the analysis results and quotations into English. These steps were designed to ensure a comprehensive and rigorous analysis of the qualitative data,

providing both depth and breadth in our exploration of the interviews.

3.4.3. Quantitative analyses

To address RQ1, we conducted descriptive analyses of learners' test scores, writing scores, and numbers of prompts reflecting different logic learning strategies, focusing on the means and SDs.

To further investigate RQ1, we performed a one-way ANOVA using Welch's F in SPSS to examine potential variations in the number of related learner prompts across learning strategy types. Tukey's HSD test was used for post hoc pairwise comparisons due to equal sample sizes and similar group variances (Field, 2018).

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 ChatGPT-based logic learning strategies and academic outcomes, controlling for learner factors. The exogenous constructs were: (a) "Gathering strategies," with the frequencies of learner prompts for "Asking for instructions" and "Asking for further elaboration of instructions" as indicators; (b) "Understanding strategies," with the frequencies of learner prompts for "Asking exploratory questions," "Answering exploratory questions," "Paraphrasing the logical concepts" and "Asking for examples" as indicators; (c) "Exercising strategies," with the frequencies of learner prompts for "Asking for exercises," "Doing exercises," and "Adjusting the format and content of exercises" as indicators; (d) "Analysing strategies," with the frequency of learner prompts for "Asking for advice for analysing logic" as the indicator; (e) "Crafting strategies," with the frequencies of learner prompts for "Asking for references to create/modify logical links" and "Asking for comments on the created/modified logical links"; (g) "Learners' prior knowledge of logic in English argumentative writing," with learners' pre-test scores as the indicator; (h) "Learners' prior quality of logic in English argumentative writing," with learners' essay writing scores as the indicator. The endogenous constructs were: (a) "ChatGPT-based logic learning outcomes concerning knowledge of logic in English argumentative writing," with learners' immediate and delayed post-test scores as indicators; (b) "ChatGPT-based logic learning outcomes concerning quality of logic in English argumentative writing," with learners' essay revision scores as the indicator. The significance of the path coefficient was tested with a significance level of 5% (2-tailed) through the standard bootstrapping of 10,000 subsamples.

To further address RQ2, we employed multiple linear regressions using SPSS based on the PLS-SEM results, focusing on the predictive powers of learning strategies on learning outcomes.

3.4.4. Validity and reliability

Our data exhibited linearity and homoscedasticity for the random and even dispersion of points in scatterplots and the alignment of dots along the diagonal in P-P plots (Field, 2018). Additionally, our data showed normality, with skewness ranging between -2 and +2 and kurtosis between -7 and +7 (Byrne, 2010). Hence, our data were suitable for conducting one-way ANOVA and multiple linear regressions. Our data were also appropriate for PLS-SEM analysis which has no distribution requirement (Hair et al., 2021).

Multicollinearity is not a concern for multiple linear regressions as the Pearson correlation coefficient between each pair of indicators is below 0.90 (Field, 2018).

We evaluated the validity and reliability of the PLS-SEM measurement model following Hair et al. (2021). We found unacceptably low loadings of indicators ($< .40$) and Cronbach's α ($< .700$) for "Crafting strategies" in the prior model. Therefore, we removed this construct and re-conducted analyses using the PLS-SEM. The revised model (Appendix G) demonstrated acceptable indicator reliability (variable loadings $> .60$), high internal consistency reliability (Cronbach's $\alpha > .70$), high convergent validity (AVE $> .50$), and high discriminant validity (HTMT $< .85$) for each predictor-outcome pair. Collinearity was not an issue (VIF < 3). Thus, the PLS-SEM

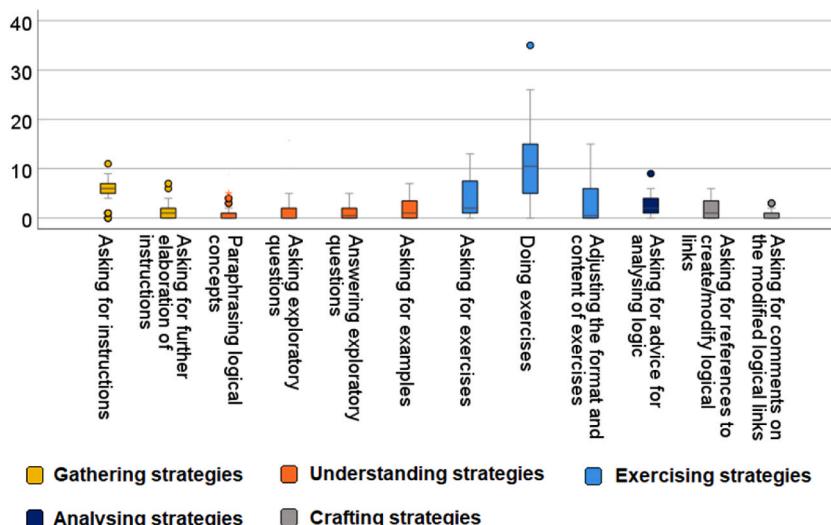


Fig. 3. Numbers of learner prompts for different learning strategies.

analyses of the revised model should be sufficiently reliable and valid.

Our sample size ($N = 40$) was large enough for one-way ANOVA (minimum of 30 observations), multiple linear regressions (minimum of 10 observations per predictor), and PLS-SEM (minimum of 10 times the largest number of indicators for any variable), according to the rules of thumb.

The data analyses were conducted individually by two authors and a colleague, and the results showed high internal validity (Cohen's $\kappa = 0.96$), with remaining differences resolved through discussion.

4. Results

4.1. RQ1. To what extent does ChatGPT support EFL learners' learning strategies for developing knowledge and quality of logic in English argumentative writing? Why?

4.1.1. Quantitative analysis results

On average, participants produced 35.95 learning-related prompts ($SD = 18.20$) in ChatGPT-based logic learning (Fig. 3). Concerning Gathering strategies, they produced 5.56 prompts ($SD = 2.56$) for "Asking for instructions" and 1.38 prompts ($SD = 1.66$) for "Asking for further elaboration of instructions." Regarding Understanding strategies, they produced 0.88 prompts ($SD = 1.75$) for "Paraphrasing the logical concepts," 1.65 prompts ($SD = 2.93$) for "Asking exploratory questions," 1.15 prompts ($SD = 1.48$) for "Answering exploratory questions," and 1.78 prompts ($SD = 1.98$) for "Asking for examples." Regarding Exercising strategies, learners produced 4.22 prompts ($SD = 3.99$) for "Asking for exercises," 10.90 prompts ($SD = 7.75$) for "Doing exercises," and 3.05 prompts ($SD = 4.11$) for "Adjusting the format and content of exercises." Concerning Analysing strategies, learners produced 2.68 prompts ($SD = 1.85$) for "Asking for advice for analysing logic." Regarding Crafting strategies, they produced 1.73 prompts ($SD = 1.88$) for "Asking for references to create/modify logical links" and 0.57 prompts ($SD = 0.93$) for "Asking for comments on created/modified logical links."

Welch's ANOVA revealed significant differences in the number of learner prompts reflecting different learning strategies, $F(11, 182.90) = 22.22, p < .001, \omega^2 = 0.41, \mu^2 = 0.41$. Post hoc Tukey's HSD test (Table 2) indicated that learners predominantly produced prompts for "Doing exercises," significantly more than other strategies, followed by "Asking for instructions" and "Asking for exercises." The least prompts were generated for "Asking for comments on the created/modified logical links."

In sum, we found overall frequent learner prompts reflecting different learning strategies, indicating ChatGPT affordance for various logic learning strategies. The number of learner prompts varied across learning strategies they related to, indicating that ChatGPT afforded different types of logic learning strategies to varying extents: Gathering strategies were supported most, while Crafting strategies were supported least.

4.1.2. Qualitative analysis results

Our interviews revealed students' different perceived frequency of applying different types of logic learning strategies in ChatGPT-based learning (Fig. 4). All interviewees reported frequent or very frequent use of Gathering strategies. They mentioned that when applying Gathering strategies, *LogicalHamster* tailored logic instruction to their individual preferences in format and style and elaborated on specific knowledge of logic points according to their needs. Seventeen interviewees (85%) indicated frequent or very frequent engagement with Understanding strategies, which involved posing exploratory, open-ended questions to *LogicalHamster*, paraphrasing the meanings and structures of logical concepts for verification, and requesting examples and explanations. Seventeen interviewees (85%) described frequent or very frequent use of Exercising strategies, detailing their process of requesting exercises, submitting responses, receiving feedback, and requesting additional exercises. Some students also requested adjustments from *LogicalHamster* in the content, difficulty levels, and format of the exercises to better meet their learning needs. Seven interviewees (35%) reported frequent or very frequent application of Analysing strategies. They submitted digital copies of their original essays to *LogicalHamster*, which then provided personalised suggestions for analysing the logic of their writings. Only six interviewees (30%) frequently used Crafting strategies. They improved the logic of their original writings based on suggestions from *LogicalHamster*. They

Table 2

Post hoc Tukey's HSD test results concerning different subsets of learner prompts for different learning strategies with significant differences in the frequency between each other ($\alpha < .05$).

	1	2	3	4
Asking for comments on the modified logical links	0.57			
Paraphrasing logical concepts	0.88	0.88		
Answering exploratory questions	1.15	1.15		
Asking for further elaboration of instructions	1.38	1.38		
Asking exploratory questions	1.65	1.65		
Asking for references to create/modify logical links	1.73	1.73		
Asking for examples	1.78	1.78		
Asking for advice for writing analyses	2.68	2.68	2.68	
Adjusting the format and content of exercises		3.05	3.05	
Asking for exercises			4.22	
Asking for instructions				5.65
Doing exercises				10.90
Sig.	0.152	0.117	0.605	0.724
				1.000

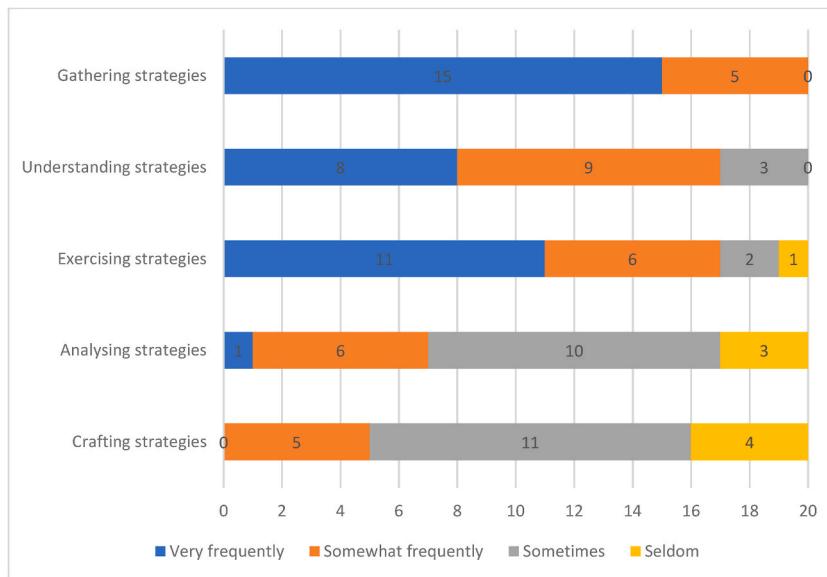


Fig. 4. Interviewees' perceived frequency of applying different learning strategies in ChatGPT-based logic learning.

then submitted their revised or newly created logical links back to *LogicalHamster* for evaluation and further suggestions. The interview findings triangulated the quantitative analysis results in which learners produced the most prompts for Gathering and Exercising and the fewest prompts for Analysing and Crafting, indicating that ChatGPT supported Gathering and Exercising most and Crafting and Analysing least.

In the interviews, we identified two main reasons for students' frequent strategy application in ChatGPT-based logic learning. One concerns *LogicalHamster*'s friend-like suggestions, guiding questions, praise and encouragement. Interviewees mentioned that *LogicalHamster* "gently encouraged" them to apply learning strategies using gentle reminders and kept praising them for strategy application "in a friendly, uncondescending way." For example, Participant 5 said: "*LogicalHamster* never pushed me to do anything. It just gave me friendly suggestions, which motivated me to follow its words and apply learning strategies." Additionally, according to the interviewees, such a communicational style enhanced learner awareness and motivation to apply learning strategies without making them stressful. Participant 11 mentioned:

In authentic learning, it's impossible to have teachers and peers constantly encouraging me and guiding me [to perform various logic learning strategies] by my side throughout the learning process. Even if it's possible, I would feel really stressful [to have human teachers and human peers constantly by my side] in the learning. However, *LogicalHamster* could keep encouraging me, praising me, and guiding me patiently and enthusiastically throughout the logic learning process, and it will never judge me. I won't feel stressful from *LogicalHamster*'s company.

The other reason for students' frequent strategy application in ChatGPT-based logic learning concerned the supportiveness of ChatGPT for active learning. Our students mentioned that *LogicalHamster* allowed them to participate in active learning and manage their learning tasks, which enhanced their awareness and confidence in applying learning strategies. Participant 3 stated, "*LogicalHamster* wouldn't give any more instructions or learning tasks unless I asked for them. So I felt really in control in the ChatGPT-based learning ... I had to manage my learning process by learning strategies." Moreover, *LogicalHamster* provided suggestions and recommendations about learning strategies based on students' learning histories, which scaffolded learners' active learning and encouraged their application of learning strategies.

However, not all students liked learning strategies in ChatGPT-based logic learning, especially Crafting and Analysing ones. Two reasons were mentioned in the interviews that demotivated students from applying learning strategies. One concerns the cognitive complexity of applying some logic learning strategies. Some interviewees admitted that Crafting and Analysing strategies were "difficult" and "exhausting," so they preferred to avoid them. They argued that "it was too much" to perform learning strategies associated with all five cognitive processes in a single-session learning.

Another factor mentioned in the interviews that may demotivate students from applying logic learning strategies is the quality of GPT responses. Some interviewees reported that *LogicalHamster* provided them with general, obscure responses when they asked exploratory, in-depth questions, which reduced their interest in applying Analysing and Crafting strategies. Participant 10 provided an example:

I submitted my essay to *LogicalHamster* and asked for some recommended perspectives for analysing and improving the logic. As response, it reminded me to provide more evidence for my arguments. Then, I asked for more detailed, specific suggestions,

while its responses remained overall shallow and general. So, I just gave up Analysing and Crafting ... Why bother thinking so hard and asking for so much? *LogicalHamster* could only feed me some superficial and general answers anyway.

4.2. RQ2. Do ChatGPT-based logic learning strategies influence EFL learners' development of knowledge and quality of logic in English argumentative writing? How and why?

4.2.1. Quantitative analysis results

Table 3 summarises the participants' scores in the knowledge of logic tests and the essay writing and revision tasks. In terms of their knowledge test, our participants, on average, obtained 6.32 out of 21 ($SD = 2.84$) in the pre-test, 16.48 out of 21 ($SD = 3.91$) in the immediate post-test, and 15.63 out of 21 ($SD = 3.78$) in the delayed post-test. In terms of their quality of logic in argumentative writing, their initial writing scored 71.38 out of 100 ($SD = 8.43$), and their revision scored 78.00 out of 100 ($SD = 8.20$).

Fig. 5 summarises the PLS-SEM results concerning the relationships between the frequency of strategy application and academic outcomes in ChatGPT-based logic learning. After controlling the learner factors, "Exercising strategies" demonstrated significant effects on "ChatGPT-based logic learning outcomes concerning quality of logic in English argumentative writing," $\beta = 0.31$, $|t| > 1.96$, 95% BCCI = (0.036, 0.623), $f^2 = 0.148$. "Gathering strategies" may have significant impacts on "ChatGPT-based logic learning outcomes concerning knowledge of logic," with coefficient close to the statistical watershed of significance, $\beta = 0.34$, $t = 1.73$, BCCI = (−0.266, 0.564), $f^2 = 0.187$.

The results of multiple regressions (**Table 4**) revealed the significant predictive power from "Gathering strategies" to "Immediate post-test scores" ($R^2_{change} = 0.17$, $F_{change}(2, 36) = 4.78$, $p < .05$), and "Delayed post-test scores" ($R^2_{change} = 0.16$, $F_{change}(1, 37) = 8.77$, $p < .05$). Calculated by G*Power, the estimated powers of the R^2 tests were solid (0.98). The independence among predictors was acceptable (Durbin-Watson value between 1 and 3).

The parameter estimates (**Table 5**) showed that the number of learner prompts for "Asking for further elaboration of instructions" had a significantly positive impact on the scores of the immediate post-test of knowledge of logic. For learners scoring 0 in the pre-test, one more prompt for "Asking for further elaboration of instructions" would lead to 0.71 points higher in the immediate post-test.

In sum, our statistical analysis results showed that in ChatGPT-based logic learning, Gathering strategies, especially "Asking for further elaboration of instructions," facilitated the development of knowledge of logic. Exercising strategies facilitated the development of quality of logic in English argumentative writing.

4.2.2. Qualitative analysis results

Our interviewees perceived the overall positive effects of logic learning strategies in ChatGPT-based settings for developing knowledge and quality of logic in English argumentative writing (**Fig. 6**). Interview feedback indicated that a significant majority of interviewees (85%) found the Gathering strategies—through which they received and processed knowledge of logic—to be highly effective in ChatGPT-based settings. Exercising strategies, which revealed students' knowledge gaps, facilitated the reinforcement of knowledge of logic, and addressed misconceptions, were considered very effective by 12 interviewees (60%). Eleven interviewees (55%) highlighted the effectiveness of Understanding strategies, noting their usefulness in helping them grasp the underlying structures of various logical concepts and aiding in their development of cognitive models about quality of logic in argumentative writing. Analysing strategies were seen as valuable by seven interviewees (35%), who emphasised their role in deepening understanding of logical concepts within authentic argumentative writings. Crafting strategies, which enables the immediate application of newly learned knowledge of logic to enhance argumentative writing, were considered effective by five interviewees (25%) in ChatGPT-based settings. The interview results triangulated the PLS-SEM analysis results that ChatGPT-based logic learning strategies had overall positive effects on learners' development, especially Gathering strategies and Exercising.

Our interviewees attributed the effectiveness of learning strategies to two main factors. One is the high-quality instructional content created by *LogicalHamster*. Many interviewees mentioned their appreciation of the "conciseness," "systematicity" and "explicitness" of the GPT-generated logic instruction. By Gathering these high-quality learning materials, learners could memorise and understand logical concepts efficiently. Participant 5 commented:

I initially assumed that ChatGPT would give me a lot of general, obscure, hard-to-understand information ... However, it turned out that the instructions created by *LogicalHamster* were very comprehensible, straightforward, and concise. It even listed the key knowledge points using bullet terms! I could grasp the key points and understand the logical concepts very quickly!

Moreover, our interviewees mentioned that *LogicalHamster* provided "detailed" feedback, which helped them identify knowledge gaps and develop new knowledge when Exercising. Interviewees also argued that the "funny" and "vivid" examples that *LogicalHamster*

Table 3

Summary of the participants' scores of the knowledge tests and the essay writing and revision tasks.

	Min	Max	M	SD
Pre-test of knowledge of logic (0–21)	0.0	12.0	6.317	2.8389
Immediate post-test of knowledge of logic (0–21)	4.0	21.0	16.476	3.9146
Delayed post-test of knowledge of logic (0–21)	5.0	21.0	15.634	3.7799
Quality of logic in the initial writing (0–100)	55	85	71.38	8.430
Quality of logic in the post-learning revision (0–100)	63	93	78.00	8.203

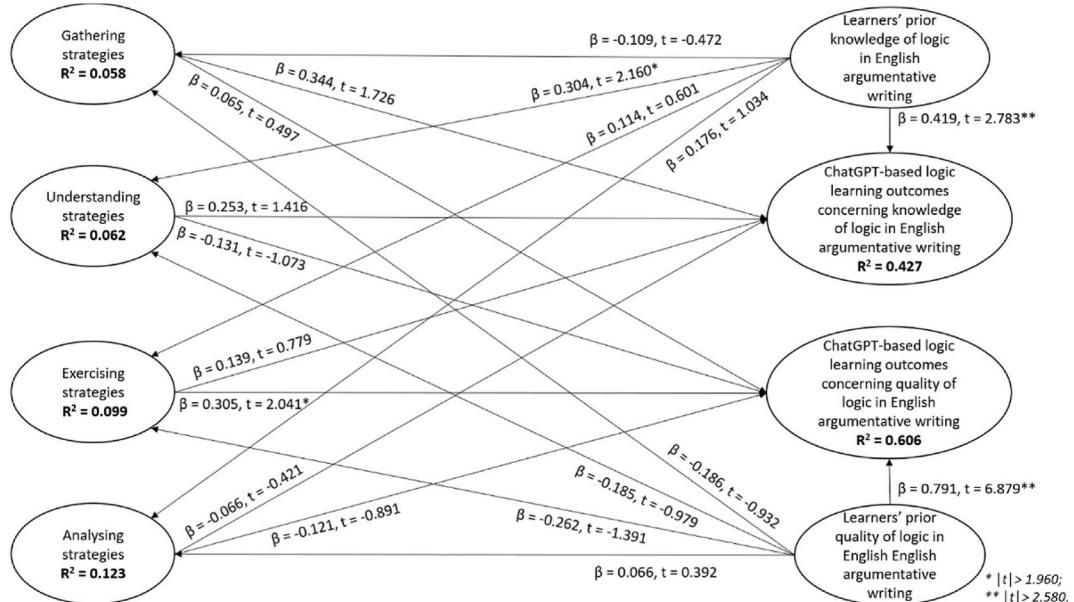


Fig. 5. Summary of the working model.

Table 4
Model estimates.

Models		R^2	R^2_{change}	F_{change}	Sig. F_{change}	Durbin-Watson
Predictors	Outcomes					
Learners' prior quality of logic in English argumentative writing	Writing scores after the learning	0.154	0.154	6.928	0.012*	n/a
Learners' prior quality of logic in English argumentative writing, Doing exercises,		0.267	0.113	1.793	0.167	2.281
Asking for exercises, Adjusting the format and content of exercises						
Learners' prior knowledge of logic	Immediate post-test scores	0.205	0.205	9.794	0.003*	n/a
Learners' prior knowledge of logic, Asking for instructions,		0.372	0.167	4.776	0.014*	1.753
Asking for further elaboration of instructions						
Learners' prior knowledge of logic	Delayed post-test scores	0.146	0.146	6.477	0.015*	n/a
Learners' prior knowledge of logic, Asking for instructions, Asking for further elaboration of instructions		0.230	0.084	1.966	0.155	1.934

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

Table 5
Parameter estimates.

Outcomes	Predictors	B	SE (B)	β	t	Sig.
Immediate post-test scores	(Constant)	9.346	1.727		5.411	<0.001 ^b
	Pre-test scores	0.690	0.185	0.501	3.728	<0.001 ^b
	Asking for instructions	0.322	0.214	0.208	1.502	0.142
	Asking for further elaboration of instructions	0.714	0.335	0.299	2.129	0.040 ^a

^a $p < .05$.

^b $p < .001$.

created helped them comprehend logical concepts, increasing the effectiveness of Understanding strategies. Participant 19 commented: "The examples presented in traditional texts are always limited and boring. However, *LogicalHamster* provided me unlimited, interesting examples in whatever contexts I required!" Some interviewees argued that the Analysing and Crafting strategies were effective due to the "constructive," "inspiring" suggestions and comments generated by *LogicalHamster*, which helped them understand and use logical concepts flexibly in authentic argumentative writing settings.

Another factor positively influencing the effectiveness of logic learning strategies was *LogicalHamster*'s supportiveness for

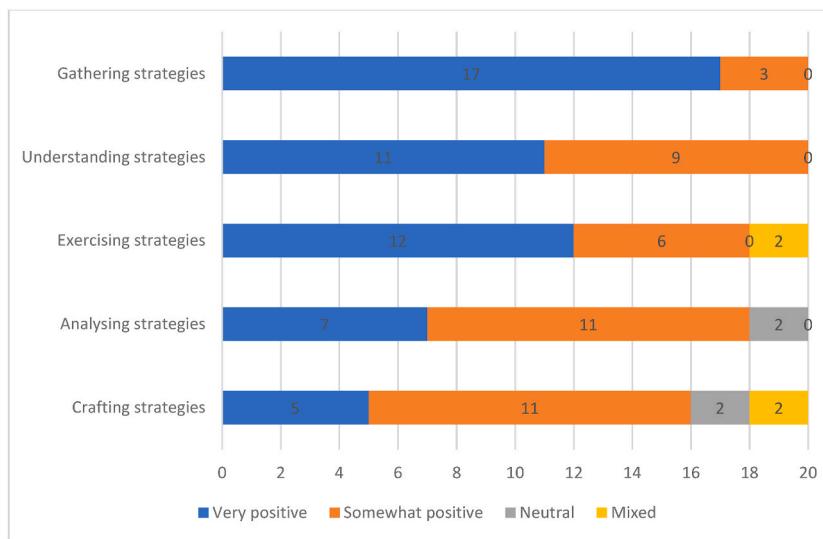


Fig. 6. Interviewees' perceived effects of different learning strategies on ChatGPT-based logic learning outcomes.

personalisation. According to the interviewees, students had various preferences when performing learning strategies: Participant 11 mentioned her preference for “bolded keywords” in instructions when Gathering; Participant 14 preferred customised exercise formats when Exercising. The interviews showed that *LogicalHamster* could personalise learning materials, presentation methods, and paces based on learner prompts and learning histories, which improved the effectiveness of learning strategies.

Additionally, our interviewees mentioned two factors that might reduce the effectiveness of learning strategies. One concerns the high difficulty in applying and generalising the knowledge learned from the single-session learning. Participant 10 stated, “I did learn a lot from this learning session, but I don’t think I could use the knowledge well in authentic setting.” According to the interviewees, it required two or more sessions of logic learning and explicit settings to fully apply Analysing and Crafting strategies to improve the quality of logic in argumentative writing significantly.

Another factor negatively influencing the effectiveness of logic learning strategies was related to the depth and complexity of ChatGPT-created learning materials. According to the interviewees, in-depth understanding and real-world application of knowledge of logic required Exercising based on “complex” and “difficult” questions (Participant 17), Understanding based on “subtle” and “close-to-real-life” examples (Participant 14), and “in-depth” explanations (Participant 17) of logical links in Crafting. However, *LogicalHamster* provided only “the over-easy exercises” (Participant 17), “superficial examples” (Participant 14), and “general advice” (Participant 5), which limited the effectiveness of learning strategies. Participant 14 mentioned, “I asked *LogicalHamster* to provide me some more challenging exercises. However, what it provided were still rather simple and superficial.”

5. Discussion

5.1. RQ1. To what extent does ChatGPT support EFL learners' learning strategies for developing knowledge and quality of logic in English argumentative writing? Why?

Our study revealed learners' overall active application of logic learning strategies of diverse types in the ChatGPT-based environment, evidencing ChatGPT's great affordance for various logic learning strategies. These findings shed light on ChatGPT as an all-around logical educational tool that can provide knowledge, feedback and scaffoldings as a teaching assistant, offer quizzes, questions and answers as a personal tutor, and interact with students as a learning partner (Deng & Yu, 2023). In particular, our participants used Gathering and Exercising strategies most frequently. This finding evidenced the expectations of previous synthesis studies on ChatGPT (e.g., Baidoo-Anu & Ansah, 2023) about its strong affordance of creating and presenting varying instructions, exercises, and feedback based on learners' individual needs, which is hardly affordable by human teachers who always have limited time and stamina (Pérez-Soler et al., 2020). By this logic, ChatGPT can be advantageous over human teachers in supporting learners' Gathering and Exercising strategies.

Additionally, we found ChatGPT afforded many strategies which were barely mentioned/affordable in previous technology-enhanced logic learning programs (e.g., Song & Sparks, 2019; Zhang et al., 2023a), such as “Paraphrasing logical concepts,” “Asking exploratory questions,” “Answer exploratory questions,” “Asking for advice for analysing logic,” and “Asking for comments on modified/created logical links.” This is because supporting these strategies requires analysis and evaluation capabilities, while most previous educational tools do not possess such capabilities and teach knowledge mainly by following pre-programmed patterns. ChatGPT, however, exhibited some analysis and evaluation capabilities in our study. As showed in the interview findings, ChatGPT could analyse and evaluate learner output and provide personalised responses, thus supporting diverse Understanding, Analysing, and

Crafting strategies. Hence, ChatGPT appeared advantageous over most conventional educational tools in logic learning for affording more learning strategies.

By analysing the results, we also identified factors influencing learners' frequency of applying logic learning strategies in a ChatGPT-based environment. Firstly, ChatGPT's emotional support may encourage learning strategy application. Our interviews showed that ChatGPT's friendly conversational styles, praise, and encouragement increased learner motivation in applying learning strategies. This phenomenon aligns with the Self-Determination Theory (Deci & Ryan, 2012), which emphasises the positive effects of the feeling of being accompanied on learner motivation. By this logic, students can feel accompanied due to ChatGPT's emotional support, thereby applying learning strategies frequently. This finding aligned with Pérez-Soler et al.'s (2020) argument for the importance of chatbots' capability of providing emotional support. It also indicated ChatGPT's high quality in detecting learner emotions and providing relevant emotional support, consistent with previous observations (Kuhail et al., 2023). Based on our findings, practitioners may increase learners' frequency of using learning strategies by utilising ChatGPT's capability of providing emotional support.

Moreover, ChatGPT's supportiveness for active learning may encourage learning strategy application. The interviews showed learners' overall high autonomy in ChatGPT-based learning, echoing Zhang et al.'s (2023c) observations of high-quality educational chatbots "facilitating active and constructive engagement in learning" (p. 15). Compared to conventional educational chatbots which are mainly based on pre-programmed patterns (Smutny & Schreiberova, 2020), ChatGPT allows even higher learner autonomy in applying learning strategies for its abilities of handling various improvised learner prompts. Our interviews showed that, in ChatGPT-based learning, students with higher autonomy felt stronger motivation and confidence in controlling their learning via learning strategies. This phenomenon aligns with the Self-Determination Theory (Deci & Ryan, 2012), which emphasises the importance of learner autonomy in learner motivation. Despite the high autonomy, learners also received suggestions and guiding questions in ChatGPT-based learning, which guided them to apply learning strategies appropriately and frequently, as mentioned in the interviews. This finding indicated the usefulness of suggestions and guiding questions in improving learner application of learning strategies, generalising Wong et al.'s (2019) findings into the ChatGPT-based learning settings. The finding also unveiled ChatGPT's robust support for active learning, as it not only enabled high autonomy in applying learning strategies but also effectively guided students in utilising their autonomy effectively. Hence, researchers and practitioners may use ChatGPT to support active learning.

However, ChatGPT may demotivate learning strategy application for its unsatisfying discussion and exploration abilities. According to our interviewees, ChatGPT sometimes provided general, superficial responses in discussion and exploration, which demotivated learners from applying strategies, especially Understanding, Analysing and Crafting. Our statistics triangulated the interview data, showing students' relatively less application of these three types of learning strategies. This finding suggests ChatGPT's current limitations in discussion and exploratory abilities. Considering the importance of in-depth analyses and application of knowledge of logic in authentic argumentative writings (Hundleby, 2010), these limitations of ChatGPT deserved particular attention of instructional designers. Practitioners should also be mindful of these limitations when implementing ChatGPT-based learning and consider strategies to address them.

Furthermore, our study revealed factors other than ChatGPT features that can influence learner application of logic learning strategies in the ChatGPT-based setting, particularly in relation to the cognitive complexity of these strategies. Most of our interviewees perceived Crafting and Analysing as cognitively demanding, which resulted in their lack of confidence and reluctance in applying these strategies. According to our interview results, a single-session logic learning might be insufficient for beginning learners to apply Crafting and Analysing. Consequently, these strategies were identified as the least utilised by our interviewees. This qualitative data was further supported by quantitative analysis, as participants generated the fewest prompts for Crafting and Analysing. The self-efficacy theory (Schunk, 1991) could explain the phenomenon, which contends that learners evaluate the difficulty of learning tasks, and their engagement in the tasks were negatively correlated with their perceived difficulty. Following the theory, learners were less motivated to use learning strategies in ChatGPT-based learning which they perceived as more challenging, i.e., Crafting and Analysing. However, considering the importance of these strategies in enhancing learners' logic in authentic English argumentative writing (Guo & Lee, 2023; Selpia & Setyarini, 2020), it is crucial to support students in frequently applying these strategies during logic learning. To address this issue, practitioners can consider conducting multiple sessions of ChatGPT-based logic learning, initially focusing on Gathering, Understanding, and Exercising, and then moving to Crafting and Analysing. This way, learners may feel more cognitively and emotionally prepared to effectively employ the more challenging strategies, i.e., Crafting and Analysing, and gradually increase their utilisation of all the five logic learning strategies.

5.2. RQ2. Do ChatGPT-based logic learning strategies influence EFL learners' development of knowledge and quality of logic in English argumentative writing? How and why?

Our study showed the overall positive effects of various ChatGPT-based logic learning strategies on developing knowledge and quality of logic in English argumentative writing. Our interview results showed that, in ChatGPT-based logic learning, learners received and memorised knowledge of logic input from Gathering, explored English argumentative logic and constructed cognitive models of it from Understanding, addressed their misunderstanding and reinforced their memories of knowledge of logic from Exercising, understood logical concepts in authentic argumentative writing settings from Analysing, and developed skills of using knowledge of logic to improve authentic argumentative writings from Crafting. Our findings generalised previous arguments for the usefulness of logical learning strategies (e.g., El Khoiri & Widiaty, 2017) into the ChatGPT-based learning settings. These findings also showed the effectiveness of ChatGPT for developing critical thinking and explicit reasoning skills in English argumentative writings, echoing Deng and Yu's (2023) observations of chatbot-based learning. Based on our findings, future practitioners may enhance

outcomes of ChatGPT-based logic learning by encouraging learner application of learning strategies.

Specifically, we identified statistically significant usefulness of applying Gathering strategies on ChatGPT-based learning outcomes, which ChatGPT's affordance of personalisation may explain. According to the interviews, ChatGPT could provide learning materials and tasks adapted to learners' needs and preferences, which enhanced the effectiveness of learning strategies. The quantitative analysis results about "Asking for further elaboration of instructions" supported the qualitative results, showing that personalised specifications of knowledge points from learners' required perspectives could significantly facilitate knowledge of logic development. This finding echoed previous arguments for the benefits of personalised learning on learning outcomes (Li & Wong, 2021). It also revealed ChatGPT's capability of detecting learner intents and needs based on prompts and contextual cues, consistent with previous arguments about ChatGPT's advantages over most educational chatbots in personalisation (Vemprala et al., 2024). Our study evidenced the usefulness of ChatGPT in "offer[ing] individualised comments and endorsements, analy[sing] individual learning requests, and provid [ing] tutoring instructions" (p. 9), suggesting it is a top-level artificial intelligence teaching assistant in Smutny and Schreiberova's (2020) model. Based on the findings, we recommend further application of ChatGPT in supporting personalised learning.

The usefulness of Gathering strategies may also result from ChatGPT's capability to create high-quality learning materials, which facilitated students' comprehension and memorisation of knowledge of logic with sufficient accuracy. This phenomenon aligns with the Input Hypothesis, emphasising the importance of high-quality input for developing learner proficiency. This finding provided empirical evidence for Kohnke et al.'s (2023) arguments for the high quality of ChatGPT-created instructions and explanations, indicating the great promise of developing and delivering learning materials via ChatGPT in future pedagogy. However, ChatGPT-created learning materials were far from perfect. Some of our participants perceived its lack of depth and complexity, which might have restricted the effectiveness of Understanding, Analysing, and Crafting strategies. Since these strategies were essential for developing knowledge and quality of logic in English argumentative writing (El Khoiri & Widiati, 2017), the insufficient depth and complexity of ChatGPT-created learning materials might have limited logic learning outcomes. Our PLS-SEM results supported this point, showing the insignificant effects of Understanding, Analysing, and Crafting strategies on ChatGPT-based logic learning outcomes. This finding revealed ChatGPT's limitation in adjusting the depth and complexity of learning materials. It also agreed with researchers' concern about ChatGPT's capability of teaching nuanced, intricate professional knowledge and skills (e.g., Baidoo-Anu & Ansah, 2023). Based on the results, we recommend cautious consideration of both the benefits and drawbacks of ChatGPT-created learning materials in implementing ChatGPT-based learning.

Our PLS-SEM analysis showed the significant effectiveness of Exercising strategies in developing the quality of logic in argumentative writing. According to the interviews, it is because Exercising enabled students to identify knowledge gaps, cover the gaps, and reinforce their memories of knowledge of logic. This finding confirmed previous arguments for the usefulness of educational chatbots in facilitating exercising (Deng & Yu, 2023). It also provided evidence for the importance of exercising in enhancing learners' logic (Song & Sparks, 2019), especially logic quality in English argumentative writing. However, our multiple regression analysis results revealed statistically insignificant effects of Exercising strategies, conflicting with the PLS-SEM analysis results. To address the conflict, future research may be conducted on the exact impact of Exercising.

This study revealed statistically insignificant effectiveness of Understanding, Analysing, and Crafting for enhancing ChatGPT-based logic learning outcomes, although our interviewees perceived an overall effectiveness of these strategies. This finding appears to contradict previous arguments for the usefulness of Understanding, Analysing, and Crafting as logic learning strategies (El Khoiri & Widiati, 2017; Guo & Lee, 2023). One possible explanation for this discrepancy could be the frequency of strategy application. As shown in the interview results and the descriptive analysis results, students applied Understanding, Analysing, and Crafting less frequently than Gathering and Exercising, especially Analysing and Crafting. The limited frequency of applying these strategies may have significantly restricted their impacts, resulting in statistically insignificant effects on learner development. Future researchers may delve into the exact effectiveness of Understanding, Analysing, and Crafting in enhancing ChatGPT-based logic learning outcomes by encouraging learners to apply these strategies with higher frequency. This way, researchers can gain a more comprehensive understanding of the effectiveness of all the five logic learning strategies in the ChatGPT-based setting.

6. Conclusion

This study has extensively explored the capabilities and advantages of ChatGPT in supporting diverse logic learning strategies, addressing two research questions that shed light on its effectiveness in logic education and English argumentative writing education. Our findings confirm ChatGPT's significant affordance for 12 specific learning sub-strategies. ChatGPT may excel human instructors in supporting Gathering and Exercising strategies, and outperforms other educational technologies in supporting Understanding, Analysing, and Crafting strategies. This usefulness is largely attributed to ChatGPT's ability to offer emotional support, enhance active learning, and facilitate robust discussion and exploration, which are crucial for effective learning in logic.

Moreover, our findings generalised the usefulness of logic learning strategies for developing logic knowledge and quality of logic in English argumentative writing into ChatGPT-based settings. The quality and depth of ChatGPT-created learning materials and ChatGPT's affordance of personalised learning demonstrated influences on the effectiveness of logic learning strategies – these attributes significantly contribute to the effectiveness of Gathering and Exercising strategies, enhancing academic outcomes of ChatGPT-based logic learning.

In conclusion, ChatGPT not only serves as a powerful educational tool that bridges the gap between theoretical knowledge and practical application but also fosters a more personalised and engaging learning environment. This study underscores the potential of AI-driven platforms like ChatGPT to facilitate educational practices, particularly in the domain of logic education and English argumentative writing education, by providing dynamic, supportive, and tailored learning experiences that are difficult to achieve through

traditional methods.

Implications and future studies

Our study holds significant theoretical implications. Firstly, most previous studies in this field focused on only one or two logic learning strategies (e.g., [Guo & Lee, 2023](#); [Song & Sparks, 2019](#)), while our study identified a comprehensive set of potent strategies through literature review, put them all together in a systematic framework, and examined their usage and effectiveness holistically in an authentic ChatGPT-based learning environment. This way, we established a systematic framework of learning strategies for logic in English argumentative writing via the combination of top-down and bottom-up approaches, which has received limited attention in previous research. According to our framework, ChatGPT can afford learner application of Gathering, Understanding, Exercising, Analysing, and Crafting, comprising a total of 12 sub-strategies. These strategies, in turn, can enhance learners' logic learning outcomes: Gathering supports knowledge of logic input; Understanding facilitates the development of cognitive models; Exercising reinforces knowledge of logic and helps correct misconceptions; Analysing deepens the understanding of logical concepts in authentic settings; Crafting enables learner use of the knowledge of logic in enhancing the quality of logic in English argumentative writing. Our framework aligns with Bloom's revised taxonomy that classifies learning strategies into Remembering, Understanding, Applying, Analysing, Evaluating, and Creating ([Bloom & Krathwohl, 2020](#)). However, unlike Bloom's taxonomy, which is criticised for its rigidity in real-life situations ([Tutkun et al., 2012](#)), our framework is rooted in the authentic pedagogy of logic in English argumentative writing and exhibits flexibility and adaptability in real-life implementation. Moreover, our framework also echoes Kolb's Experiential Learning Theory, which organises strategies into four cognitive stages: Concrete Experience, Reflective Observation, Abstract Conceptualization, and Active Experimentation ([Kolb et al., 2014](#)). Nonetheless, unlike Experiential Learning Theory, which is criticised for oversimplifying learning processes ([Schenck & Cruickshank, 2015](#)), our framework is systematic and detailed and caters to learners' spontaneous learning behaviours and natural cognitive processes in the ChatGPT-based learning setting. Given its flexible, adaptable, and meticulous nature, our framework for learning logic in English argumentative writing holds potential beyond the ChatGPT-based settings. Future researchers can refer to our framework for further exploration of this topic in various contexts.

Secondly, our study shed light on ChatGPT-based learning of logic in English argumentative writing, which may be generalised in other domains and disciplines. On the one hand, we identified ChatGPT's strong support for various logic learning strategies, including those that were unaffordable by most other educational technologies, such as Analysing ([Smutny & Schreiberova, 2020](#)). The findings provide empirical evidence for previous synthesis study results concerning the usefulness of ChatGPT in presenting personalised learning materials and comments ([Kasneci et al., 2023](#)), asking and answering open questions ([Herft, 2023](#)), providing immediate feedback on learner performance in exercises ([Baidoo-Anu & Ansah, 2023](#)), and analysing writings from the required perspectives ([Kohnke et al., 2023](#)). The findings also revealed ChatGPT's high quality in teaching, humanity, affect and accessibility aspects, aligning with [Smutny and Schreiberova's \(2020\)](#) criteria for top-level educational chatbots. On the other hand, our study uncovered ChatGPT's unsatisfying ability to create subtle, in-depth learning materials and provide in-depth, insightful responses in discussions and explorations, consistent with researchers' concerns about ChatGPT's effectiveness of teaching specific domain knowledge and professional skills (e.g., [Baidoo-Anu & Ansah, 2023](#)). According to our interviewees, this limitation of ChatGPT could restrict the application and effectiveness of learning strategies, especially Understanding, Analysing, and Crafting, thus preventing the development of knowledge and quality of logic in English argumentative writing. Future practitioners may design and implement ChatGPT-based logic learning based on comprehensive consideration of ChatGPT's affordances and limitations, with reference to our study results. We also call for future efforts to address ChatGPT's limitations in teaching specific domain knowledge and professional skills.

Methodologically, our study highlights the significance of conducting a pilot study in the investigation and implementation of ChatGPT-based logic learning. The pilot study conducted in our research played a crucial role in validating and refining our ChatGPT-based logic learning tool, *LogicalHamster*. By involving five students with backgrounds similar to our target demographic, the pilot study provided an opportunity for participants to interact with *LogicalHamster* for over 45 min, followed by unstructured interviews to gather detailed feedback. Two authors meticulously analysed *LogicalHamster*'s responses in its interactions with the five students, to ensure its conversational style remained consistently friendly and positive. Additionally, we cross-referenced the logic instruction provided by *LogicalHamster* and verified its accuracy and unbiasedness. The feedback from learners confirmed the ease of use and effectiveness of *LogicalHamster* in enhancing knowledge and quality of logic in English argumentative writing. Notably, learners recommended limiting session durations to between 45 and 75 min, which was invaluable in standardising the learning conditions for the subsequent phases of our study. Based on learner feedback from the pilot study, we made important adjustments to the structure and content of ChatGPT-based logic learning. For example, we initially had asked students to learn 11 logical concepts with the assistance of *LogicalHamster*. However, our students admitted that logic learning was cognitively demanding, and they felt that 11 concepts might be overwhelming for a single session. Taking their feedback into consideration, we reduced the target knowledge to seven logical concepts. This refinement was a direct result of the insights gained from the pilot study, which played an instrumental role in improving our educational tool. Given the essential role of the pilot study in our main study and the potential risks associated with ChatGPT producing biased, general, or inaccurate information ([Baidoo-Anu & Ansah, 2023](#)), we strongly recommend that future researchers conduct a pilot study before adjusting and implementing ChatGPT-based learning with reference to our study, so as to ensure the effectiveness of the learning tool.

We also recommended the long-term implementation of ChatGPT-based logic learning. Our interview findings revealed learners' overall difficulty in understanding and utilising knowledge of logic in English argumentative writing, consistent with [El Khoiri and Widiati \(2017\)](#) and [Zhang et al. \(2023b\)](#). According to the interviewees, it requires two or more sessions of ChatGPT-based logic

learning, along with consistent effort, to fully master knowledge of logic. Specifically, Analysing and Crafting strategies proves essential in enhancing quality of logic in English argumentative writing (El Khoiri & Widiati, 2017; Guo & Lee, 2023), while learners may only feel adequately prepared cognitively and emotionally to employ these strategies after participating in one or more sessions of logic learning. By this logic, within the context of ChatGPT-based logic learning, the impacts of Analysing and Crafting strategies may be pronounced only when the learning process continues over an extended period, encompassing no fewer than two sessions. Therefore, following Zhang et al.'s (2024) suggestions on technology-enhanced learning, we strongly encourage future practitioners to implement ChatGPT-based logic learning for a duration exceeding two sessions. This extended implementation will foster noticeable changes in learners' knowledge and quality of logic in English argumentative writing.

Our study has limitations. Firstly, our sample was gender-imbalanced with tertiary or advanced educational levels and no experience in chatbot-based learning or prompt engineering. Further exploration on this topic may be needed based on gender-balanced samples with diverse educational and training backgrounds. Secondly, our study focused on individual engagement in ChatGPT-based learning, while ChatGPT-based learning programmes can also be based on collaborative engagement (Hamid et al., 2023). Researchers may further explore ChatGPT-based learning strategies in collaborative contexts to obtain a more comprehensive understanding of this topic. Thirdly, in this study, we focused on the influences of ChatGPT-based logic learning strategies on learners' quality of logic in argumentative writing in general. However, different strategies may affect learners' mastery of specific logical elements within argumentative writings differently. For instance, our interview findings indicate that Understanding strategies helped students distinguish between various logical concepts. Hence, Understanding strategies might be particularly effective in clarifying logical misconceptions, such as misconceptions between Hasty Generalisation and Post Hoc fallacies. Future research could greatly benefit from a detailed comparative analysis of the initial essays and the revisions, exploring the specific logical elements learners may have addressed through the post-learning revision. Such analysis would provide a clearer illustration of the specific impacts of ChatGPT-based logic learning at a micro-level.

Recently, we have witnessed the rapidly-growing trend of ChatGPT-based learning in various domains and disciplines, while empirical, experimental studies remained scarce on this innovative learning approach. This study looked specifically at this approach in logical education, illuminating its mechanisms and effects, which may remain applicable in other fields. To further explore the potential of ChatGPT-based learning, we call for more contributions in this field, such as ChatGPT-based collaborative learning and ChatGPT-based educational games.

CRediT authorship contribution statement

Ruofei Zhang: Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Di Zou:** Writing – review & editing, Visualization, Supervision, Resources, Methodology, Formal analysis, Conceptualization. **Gary Cheng:** Writing – review & editing, Supervision, Methodology, 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.

Declaration of generative AI in scientific writing

During the preparation of this work, the authors used *Grammarly* in order to polish the language. After using this tool, the authors meticulously reviewed and edited the content and take full responsibility for the content of the publication.

Funding

The research has been supported by the Interdisciplinary Research Scheme of the Dean's Research Fund 2021/22 (FLASS/DRF/IDS-3 2022) of The Education University of Hong Kong.

Appendix A. Participant information

Participant	Age	Gender	Educational level
1.	26	Female	Graduate
2.	18	Female	Undergraduate
3.	27	Female	Graduate

(continued on next page)

(continued)

Participant	Age	Gender	Educational level
4.	19	Male	Undergraduate
5.	22	Female	Graduate
6.	23	Male	Undergraduate
7.	24	Female	Undergraduate
8.	23	Female	Graduate
9.	20	Female	Undergraduate
10.	25	Female	Graduate
11.	23	Female	Graduate
12.	23	Female	Graduate
13.	26	Female	Graduate
14.	26	Female	Graduate
15.	19	Male	Undergraduate
16.	27	Female	Graduate
17.	27	Female	Graduate
18.	22	Male	Graduate
19.	22	Female	Graduate
20.	27	Female	Graduate
21.	24	Male	Undergraduate
22.	19	Female	Undergraduate
23.	27	Female	Graduate
24.	20	Female	Undergraduate
25.	19	Male	Undergraduate
26.	23	Female	Undergraduate
27.	18	Female	Undergraduate
28.	18	Female	Undergraduate
29.	19	Female	Undergraduate
30.	18	Female	Undergraduate
31.	25	Female	Graduate
32.	22	Female	Undergraduate
33.	25	Female	Graduate
34.	27	Female	Graduate
35.	27	Female	Graduate
36.	22	Female	Undergraduate
37.	21	Male	Undergraduate
38.	23	Female	Graduate
39.	25	Male	Graduate
40.	23	Female	Graduate
M = 22.85,		Female: 32;	Undergraduate: 18;
SD = 3.05		Male: 8	Graduate: 14

Appendix B. Screenshots of *LogicalHamster* supporting learning strategies

What is False Analogy?

 LogicalHamster 

False Analogy is when you compare two things that aren't really alike, but you pretend they're twinsies! 🐹🐹 It's like saying, "Because I'm a hamster and I love running in my wheel, humans must love running in giant wheels too!" Silly, right? 😂

Here's the logical structure:

1. Thing A has property X.
2. Thing B has property Y.
3. X and Y are kinda sorta similar.
4. Therefore, things A and B are compared as if they're the same.

To spot it, look for comparisons that are more stretchy than my little hamster legs during yoga! 🧘🐹

Ready to see if you can catch one in the wild? Let's test your skills with an example!

Please further explain the meaning of False Analogy.

 LogicalHamster 

Alrighty! 😊 A False Analogy is a bit like dressing up a cat in a hamster costume and saying, "Oh look, it's a hamster!" We all know it's still a cat, right? It's using a comparison that doesn't hold water because the two things being compared are as different as carrots and cupcakes. 🥕🧁

Here's how you can spot it:

- Look for comparisons that are based on superficial similarities.
- Check if the comparison ignores important differences.
- Ask yourself, "Does this really make sense?" 🤔

So, in an argument, if someone says, "Employees are like nails. Just as nails must be hit on the head to get them to work, so must employees." That's a False Analogy! Employees aren't nails; they're people with feelings and don't need hammering to work. 🦴❤️

Got it, or should I roll out another example? 🎲📝

Fig. 1. LogicalHamster supporting Gathering strategies.

So, False Analogy is about drawing comparisons between incomparable subjects and leading to invalid conclusion -- Am I correct?

 LogicalHamster 

Bingo! 🎉 You got it, Fanny! It's all about those mismatched comparisons leading to a wonky conclusion. It's like saying, "Because I'm a super cute hamster and I can escape my cage, you can escape from your room too!" Nope, not the same thing. 🐹🔒

Please give me some examples about False Analogy.

 LogicalHamster  ...

You got it! Here are some False Analogy examples for you:

1. "Just like a seed grows into a tree with proper care, a business will turn profitable with good management." – It's cute, but not quite right! 🌱💼 A seed's growth involves biological processes, while a business has lots of complex human factors.
2. "Cars have wheels and go vroom, and so do airplanes. Therefore, driving a car is just like flying a plane." – Silly comparison, right? 🚗✈️ They might both transport us, but the skills and mechanics are worlds apart.
3. "A school is like a factory; it takes in raw materials (students) and produces products (graduates). So, schools should be run just like factories to be efficient." – Ouch! Students aren't widgets, and education is way more than assembly lines. 🏫🏭

Remember, just because two things share a trait, doesn't mean they're the same or work the same way. Keep an eye out for those tricky comparisons! 😊 Want to take a shot at figuring out why these analogies are more wobbly than a hamster on a tightrope? 🐹🎪

We use False Analogy when writing stories, novels, and narratives all the time. Why is False Analogy a logical fallacy?

 LogicalHamster  ...

Oh, Fanny, you've got a point as sharp as my little claws! 🐹sharp In stories, novels, and narratives, we use metaphors and similes to spice things up, like adding a bit of seasoning to a salad. Yum! 🥗👉 They're not meant to be taken literally, just to add flavor to the writing.

But in arguments, where we're trying to convince someone of something, we need to be as logical as a hamster planning the most efficient route to its snacks. 🍪💡 A False Analogy in an argument is like trying to use a slice of cheese to open a locked door—it just doesn't work! It's a logical fallacy because it misleads by comparing things that aren't truly comparable to draw a conclusion. It's not about creativity; it's about correctness. 🚫💡

So, in creative writing, be as wild as a hamster on a sugar rush! 🐹➡️ But in arguments, stick to comparisons that are as solid as my love for sunflower seeds! 🌻❤️ Got any other nutty questions, or are you ready for some exercises to test your fallacy-finding skills? 🤓📝

Fig. 2. LogicalHamster supporting Understanding strategies.

Please give me some T/F questions about False Analogy.

 LogicalHamster 

Absolutely, let's do this!   Here come some True/False questions for you:

1. True or False: A False Analogy is when two things with similar appearances are compared as if they are identical in function or nature.
2. True or False: "Students are like batteries; if you charge them with knowledge, they will go a long way." This is a valid analogy because it encourages education.
3. True or False: "Just as watering plants helps them grow, providing feedback to employees helps them improve." This statement is a False Analogy.
4. True or False: False Analogies are acceptable in argumentative writing as long as they make the writing more interesting.

Nuzzle your brain and tell me what you think!   When you're ready, I'll reveal the truths and the fluffy falsehoods!   

...
1. True. False Analogy is about assuming what are similar in appearances are also similar in nature.

 LogicalHamster 

Oopsie, let's nibble that one a bit more!   False Analogies are not just about appearances, but also about assuming that because two things share one or two similar traits, they are similar in other ways, which leads to an incorrect conclusion. It's a tricky one, so keep your whiskers sharp! 

Would you like to take a crack at the other statements?  

Fig. 3. LogicalHamster supporting Exercising strategies.

From which perspectives can I analyse the following essay from the logical perspective? "How to Keep Psychologically Healthy?"



LogicalHamster Poe

Alright, you brainy critter, let's put on our detective hats and sniff out the logic! 🐹🐹 Here's how we can analyze the essay:

- Check for Unsupported Claims:** Are there statements without evidence? Like, it says "one in four people will experience some kind of mental health problem," but where's the data to back it up? 📊
- Look for Hasty Generalizations:** Does the essay make broad claims from a small sample? Like saying mental illness always leads to suicide without considering other outcomes. 🚨
- Search for Causal Fallacies:** Does it imply a cause-effect relationship that might not be there? For example, it assumes that not taking care of psychological health will definitely worsen it and raise suicide rates. 🔗
- Spot Any False Analogies:** Are there comparisons that don't quite match up? This essay seems free from this furry faux paw, though. 🐾
- Beware of Slippery Slopes:** Does the conclusion predict a dire outcome without clear evidence? The essay suggests that neglecting psychological health will surely increase suicide rates, which might be a bit of a leap. 🚫
- Examine the Structure of Arguments:** Are claims logically connected to the conclusion? We want a solid bridge, not a wobbly one! 🏗

Let's rate this essay's logical quality... I'd give it a cute 70/100! 🐹💯 It's got some good points, but we need to patch up a few logic holes.

Fig. 4. LogicalHamster supporting Analysing strategies.

Please give me some references to modify the sentence which implies cause-effect relationship.

 LogicalHamster  ...

Oh, you're diving deep into the logic burrow!   Here's how we can tweak that cause-effect sentence to be as logically cozy as a hamster in fresh bedding:

Original: "If we do not take care of our psychological health, our psychology may get worse and worse. Eventually, the suicide rate may raise."

Modified: "Neglecting our psychological health could contribute to its deterioration, which may be one of several factors that, if not addressed, could lead to an increased risk of severe mental health issues."

Look at the change! We're not saying it's the only cause, just a potential contributor. It's like saying, "If I don't fill my cheeks with seeds, I might get hungry," instead of "If I don't fill my cheeks with seeds, I'll definitely starve." One's a possibility, the other's a bit dramatic, even for a hamster!  

Got it? Or is your brain doing somersaults? Let me know, and I'll help you land back on your cute little feet!  

I have modified the logic of the last sentence as follows. What do you think about it? "If we consistently neglect our psychological health, it may deteriorate, potentially contributing to an increased risk of more severe mental health issues."

 LogicalHamster  ...

Oh, you clever human, you've done it!  That sentence is now as logically sound as a hamster wheel in full spin!  You've removed the slippery slope by not directly linking neglect to suicide rates. Instead, you've shown a potential risk for more severe issues, which is a more cautious and logical approach. A round of applause and a tiny high-paw to you!  

Let's give that sentence a sparkly 100/100 for logical improvement!   You've taken the acorn and turned it into a mighty oak of reasoning!  

Keep up the great work, and if you have more sentences to tweak or want to dive deeper into the world of logical fallacies, just let this tiny teacher know!  

Fig. 5. LogicalHamster supporting Crafting strategies.

Appendix C. Knowledge of logic test (Zhang et al., 2023a)

Instruction:

Each sentence in **column [A]** typically commits one of the seven commonly-seen reasoning errors in argumentative writing, as follows.

- | | | |
|-------------------------|-------------------------|-----------------------|
| A. Begging the Question | B. Red Herring | C. False Alternatives |
| D. Post hoc fallacy | E. Hasty Generalization | F. Slippery Slope |
| G. Faulty Analogy | | |

Please read the seven sentences in **column [A]** carefully.

- a) explain how and why each sentence commits a reasoning error in **column [B]**,
 b) and match the term of each reasoning error in **column [C]** (e.g., A. Begging the Question).

[A]	[B] How and why does this sentence commit a reasoning error?	[C] What's the term of the reasoning error?
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 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 D. Scoring system of the tests on knowledge of logic (Zhang et al., 2023a)

Categories	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 E. Scoring system of English argumentative writing (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
Commitment of logical fallacies	The frequencies and depths of logical fallacies 0~25
Coherence	The linking of ideas through logical sequencing throughout the students' writings 0~25
Total	0~100

Appendix F. Coding scheme of the interview data with a pseudo sample

Interviewee number	Logic learning strategies	Regarding RQ1: The frequency of learner application of different types of logic learning strategies		Regarding RQ2: Perceived impacts of different types of logic learning strategies on learning outcomes	
		Very frequently, somewhat frequently, sometimes, seldom, or never	Reason	Positive, negative, neutral, or mixed	Reason
1	Gathering	Somewhat frequently	Gathering is easy to apply.	Positive	I learned knowledge of logic mainly from reading the instructions.

(continued on next page)

(continued)

Interviewee number	Logic learning strategies	Regarding RQ1: The frequency of learner application of different types of logic learning strategies		Regarding RQ2: Perceived impacts of different types of logic learning strategies on learning outcomes	
		Very frequently, somewhat frequently, sometimes, seldom, or never	Reason	Positive, negative, neutral, or mixed	Reason
	Understanding	Somewhat frequently	Discussing logic concepts with <i>LogicalHamster</i> is interesting.	Positive	I understood logical concepts better by discussing them with <i>LogicalHamster</i> .
	Exercising	Very frequently	<i>LogicalHamster</i> 's immediate feedback motivated me to do more exercises.	Positive	My memory of logical concepts was reinforced by doing the exercises.
	Analysing	Seldom	Analysing is too difficult.	Neutral	<i>LogicalHamster</i> 's suggestions for analysing logic were too general.
	Crafting	Seldom	Crafting is too difficult.	Positive	I become able to use logical concepts in my own writing.

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

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

Variables	Indicators	Factor loadings	rhoC	AVE
Gathering strategies	Asking for instructions	0.653	0.769	0.632
Exercising strategies	Asking for further explanations	0.915		
	Asking for exercises	0.795	0.874	0.777
	Doing exercises	0.960		
	Adjusting the format and content of exercises	0.700		
Understanding strategies	Asking exploratory questions	0.752	0.806	0.511
	Paraphrasing the logical concepts	0.736		
	Asking for examples	0.641		
	Answering exploratory questions	0.724		
Analysing strategies	Asking for advice for analysing logic in authentic writings	1.000	1.000	1.000
Learners' prior knowledge of logic in English argumentative writing	Pre-test scores	1.000	1.000	1.000
Learners' prior quality of logic in English argumentative writing	Pre-writing scores	1.000	1.000	1.000
ChatGPT-based logical learning outcomes concerning knowledge of logic in English argumentative writing	Immediate post-test scores	0.948	0.936	0.880
ChatGPT-based logical learning outcomes concerning quality of logic in English argumentative writing	Delayed post-test scores	0.928		
	Writing scores after the learning	1.000	1.000	1.000

Table 2
The heterotrait-monotrait ratio of correlations (HTMT) indicating the discriminant validity of the measurement models of PLS-SEM 1

Outcomes	Predictors					
	Learners' prior knowledge of logic in English argumentative writing	Learners' prior quality of logic in English argumentative writing	Gathering strategies	Exercising strategies	Understanding strategies	Analysing strategies
Pre-test scores
Pre-writing scores	0.038
Gathering strategies	0.176	0.304
Exercising strategies	0.140	0.284	0.482	.	.	.
Understanding strategies	0.320	0.211	0.287	0.231	.	.
Analysing strategies	0.314	0.059	0.237	0.580	0.333	.
ChatGPT-based logical learning outcomes concerning knowledge of logic in English argumentative writing	0.488	0.167	0.499	0.310	0.485	0.165
ChatGPT-based logical learning outcomes concerning quality of logic in English argumentative writing	0.036	0.717	0.117	0.038	0.300	0.041

Table 3

Variance inflation factors (VIF) indicating the collinearity of the PLS-SEM structural model

Exogenous constructs	Endogenous constructs					
	Learners' prior knowledge of logic in English argumentative writing	Learners' prior quality of logic in English argumentative writing	Gathering strategies	Exercising strategies	Understanding strategies	Analysing strategies
Gathering strategies	1.146	1.002	n/a	n/a	n/a	n/a
Exercising strategies	1.146	1.002	n/a	n/a	n/a	n/a
Understanding strategies	1.146	1.002	n/a	n/a	n/a	n/a
Analysing strategies	1.146	1.002	n/a	n/a	n/a	n/a
ChatGPT-based logical learning outcomes concerning logical knowledge of English argumentative writing	1.151	n/a	1.094	1.441	1.159	1.470
ChatGPT-based logical learning outcomes concerning quality of logic in English argumentative writing	n/a	1.219	1.078	1.596	1.145	1.576

References

- Adamopoulou, E., & Moussiades, L. (2020). An overview of chatbot technology. In I. Maglogiannis, L. Iliadis, & E. Pimenidis (Eds.), *IFIP advances in information and communication technology: vol. 584. Artificial intelligence applications and innovations 2020* (pp. 373–383). Cham: Springer. https://doi.org/10.1007/978-3-030-49186-4_31.
- AlAfnan, M. A., Dishari, S., Jovic, M., & Lomidze, K. (2023). Chatgpt as an educational tool: Opportunities, challenges, and recommendations for communication, business writing, and composition courses. *Journal of Artificial Intelligence and Technology*, 3(2), 60–68. <https://doi.org/10.37965/jait.2023.0184>
- Alharbi, W. (2023). *AI in the foreign language classroom: A pedagogical overview of automated writing assistance tools*. Education Research International. <https://doi.org/10.1155/2023/4253331>, 2023.
- Bai, B., & Wang, J. (2023). The role of growth mindset, self-efficacy and intrinsic value in self-regulated learning and English language learning achievements. *Language Teaching Research*, 27(1), 207–228. <https://doi.org/10.1177/1362168820933190>
- 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. <https://dergipark.org.tr/en/pub/jai/issue/77844/1337500>.
- Bassham, G., Irwin, W., Nardone, H., & Wallace, J. (2010). Logical fallacy. In M. P. Brown (Ed.), *Critical thinking: A student's introduction* (6th ed., pp. 140–178). McGraw-Hill Education.
- Bloom, B. S., & Krathwohl, D. R. (2020). *Taxonomy of educational objectives: The classification of educational goals. Book 1, Cognitive domain*. Longman.
- 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>
- Chu, S. T., Hwang, G. J., Chien, S. Y., & Chang, S. C. (2023). Incorporating teacher intelligence into digital games: An expert system-guided self-regulated learning approach to promoting EFL students' performance in digital gaming contexts. *British Journal of Educational Technology*, 54(2), 534–553. <https://doi.org/10.1111/bjet.13260>
- de Swart, H. H. (2018). Fallacies and unfair discussion methods. In *Philosophical and mathematical logic* (pp. 489–530). Cham: Springer.
- Deci, E. L., & Ryan, R. M. (2012). Self-determination theory. In P. A. M. Van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 416–436). Sage Publications Ltd. <https://doi.org/10.4135/9781446249215.n21>
- DeLancey, C. (2017). *A concise introduction to logic*. Open SUNY Textbooks.
- 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>
- El Khoiri, N., & Widjati, U. (2017). Logical fallacies in EFL learners' argumentative writings. *Dinamika Ilmu*, 17(1), 71–81. <https://doi.org/10.21093/di.v17i1.638>
- Feuerriegel, S., Hartmann, J., Janiesch, C., & Szcech, P. (2024). Generative AI. *Business & Information Systems Engineering*, 66(1), 111–126. <https://doi.org/10.1007/s12599-023-00834-7>
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications Ltd.
- Gloger, I., Schwonke, R., Holzapfel, L., Nückles, M., & Renkl, A. (2012). Learning strategies assessed by journal writing: Prediction of learning outcomes by quantity, quality, and combinations of learning strategies. *Journal of Educational Psychology*, 104(2), 452–468. <https://doi.org/10.1037/a0026683>
- 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.
- Hale, B. (2002). Basic logical knowledge. *Royal Institute of Philosophy Supplements*, 51, 279–304. <https://doi.org/10.1017/S1358246100008171>
- Hamid, H., Zulkifli, K., Naimat, F., Yaacob, N. L. C., & Ng, K. W. (2023). Exploratory study on student perception on the use of chat AI in process-driven problem-based learning. *Currents in Pharmacy Teaching and Learning*, 15(12), 1017–1025. <https://doi.org/10.1016/j.cptl.2023.10.001>
- Herft, A. (2023). A teacher's prompt guide to ChatGPT aligned with 'What Works Best' guide. @HerftEducator. Retrieve on November 20th 2024 from. <https://www.herfeducator.com/>.
- Hsiao, J. C., & Chang, J. S. (2023). Enhancing EFL reading and writing through AI-powered tools: Design, implementation, and evaluation of an online course. *Interactive Learning Environments*, 1–16. <https://doi.org/10.1080/10494820.2023.2207187>
- Hundley, 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>.
- Hwang, G. J., & Chang, C. Y. (2021). A review of opportunities and challenges of chatbots in education. *Interactive Learning Environments*, 1–14. <https://doi.org/10.1080/10494820.2021.1952615>
- 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>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, , Article 00336882231162868. <https://doi.org/10.1177/00336882231162868>

- Kolb, D. A., Boyatzis, R. E., & Mainemelis, C. (2014). Experiential learning theory: Previous research and new directions. In R. J. Sternberg, & L. F. Zhang (Eds.), *Perspectives on thinking, learning, and cognitive styles* (pp. 227–247). Routledge.
- Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, 28 (1), 973–1018. <https://doi.org/10.1007/s10639-022-11177-3>
- Learning Everest. (2024). Merrill's first principles of instruction. Retrieved March 7, 2024, from <https://www.learningeverest.com/merrills-first-principles-of-instruction/>.
- Li, K. C., & Wong, B. T. M. (2021). Features and trends of personalised learning: A review of journal publications from 2001 to 2018. *Interactive Learning Environments*, 29(2), 182–195. <https://doi.org/10.1080/10494820.2020.1811735>
- Mayes, T., & de Freitas, S. (2013). Technology-enhanced learning: The role of theory. In H. Beetham, & R. Sharpe (Eds.), *Rethinking pedagogy for a digital age* (2nd ed., pp. 41–54). Routledge.
- Mayring, P. (2004). Qualitative content analysis. *A companion to qualitative research*, 1(2), 159–176.
- Merrill, M. D. (2002). First principles of instruction. *Educational Technology Research & Development*, 50, 43–59. <https://doi.org/10.1007/bf02505024>
- Murray, N. (2012). *Writing essays in English language and linguistics: Principles, tips and strategies for undergraduates*. Cambridge University Press.
- Nazari, N., Shabbir, M. S., & Setiawan, R. (2021). Application of artificial intelligence powered digital writing assistant in higher education: Randomized controlled trial. *Heliyon*, 7(5), Article e07014. <https://doi.org/10.1016/j.heliyon.2021.e07014>, 1–9.
- 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>.
- Nisbet, J., & Shucksmith, J. (2017). *Learning strategies*. Routledge.
- Oaksford, M., & Chater, N. (2020). New paradigms in the psychology of reasoning. *Annual Review of Psychology*, 71, 305–330. <https://doi.org/10.1146/annurev-psych-010419-051132>
- Pérez-Soler, S., Guerra, E., & De Lara, J. (2020). Model-driven chatbot development. In *International conference on conceptual modeling* (pp. 207–222). Cham: Springer International Publishing.
- Riazi, A. M., & Farsani, M. A. (2023). Mixed-methods research in applied linguistics: Charting the progress through the second decade of the twenty-first century. *Language Teaching*, 1–40. <https://doi.org/10.1017/S0261444823000332>
- 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.ezproxy.eduhk.hk/10.30495/JAL.2020.675547>.
- Schenck, J., & Cruickshank, J. (2015). Evolving Kolb: Experiential education in the age of neuroscience. *Journal of Experiential Education*, 38(1), 73–95. <https://doi.org/10.1177/1053825914547153>
- Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist*, 26(3–4), 207–231. <https://doi.org/10.1080/00461520.1991.9653133>
- Selpia, D., & Setyarini, S. (2020). Students' reasoning and fallacies in Indonesian EFL classroom. In *Twelfth conference on applied linguistics (CONAPLIN 2019)* (pp. 20–24). Atlantis Press. <https://doi.org/10.2991/assehr.k.200406.005>
- Smutny, P., & Schreiberova, P. (2020). Chatbots for learning: A review of educational chatbots for the facebook messenger. *Computers & Education*, 151, Article 103862. <https://doi.org/10.1016/j.compedu.2020.103862>
- 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>
- Sooodmand Afshar, H., Movassagh, H., & Radi Arbab, 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>
- Tutkun, O. F., Güzel, G., Körögü, M., & İlhan, H. (2012). Bloom's revised taxonomy and critics on it. *The Online Journal of Counselling and Education*, 1(3), 23–30. https://www.researchgate.net/profile/Ozkan-Cikrikci/publication/259642504_University_Students'_Ways_Of_Coping_With_Stress_Life_Satisfaction_And_Subjective_Well-Being/links/59200e71a6fdcc4443efb172/University-Students-Ways-Of-Coping-With-Stress-Life-Satisfaction-And-Subjective-Well-Being.pdf#page=28
- Vemprala, S. H., Bonatti, R., Bucker, A., & Kapoor, A. (2024). ChatGPT for Robotics: Design principles and model abilities. *IEEE Access*, 12, 55682–55696. <https://doi.org/10.1109/access.2024.3387941>
- Wong, J., Baars, M., Davis, D., Van Der Zee, T., Houben, G. J., & Paas, F. (2019). Supporting self-regulated learning in online learning environments and MOOCs: A systematic review. *International Journal of Human-Computer Interaction*, 35(4–5), 356–373. <https://doi.org/10.1080/10447318.2018.1543084>
- Zhang, R., & Zou, D. (2022). Types, purposes, and effectiveness of state-of-the-art technologies for second and foreign language learning. *Computer Assisted Language Learning*, 35(4), 696–742. <https://doi.org/10.1080/09588221.2020.1744666>
- 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. (2023a). Chatbot-based learning of logical fallacies in EFL writing: Perceived effectiveness in improving target knowledge and learner motivation. *Interactive Learning Environments*, 1–18. <https://doi.org.ezproxy.eduhk.hk/10.1080/10494820.2023.2202374>
- Zhang, R., Zou, D., & Cheng, G. (2023b). Chatbot-based training on logical fallacy in EFL 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-assisted 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. (2024). Technology-enhanced language learning with null and negative results since 2000: A systematic review based on the activity theory. *Education and Information Technologies*, 29(4), 5017–5077. <https://doi.org/10.1007/s10639-023-11993-1>

Ruofei Zhang is a lecturer at Centre for English and Additional Languages, Lingnan University, Hong Kong SAR, China. Her research interests include second language acquisition, technology-enhanced language learning, game-based learning, chatbot-based learning, and self-regulated learning.

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*.

Gary Cheng is an Associate Professor of the Department of Mathematics and Information Technology at The Education University of Hong Kong. He has been involved in research projects funded by EduHK and the Research Grant Council of Hong Kong on a range of topics mainly related to technology enhanced learning. His research interests include Information Technology Supported L2 Learning, ePortfolio-mediated Learning, Computer Programming Education, Online Learning Management System, e-Assessment, and Learning Analytics.