

Collaborating with ChatGPT in argumentative writing classrooms[☆]

Yanfang Su ^a, Yun Lin ^{b,*}, Chun Lai ^c

^a Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University, Hong Kong, China

^b College of Chinese Language and Culture, Jinan University, Guangzhou, China

^c Faculty of Education, The University of Hong Kong, Hong Kong, China



ARTICLE INFO

Keywords:

Argumentative writing
ChatGPT
Chatbots

ABSTRACT

Composing high-quality argumentative writing requires consideration of the dialogical, structural, and linguistic aspects of argumentation. However, finding an ideal peer student to practice argumentation skills can be challenging, and providing feedback on lower-level language concerns and complex issues related to content and organization can be time-consuming for teachers. In this article, we suggest the integration of ChatGPT, a chatbot released by OpenAI in 2022, into argumentative writing classrooms as a promising solution. The article explores the possibilities of ChatGPT in assisting students with tasks such as outline preparation, content revision, proofreading, and post-writing reflection. We also acknowledge the limitations of ChatGPT and offer suggestions for future considerations in teaching and research.

1. Introduction

Argumentative writing is widely recognized as an essential aspect of academic literacy (Zhu, 2001) and is a high-stake task in standardized tests such as Scholastic Aptitude Test (SAT) (Aull & Ross, 2020; Hartwell & Aull, 2022). Argumentative writing involves formulating a clear and logically-sound claim supported by evidence and reasoning to persuade others to accept one's position on a topic. High-quality argumentative writing demands the soundness of the structural, dialogical, and language aspect of argumentation. From the structural perspective, high-quality argumentative writing requires students to incorporate different structural components such as valid claim, relevant, adequate, and credible evidence, appropriate reasoning, and effective rebuttal (McNeil & Krajcik, 2009; Su, Liu, Lai, & Jin, 2021). The dialogical aspect emphasizes acknowledging and addressing multiple perspectives (Chin & Osborne, 2010). The language aspect underscores effective language use, such as careful selection of words to avoid ambiguity in expressing

[☆] Metanote: Purposeful and principled use of generative AI tools like ChatGPT can benefit scholars' work as well, assisting different aspects of academic research and publishing, such as brainstorming, language polishing, and format organizing. However, it is vital to check the accuracy of AI-generated content by consulting multiple reliable sources rather than blindly trusting it without critical evaluation. Researchers should also acknowledge the use of generative AI by adhering to appropriate referencing conventions. In the process of drafting this essay, ChatGPT functioned as a database for clarifying technical terms in the preparation stage. We searched ChatGPT for layman's explanations on some technical terms, such as reinforced learning, to gain a better understanding of the mechanism of large language models and ChatGPT. Throughout the writing process, we constantly referred to ChatGPT when checking the meaning and translation of certain expressions and eliciting alternative expressions. Upon finishing the review draft, we also used ChatGPT to polish the language.

* Corresponding author.

E-mail addresses: yvonesuuu1896@gmail.com (Y. Su), toyunlin@163.com (Y. Lin), laichun@hku.hk (C. Lai).

opinions, strategic use of linking words to enhance coherence and cohesion of writing, as well as maintenance of persuasive writing style.

Although high-stake argumentative writing tasks often prioritize one-sided stance-taking with effective refutation to deconstruct the counterarguments (Bakhtin, 1986; Hartwell & Aull, 2022), argumentative writing is an activity embedded in social contexts and relies on dialogical interaction that acknowledges multiple perspectives and addresses disagreement (Aull & Ross, 2020, p. 24). Dialogical interaction is particularly important for both L1 and L2 language learners, as they often encounter challenges in argumentative writing from structural, dialogical, and language-related perspectives (Wagner, Parra, & Proctor, 2017). In this regard, timely and effective feedback is essential. However, teachers often find it time-consuming to provide detailed feedback on both language and content issues (Guo, Wang, & Chu, 2022; Xu & Zhang, 2022). Educators have hence explored the potential of various technologies in providing support, including the use of automated writing evaluation (AWE) systems to provide corrective feedback on language features and substantial feedback on argument structure (Shi, Liu, Lai, & Jin, 2022), and the utilization of AI-based chatbots as learning partners to interact with students and scaffold their argumentative writing (Guo, Wang, & Chu, 2022; Lin & Chang, 2020).

However, both technological solutions have limitations: AWE systems fail to attend to the social aspect of argumentative writing (Godwin-Jones, 2022). Whereas early versions of chatbots frequently encounter communication breakdown (Yin & Satar, 2020), and have limited capacity to provide detailed feedback on students' argumentative writing (Guo, Wang, & Chu, 2022; Lin & Chang, 2020).

ChatGPT, the new chatbot model fine-tuned from GPT 3.5, exhibits an enhanced ability to understand natural language, improved efficiency and accuracy in responding to queries, and boosted adaptability (Rudolph, Tan, & Tan, 2023). It hence has the potential to assist students in coping with the dialogical, structural and linguistic challenges in argumentative writing, as well as adapting to different pedagogical contexts. In this article, we explore the integration of ChatGPT in argumentative writing classrooms, focusing on its potential to support pre-writing planning of outline, during-writing revision of argument content and language, and post-writing reflection.

2. Chatbots and argumentative writing

Chatbots are intelligent conversational systems designed to interact with human users through natural languages (Luo, Lau, Li, & Si, 2022). While chatbots have been used in language learning settings for decades (Huang, Hew, & Fryer, 2022), research exploring the use of chatbots for argumentative writing development demonstrated mixed results. Goda, Yamada, Matsukawa, Hata, and Yasunami (2014) tried to use an ELIZA-based chatbot to facilitate group argumentation activities but observed no significant effects. In contrast, Lin and Chang (2020) observed the positive effects of using a chatbot to assist in the outline drafting of argumentative essays. Nevertheless, they attributed the positive results largely to the novelty effect. Guo, Wang, & Chu, 2022 designed *Argumate* to scaffold rebuttals and counterarguments during the argumentative writing process. Although *Argumate* was well-intentioned, it cannot understand students' input and only provides pre-defined scaffoldings on limited argumentation topics step-by-step, regardless of students' input. For example, when Argumate guided the student to provide specific reasons for the claim, even if the student responded with "I don't have any specific reasons for this opinion", Argumate would still reply with "Yes, that's a good point!" and prompt the students to further elaborate on the evidence supporting the claim. Despite the potential of these early chatbots in utilizing the conversational nature of chatbots to support the dialogic process of argumentative writing, some limitations can be found. Early chatbots are not based on generative models and, therefore, can only provide pre-set responses rather than dynamic and personalized feedback (Goda et al., 2014; Guo, Wang, & Chu, 2022). They are also limited in natural language understanding (NLU). Hence, students can neither converse with them freely on problems encountered during writing nor seek detailed feedback on a full text as AWE systems do. In addition, chatbots generally do not offer scaffolding for students' language use in argumentative writing.

AWE systems also have limitations. Although much acknowledged as an effective technological tool for writing development with an overall effect size of 0.59 (Ngo, Chen, & Lai, 2022), AWE systems are primarily designed to enhance meta-cognitive awareness and self-monitoring of writing through timely feedback without possibilities for further negotiation and discussion. Scholars hence worry that AWE might exclude students from authentic writing contexts (Herrington & Moran, 2012; Vojak, Kline, Cope, McCarthey, & Kalantzis, 2011), and therefore cannot cope with the dialogic challenges of argumentative writing. Moreover, AWE systems are limited in providing higher-order feedback from the structural perspective, such as the effectiveness of argumentation and the coherence and cohesion of writing (Godwin-Jones, 2022). AWE systems might also provide abstract feedback that is hard to understand (Ranalli, 2018; Zhang, 2020).

Recently, AI-powered agents using a machine learning approach have made significant progress. OpenAI's GPT-series models, especially ChatGPT, represent the most recent technical leap of chatbots (Godwin-Jones, 2022). The GPT models were trained on a large tokenized language dataset to create a deep learning neural network, which allows the models to comprehend the text data and generate human-like responses by predicting the subsequent words within a given string of words. To refine ChatGPT's ability to respond to a variety of prompts, the model was fine-tuned using Reinforcement Learning from Human Feedback (RLHF) (OpenAI, 2022), where the AI model learns to adjust its behavior and improve performances based on the reward model created by human feedback. ChatGPT also possesses the ability to maintain conversational context so that it can utilize prior exchanges to inform subsequent interactions and respond to users' requests for modifications or clarifications of their discussion (OpenAI, 2022). The reiterative training and refinement process enhances the capacity of ChatGPT to understand natural language, process long input, provide creative responses to a variety of prompts, and maintain a natural conversation flow (Adamopoulou & Moussiades, 2020; Floridi, 2023; Kasneci et al., 2023). However, it should also be admitted that due to the potential bias in the training data, ChatGPT might generate some inappropriate content and need ongoing refinement based on user feedback (OpenAI, 2022).

In argumentative writing contexts, ChatGPT can integrate the affordances of AWE and early chatbot models to support

ChatGPT prompt	Student outline	ChatGPT feedback
<p>ChatGPT prompt</p> <p>Act as an argumentation outliner who is currently working on providing feedback to the outline of efl students' argumentative writing, i.e., providing critical, specific, and constructive feedback to the outline of the writing.</p> <p>Please follow the checklists to evaluate whether the outline has incorporated the essential elements of argumentative writing and whether the outline can address the topic well. Provide critical and specific feedback to the outline with some explanations.</p> <p>Claim: -the outline has a main claim that is relevant and well-suited to the topic; -the subclaims are closely connected and progressed logically to support the main claim</p> <p>Evidence: -the evidence can support the subclaims; -the evidence is from credible and trustworthy sources;</p> <p>Reasoning: -the reasoning provides a justification of why the evidence supports the claim;</p> <p>Counterargument and rebuttal: -the outline includes a counterargument. -the rebuttal can rebut the counterargument to strengthen the argumentation.</p> <p>I will give you the student outline later. Do not start until I provide you with a student writing. Please be specific. Do you understand?</p>	<p>Student outline</p> <p>Topic: should primary schools have a longer school day?</p> <p>Main claim: primary schools should not have a longer school day.</p> <p>Subclaim 1: longer school time increases the burden of students. Evidence 1: primary school students need to stay in school for more than 6 hours.</p> <p>Reasoning 1: longer hours leads to fatigue. Fatigue leads to lack of attention. Lack of attention leads to poor learning efficiency.</p> <p>Subclaim 2: longer school time increases the burden of teachers. Evidence 2: primary school teachers need to spend more time having classes.</p> <p>Reasoning 2: teachers will have less time preparing their lessons.</p> <p>Subclaim 3: longer school time influences students' development as a whole person. Evidence 3: students will have no time for extra-curriculum activities.</p> <p>Reasoning 3: extra-curriculum activities help develop various abilities of students.</p> <p>Counterargument: increased study time leads to better academic performances. Rebuttal: when both teachers and students are too tired and stressful, it is impossible to improve academic performances.</p>	<p>ChatGPT feedback</p> <p>Thank you for providing the outline of the efl student's argumentative writing on whether primary schools should have a longer school day. Here is my feedback:</p> <p>Claim: The main claim is clear, specific, and directly addresses the topic. It is relevant to the topic and takes a position that opposes the idea of longer school days for primary school students.</p> <p>Subclaims: The subclaims are relevant to the main claim, and they logically progress to support the main claim. Each subclaim provides a distinct reason why primary schools should not have a longer school day.</p> <p>Evidence: The evidence provided for each subclaim is relevant and supports the subclaims. However, evidence 1 could benefit from further specification to clarify what is meant by staying in school for more than 6 hours.</p> <p>Reasoning: The reasoning behind each subclaim is well-explained and provides justification for why the evidence supports the subclaims. The reasoning is logical and easy to follow, providing a clear connection between the evidence and subclaims.</p> <p>Counterargument and rebuttal: The outline includes a clear counterargument that suggests longer study time leads to better academic performance. The rebuttal is strong and addresses the counterargument by pointing out that increased study time is not the only factor in academic performance, and it is impossible to improve academic performance when both students and teachers are too tired and stressed.</p> <p>Overall, the outline is well-structured, and the argumentation is strong, specific, and relevant. However, to strengthen the outline, the student could consider providing more specific evidence for subclaim 1 and adding additional evidence to further support subclaims 2 and 3. Additionally, the student could consider incorporating a counterargument related to the potential benefits of longer school days and rebutting it with specific evidence and reasoning.</p>

Fig. 1. Using ChatGPT for outline preparation.

argumentative writing from structural, language, and dialogical perspectives while mitigating their constraints. Similar to AWE systems, the improved ability to process longer input enables ChatGPT to provide feedback on writing samples, which is not achievable by early chatbots (Guo, Wang, & Chu, 2022; Lin & Chang, 2020). Since ChatGPT shows an improved ability to respond to different prompts (OpenAI, 2022), it might also provide adaptive and differentiated support to students' argumentative writing problems, including both lower-level language problems and higher-order problems on the organization and strength of argument. Meanwhile, ChatGPT can potentially address the dialogical challenges of argumentative writing by serving as a virtual learning partner that allows social interaction (Park et al., 2021), and reduces communication hurdles such as social anxiety (Abrams, 2005). With its improved NLU abilities, ChatGPT can deal with communication breakdown better and impose fewer demands on user language accuracy than early chatbots (Fryer, Nakao, & Thompson, 2019; Kasneci et al., 2023; OpenAI, 2022; Yin & Satar, 2020).

While ChatGPT made some improvements compared with early chatbots and AWE systems, the prospects of incorporating ChatGPT into argumentative writing classrooms have yet to be explored. This article aims to bridge the gap and explores the potential of ChatGPT in supporting the writing process.

3. Potential

Argumentation is a complex reasoning process in social contexts to construct knowledge or solve problems through active language use (Duschl & Osborne, 2002; van Eemeren, Henkemans, & Grootendorst, 2002). Considering the structural, dialogic, and language demands of quality argumentative essays and the generative function of ChatGPT (McNeil & Krajcik, 2009; OpenAI, 2022), we conceptualize a dual role of ChatGPT in supporting this process. ChatGPT can serve as both a writing evaluator that provides feedback to scaffold the structural and language aspect of argumentative essays, and a virtual peer that engages in conversations around the writing topic, troubleshoots the writing process, and offers tips to strengthen the dialogic aspect of argumentative writing (Hartwell & Aull, 2022). Given that the process approach to writing is a widely acknowledged effective writing pedagogy (Graham, Gillespie, & McKeown, 2013; Williams & Beam, 2019), we examined the potential of ChatGPT at different stages of writing to provide structural, dialogic, and language support. We test this conceptualization by simulating the use of ChatGPT in scaffolding the writing process of an essay: "Should schools have longer instruction time?".

The screenshot shows a conversation between a student and ChatGPT. The student asks:

**I don't know how to provide specific evidence and reasoning to strengthen the subclaim:
"longer school hours lead to students being "very tired"**

ChatGPT responds:

Here are some ways you could provide specific evidence and reasoning to strengthen the subclaim:

1. Look for research studies or articles that explore the impact of longer school hours on students' well-being. You could cite specific studies that show how extended school hours can negatively affect students' physical and mental health, including increased stress, fatigue, and sleep deprivation.
2. Provide examples of specific situations where students have experienced negative consequences from extended school hours, such as struggling to stay awake during class, feeling too exhausted to participate in extra-curricular activities, or experiencing mental health issues like anxiety or depression.
3. Explore the reasons why longer school hours might lead to fatigue and tiredness. For example, you could discuss how sitting for long periods can be physically exhausting, how lack of breaks or opportunities for physical activity can contribute to mental and physical fatigue, or how high-stress environments can wear students down over time.
4. Compare the effects of longer school hours to those of shorter school days or more flexible schedules. For example, you could compare the mental and physical well-being of students in schools with longer hours to those in schools with shorter hours, or look at how schools that offer more breaks or opportunities for physical activity compare to those with more rigid schedules.

By providing specific evidence and reasoning, you can help support your subclaim and make your argument more persuasive.

Fig. 2. ChatGPT answering a student question on evidence selection.

ChatGPT prompt	Student writing	ChatGPT feedback
<p>ChatGPT prompt</p> <p>Act as a professional argumentative writing tutor who is currently working on tutoring EFL students to write argumentative writing, i.e., providing critical, specific, and constructive feedback to the content and organization of the writing</p> <p>Please follow the checklists to evaluate and give critical, specific, and constructive feedback to a student's writing from the following aspects: claim, evidence, and reasoning</p> <p>Claim:</p> <ul style="list-style-type: none"> (1) The writing is relevant and well-suited to the topic and reflects clear insights and purposes; (2) The main claims and subclaims of the writing demonstrate depth and complexity of thought supported by rich, pertinent explanations, exemplifications and/or details; (3) The subclaims are closely connected and progressed logically to support the main claim; <p>Evidence:</p> <ul style="list-style-type: none"> (4) The writing has relevant evidence to support the claims from credible and trustworthy sources. (5) The writing gives detailed description of the evidence; (6) The writing explicitly and appropriately links evidence with the claim; <p>Reasoning and counterargument:</p> <ul style="list-style-type: none"> (7) The writing makes valid inferences from the evidence, drawing appropriate conclusions and highlighting the implications of the information; (8) The writing incorporates appropriate counterargument and to strengthen the argument being made; (9) The writing uses linking words effectively to ensure that ideas are introduced in a logical sequence, helping the reader to follow the flow of the argument and understand how each point relates to the others. <p>I will give you the student writing later. Please be specific and critical. Do not start until I provide you with a student writing. Do you understand?</p>	<p>The start or the end of each semester has always bring about discussions about whether we should shorter school days. It is a tradition in countries that take education very seriously to have longer school days. However, we propose that it is the time to make the change! Our children need shorter school days.</p> <p>First, longer school time might lead to fatigue and loss of attention rather than higher scores. As a junior high school student, I go to school 7 o'clock every day for morning reading and go back home at about 5:30 in the afternoon. Sometimes I do not have time for my breakfast. So I can't stay focused during morning reading. In the afternoon after 4 o'clock, I am very tired and get hungry easily so can't pay attention to the teachers. If we have longer school days, I can't imagine my learning quality. So it is also impractical to longer the time of each class.</p> <p>Most importantly, longer school time decreases the opportunities of students to become whole person. The time of extracurriculum activities are less. Extracurriculum activities are very important for us to get closer to nature and serve the society which is totally different from learning the textbooks at school. Last week, I went to the Yangtai Mountain with my friends. We learn about Yangtai Mountain's history. It was very interesting. Then we pick up the rubbish on the mountain. We were tired but we all felt very happy. I thought it was unforgettable and educational. Because we learnt protect the Earth from it. SEOW, Poh Sun and PAN, Gary from Singapore Management University found that extracurriculum activities can lead to better academic performance.</p> <p>Although some parents think longer school time means higher scores. This is not always true. When students in school for too long, they are too tired to concentrate. How could they get higher scores in tests?</p> <p>Therefore, for all the reasons above, I vote against longer school time.</p>	<p>ChatGPT feedback</p> <p>Overall, the student writing shows potential but could benefit from some improvements in organization and development of ideas.</p> <p>Claim:</p> <ul style="list-style-type: none"> (1) The writing is relevant and well-suited to the topic and reflects clear insights and purposes; <p>The writing effectively presents the topic of shorter school days and offers a clear and concise claim that is relevant to the discussion.</p> <ul style="list-style-type: none"> (2) The writing provides two well-developed subclaims that support the main claim, providing evidence and personal anecdotes to demonstrate the impact of longer school days on student learning and extracurricular activities. However, the subclaims could benefit from clearer organization and more explicit linking to the main claim. (3) The subclaims are connected to the main claim and follow a logical progression, but could benefit from more explicit linking words and phrases to help the reader follow the flow of the argument. <p>Evidence:</p> <ul style="list-style-type: none"> (4) The writing includes a reference to a study by SEOW, Poh Sun and PAN, Gary from Singapore Management University to support the claim that extracurricular activities can lead to better academic performance. However, the evidence could be more specific and directly linked to the subclaims. (5) The writing does not provide a detailed description of the evidence used. (6) The writing explicitly and appropriately links evidence with the claim; The writing could more effectively link the evidence to the subclaims and main claim. <p>Reasoning and counterargument:</p> <ul style="list-style-type: none"> (7) The writing presents valid inferences from the evidence and personal anecdotes to draw appropriate conclusions about the impact of longer school days on student learning and extracurricular activities. (8) The writing does not incorporate an explicit counterargument, which could strengthen the argument by acknowledging opposing views and responding to them. (9) While the writing does use some linking words and phrases to connect ideas, more could be used to ensure that the reader can follow the progression of the argument and understand the relationships between the claims, evidence, and reasoning. <p>In conclusion, the student writing effectively presents a claim about the need for shorter school days, and supports this claim with evidence and personal anecdotes. However, the writing could benefit from clearer organization, more explicit linking of subclaims to the main claim, and more specific evidence. Additionally, the writing would be strengthened by incorporation of counterarguments and more effective use of linking words and phrases to ensure logical progression of ideas.</p>

Fig. 3. Using ChatGPT for content editing (OpenAI, 2023a).

3.1. The preparation stage – facilitating idea generation and providing feedback on outlines

Conceptual planning is the starting point of argumentative writing (Coirier, Andriessen, & Chanquoy, 1999). ChatGPT can facilitate conceptual planning by playing the virtual tutor role to support idea generation (Guo, Wang, & Chu, 2022) and the evaluator role to assist in outline creation and refinement. To ensure that ChatGPT improves rather than substitutes student efforts, the teacher can ask students to generate an outline for their argumentative writing first by themselves and then use ChatGPT to seek feedback and additional ideas. To tailor the feedback, the teacher should provide students with differentiated evaluation rubrics for argumentative writing outlines based on their learning stages when eliciting feedback from ChatGPT. For instance, if students are unfamiliar with argumentative writing or have a lower level of English proficiency, teachers can give an evaluation rubric based on a simplified CER (claim-evidence-reasoning) framework (McNeil & Krajcik, 2009). For students who have a solid understanding of the simplified structure of argumentation (e.g., the CER framework), evaluation rubrics based on the more complex TAP framework (including claim, data, warrant, qualifier, rebuttal, and backing) (Toulmin, 2003) may be provided (Su, Liu, Lai, & Jin, 2021).

To test how this proposed design might work out, we provided ChatGPT with an outline and an evaluation rubric to assess the feedback ChatGPT generates (see Fig. 1).

ChatGPT feedback pinpointed the lack of supportive evidence for some subclaims in the outline. We then followed up with a demand for tips on providing specific evidence for a claim (see Fig. 2). ChatGPT offered detailed suggestions about how to select evidence with elaborated examples, which could facilitate idea generation.

Thus, in the preparation stage, ChatGPT can provide feedback on the argumentation outline by assessing the connection between claims and subclaims, evaluating the quality of the supporting evidence and rebuttals, and suggesting potential counterarguments and claims, thus potentially helping enhance the structural aspect of the argumentation. However, the quality of feedback depends crucially on the evaluative rubrics that users provide. When evaluative rubrics focus solely on the quality of evidence and claims, ChatGPT may fail to recommend the inclusion of missing components such as counterarguments and rebuttals.

3.2. The editing stage – providing feedback on the draft and supplying different perspectives

The focus at the editing stage is the quality of the content. Quality content involves delineating the claims, organizing a logical flow of the subclaims, selecting evidence from credible sources and elaborating on the evidence with adequate details, justifying how the evidence could support the claim, and rebutting the counterarguments (Berland & McNeill, 2010; Chin & Osborne, 2010; Coirier,

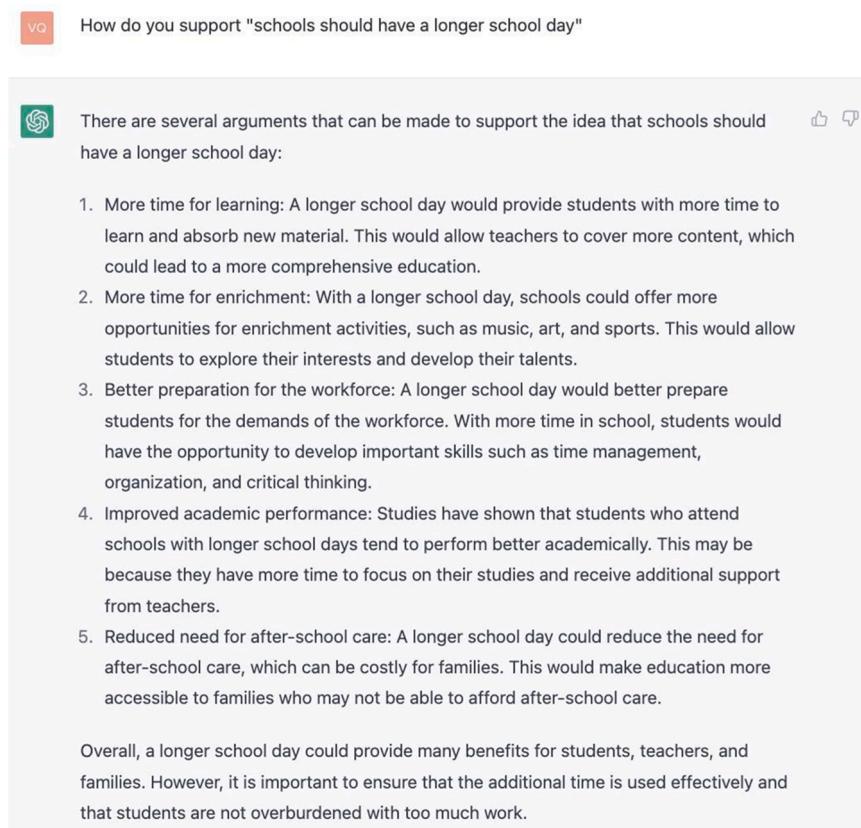


Fig. 4. ChatGPT response to a student question on counterargument (OpenAI, 2023a).

ChatGPT prompt	Student writing	ChatGPT feedback
<p>ChatGPT prompt</p> <p>Act as a professional writing tutor who is currently working on giving feedback to the language of students argumentative writing.</p> <p>Please follow the checklists to proofread a learner writing:</p> <ul style="list-style-type: none"> (1) The writing uses an extensive range of vocabulary and demonstrates very natural and sophisticated control over lexical features in the writing; (2) The writing demonstrates appropriate word choice and idiomacticity; (3) The writing shows a good command of word forms; (4) The writing employs diversified structures with full flexibility and accuracy; (5) The writing contains no grammatical errors, including agreement, tense, number, word order/function, articles, pronouns, and prepositions; (6) The ideas are easy to understand and expressed with clarity; (7) The writing demonstrates a good mastery of conventions; (8) The writing exhibits consistency and no errors in spelling, spacing, punctuation, and capitalization. (9) The writing is appropriate for its intended text type, i.e., argumentative writing. <p>I will give you the student writing later. Please be specific and critical. Do not start until I provide you with a learner writing. Do you understand?</p>	<p>Student writing</p> <p>The start or the end of each semester has always bring about discussions about whether we should shorter school days. It is a tradition in countries that take education very seriously to have longer school days. However, we propose that it is the time to make the change! Our children need shorter school days.</p> <p>First, longer school time might lead to fatigue and loss of attention rather than higher scores. As a junior high school student, I go to school 7 o'clock every day for morning reading and go back home at about 5:30 in the afternoon. Sometimes I do not have time for my breakfast. So I can't stay focused during morning reading. In the afternoon after 4 o'clock, I am very tired and get hungry easily so can't pay attention to the teachers. If we have longer school days, I can't imagine my learning quality. Also, research shows that it is hard for students to stay focused for longer than 45 minutes. So it is also impractical to longer the time of each class.</p> <p>Most importantly, longer school time decreases the opportunities of students to become whole person. However, the Ministry of Education promote the integration of education across different aspects, such as physics, ethical, intellectual and aesthetics, and learning should not be the only purpose of education. Study from the University of Southern Mississippi found that taking part in different activities out of school help students perform better academically and become a better person in their whole life.</p> <p>Although some parents think longer school time means higher scores. This is not always true. When school hours are too long, the teachers will have no time to prepare the class, the students will be too tired to concentrate. How could students achieve better academic performance? Therefore the time of extracurriculum activities after school is very important.</p> <p>Therefore, for all the reasons above, I vote against longer school time.</p>	<p>ChatGPT feedback</p> <ul style="list-style-type: none"> (1) The student writing demonstrates a decent range of vocabulary and adequate control over lexical features. However, there is room for improvement in using more sophisticated vocabulary to enhance the writing's quality. (2) The writing generally shows appropriate word choice and idiomacticity, but there are a few instances where the use of certain words could be improved for clarity and precision. (3) The writing shows a good command of word forms, and there are no major issues in this aspect of the writing. (4) The writing employs some diversified structures, but there is room for more full flexibility and accuracy. Some sentences could be rephrased for clarity and coherence. (5) The writing contains a few minor grammatical errors, such as subject-verb agreement and word order/function. For example, "However, we propose that it is time to make the change" could be improved to "However, we propose that it is time to make a change." Additionally, "If we have longer school days, I can't imagine my learning quality" should be written as "If we had longer school days, I couldn't imagine my learning quality." (6) The ideas are easy to understand and expressed with clarity. (7) The writing demonstrates a good mastery of conventions, but there are a few areas where the writer could improve, such as consistency in capitalization and punctuation. (8) The writing exhibits consistency and no errors in spelling, spacing, punctuation, and capitalization. (9) The writing is appropriate for its intended text type, argumentative writing. <p>Overall, the student writing demonstrates a good understanding of the topic and presents a clear argument against longer school days. However, there is room for improvement in the areas mentioned above to enhance the writing's quality and effectiveness.</p>

Fig. 5. Using ChatGPT to obtain language feedback (OpenAI, 2023a).

(Andriessen, & Chanquoy, 1999; Su, Liu, Lai, & Jin, 2021). To enhance the quality of the writing content, after students finish their first draft based on the argumentative writing outline, they could seek specific content feedback from ChatGPT. The teachers can give ChatGPT evaluation rubrics designed for a specific group of students. We tested the potential of ChatGPT by providing it with evaluation rubrics for argumentative writing content based on Su, Liu, Lai, & Jin, 2021 argument construction scheme and Duncan, Chinn, and Barzilai (2018) "Grasp of Evidence" framework (see Fig. 3). Overall, the feedback provided by ChatGPT accurately identified the strengths and weaknesses of the writing sample and offered some suggestions for improvement. Due to the limited generative ability of ChatGPT, the suggestions were somewhat vague, with abstract comments like "clearer organization," "explicit linking of subclaims to the main claim," and "specific evidence," and required further elaboration. In this case, students can seek further clarification on the suggestions provided by ChatGPT.

We further followed up on one abstract comment in the feedback, "would be strengthened by incorporation of counterarguments", with a question, "How do you support 'schools should have a longer school day'" (see Fig. 4). The response from ChatGPT provided a list of elaborated subclaims supportive of the counterargument, which makes it possible for students to be exposed to multiple perspectives and critically evaluate their own claims and evidence (Anderson, Chaparro, Smolkowski, & Cameron, 2023). This capacity of ChatGPT to generate responses to user-raised questions enables it to scaffold the dialogical aspect of argumentative writing as it provides a relatively low-stake conversation space for students to gain multiple perspectives on the topic (Guo, Wang, & Chu, 2022). However, the scaffolding is limited as ChatGPT can only generate content upon request and cannot prompt guiding questions to scaffold students' thinking. Moreover, the different arguments provided by ChatGPT are overlapping and superficial. They might serve better as idea prompts rather than ready-to-use content in writing.

Thus, ChatGPT can contribute to the dialogical aspect of augmentative writing by exposing students to different perspectives. However, its virtual peer role is constrained by the nature of its generative affordance. It functions at most as a knowledgeable peer who is unfortunately not able to initiate conversations and guide thinking development.

3.3. The proofreading stage – providing error corrections

Adept use of lexical devices, syntactic structures, and textual devices is essential for organizing ideas into coherent argumentative writing, which is always challenging for English learners due to limited language proficiency (Coirier et al., 1999). ChatGPT could help enhance the effectiveness of language use by providing language-related feedback. Likewise, teachers could give ChatGPT specific rubrics of argumentative writing language so that the chatbot can provide specific feedback to the targeted students. Fig. 5 shows an

The ~~startbeginning~~ or ~~the~~-end of each semester ~~has always bring about often sparks~~ discussions about whether we should ~~shortershortsen~~ school days. ~~It is a tradition in-In~~ countries that ~~takeplace great value on~~ education ~~very seriously to have~~, longer school days. ~~are a tradition~~ However, we ~~proposebelieve~~ that ~~it is the time has come~~ to make ~~thea~~ change! ~~Our our~~ children need shorter school days..

~~FirstFirstly~~, longer school ~~time mighthours may~~ lead to fatigue and ~~lesslack~~ of attention ~~rather than higherconcentration, which could result in lower~~ scores. As a junior high school student, I ~~go toarrive at~~ school ~~at 7 o'clockam~~ every day for morning reading and ~~go back home at aboutleave around~~ 5:30 pm in the ~~afternoonevening~~. Sometimes, I ~~do notdon't even~~ have time for ~~my~~-breakfast. ~~So I can't, making it difficult for me to~~ stay focused during morning reading. ~~InBy~~ the afternoon ~~after, around~~ 4 o'clockpm, I ~~andam usually~~ very tired and ~~get hungry-easily so can't, making it difficult for me to~~ pay attention to ~~the teachersin class~~. If we ~~were to~~ have longer school days, I ~~can'tcannot~~ imagine my learning quality. ~~Also improving. Moreover,~~ research ~~showsindicates~~ that ~~it is hard for students find it difficult to stayremain~~ focused for ~~longermore~~ than 45 minutes. ~~So, so~~ it is ~~also impractical to longerlengthen~~ the ~~timeduration~~ of each class..

Most importantly, longer school ~~time decreases the-hours limit students' opportunities of students to become whole person. However well-rounded individuals. While~~ the Ministry of Education ~~promote the promotes education~~ integration of education across different ~~aspectsareas~~, such as physics, ethical, intellectual, aesthetics, and aesthetics, and learning should not be the only ~~purposgoal~~ of education. Study from A study conducted by the University of Southern Mississippi found that taking partparticipating in differentextracurricular activities ~~outoutside~~ of school ~~help~~ helps students perform better academically and become a better person in their whole life. individuals overall.

Although some parents ~~thinkbelieve that~~ longer school ~~time means higher scores. Thishours result in better grades, this~~ is not always ~~true~~ the case. When school hours are too long, ~~the~~-teachers ~~will have no little~~ time to prepare ~~the~~-class, ~~thefor~~ classes, and students ~~will bebecome~~ too tired to concentrate. ~~How~~This could students achieve ~~betternegatively impact~~ academic performance². Therefore, the time of ~~extra~~curriculumally allocated for extracurricular activities after school is ~~very important critical~~.

Therefore, ~~for allbased on~~ the ~~above~~ reasons ~~above~~, I vote against longer school ~~time-hours~~.

Fig. 6. ChatGPT-polished writing (OpenAI, 2023a).

evaluation checklist we provided to ChatGPT and one version of language feedback it generated in response to the checklist. The feedback ChatGPT provided was primarily evaluative but not elaborative, thus lacking learning potential. Moreover, ChatGPT provided different feedback each time the student writing was submitted anew. The inconsistency of the feedback limits the usefulness of the feedback. To cope with the limitations of ChatGPT, teachers or students can run the writing through ChatGPT multiple times and check the consistency of the feedback.

However, ChatGPT is rather useful as a proofreading aid. When submitting individual sentences for grammar check, ChatGPT can judge the grammatical accuracy, elaborate on the meaning that the sentence was trying to convey, and suggest alternative ways of expression. When submitting full text for proofreading, ChatGPT revised grammatical mistakes, including those related to articles, modal verbs, and tense, as well as typos (see Fig. 6). It also modified the lexical choice and syntactic structures to ensure accurate wording, enhance the conciseness of the text, and improve the academic tone. It is the students' responsibility to critically analyze and decide whether to accept the revision made by ChatGPT. Students can further use ChatGPT as a query tool, such as searching for collocation, discerning similar words, and checking grammatical usage.

3.4. The reflection stage – facilitating reflection through the chat history

Reflection is essential for students to incorporate the newly acquired information into long-term memory (Bitchener & Storch, 2016; Su, Liu, Lai, & Jin, 2021). The detailed chat history between students and ChatGPT retained on the system can serve as records of the student writing process and progression. Students can build e-portfolios by exporting the chat history (see Fig. 7) and organizing history based on the type of query or feedback. By frequently and repeatedly reviewing the argumentative writing feedback and conversations, students can consolidate their knowledge and improve their writing quality in future assignments (Malecka, Boud, & Carless, 2022).

To sum up, our test of ChatGPT in scaffolding the writing process revealed that interaction with it might potentially support the structural, dialogical, and language aspects of argumentative writing. Being a generative AI, ChatGPT can provide personalized responses and feedback, evaluate and make suggestions on content organization, conduct linguistic analysis of the language, and proofread the text. However, this generative function is of reactive nature (i.e., in response to questions and to the evaluation criteria provided), depending crucially on the quality of questions initiated and rubrics provided by the users. It is limited in its capacity to provide proactive guidance on writing.

4. Limitations

4.1. Ethical issues

One limitation of using ChatGPT in writing classrooms concerns ethical issues. First, authorship and plagiarism are worth attention. Since ChatGPT is a generative AI model, it has drastically transformed the role of AI-based writing assistance tools from editors or evaluators to authors, or at least co-authors (Dale & Viethen, 2021). The development of AI-based writing tools necessitates an expansion of the conventional definition of plagiarism (Hartwell & Aull, 2021), i.e., to regard copying chatbot-generated texts without appropriate attribution as plagiarism in addition to copying other people's work. While authorship and plagiarism are related to academic integrity, other ethical issues concern the dark side of large language models. Even though OpenAI has endeavored to filter harmful content from ChatGPT, some unwanted biased texts related to violence, sexism, and racism might be generated (Hutson, 2021; Godwin-Jones, 2021). Considering the multiple ethical issues, it is suggested that ChatGPT should be integrated with caution as argumentative writing assistance.

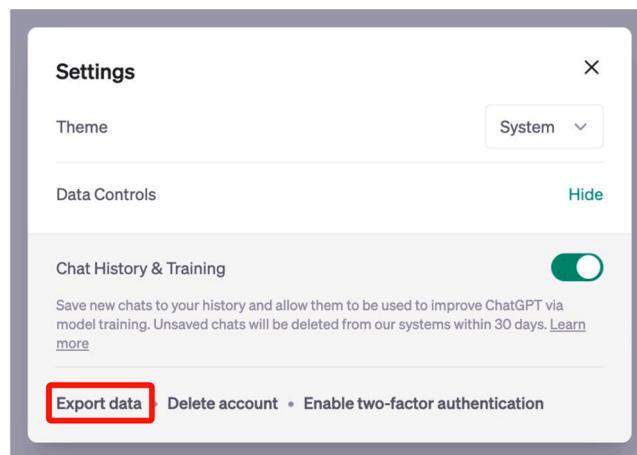


Fig. 7. Export Chat History (OpenAI, 2023b).

4.2. Not intelligent enough

Although ChatGPT can make conversations with users across a wide variety of topics, the system still has not fulfilled expectations of logical reasoning (Zhang & Li, 2021; Thorp, 2023). It guesses the meaning of users' ambiguous expressions rather than asking for clarification, and sometimes might produce "plausible-sounding but incorrect or nonsensical" content or generate fake or misleading responses (Thorp, 2023; OpenAI, 2022). All these problems might be particularly true when the teachers do not set specific goals and clear instructions when using ChatGPT as argumentative writing assistance in classrooms.

5. Future considerations

Proficiency in using AI tools is an important dimension of digital literacy (Godwin-Jones, 2022). The launch of ChatGPT brings both prospects and challenges to argumentative writing instruction. Instead of banning students from using chatbots as writing assistants, training on the effective and ethical use of ChatGPT is essential, where teachers and learners need to be guided on asking effective questions, on ways of using ChatGPT as a thinking aid and reference for personal development, and on addressing ethical issues related to ChatGPT. Training that raises teachers' and students' awareness of the pedagogical strengths and weaknesses of ChatGPT and guides them in exploring ways to maximize the potential of ChatGPT while minimizing its negative consequences in argumentative writing is needed. This essay illustrates one possible way of using ChatGPT to facilitate argumentative writing, and there are other ways of using ChatGPT to develop argumentative writing skills. Regardless, the limitation in the generative feature of ChatGPT demands a collaborative model, where humans and ChatGPT coordinate their strengths in argumentative writing development. Teachers, educators, and schools should work together to explore effective pedagogical approaches and establish robust pedagogical guidelines. Educational software developers could also integrate the ChatGPT model into their apps through the API to capitalize on the educational power of ChatGPT (OpenAI, 2023c).

To cope with the possible issues of academic integrity, detectors of AI-generated texts might help regulate students' use of ChatGPT in argumentative writing. More importantly, teachers are suggested to assign students higher-order tasks that highlight creative ideas, critical thinking, and logical reasoning, which are not the strengths of the system (Stokel-Walker, 2022), rather than the one-sided stance-taking tradition (Aull & Ross, 2020). In addition, even though ChatGPT is intelligent in providing feedback and revising different aspects of student argumentative writing, the problems of nonsensical or inaccurate output exist. Therefore, students need to be guided in critically evaluating the ChatGPT output, and teachers should provide scaffolding and supplementary feedback to students when incorporating ChatGPT into writing classrooms. As for future research, empirical studies on whether and to what extent students can benefit from using ChatGPT as argumentative writing assistance and how students engage with the system during the process are much needed. While both panic and hype around the system have not calmed down, research of this kind will help teachers make more sensible decisions regarding whether and how to integrate ChatGPT into classrooms.

Data Availability

Data will be made available on request.

References

- Abrams, Z. (2005). Asynchronous CMC, collaboration and the development of critical thinking in a graduate seminar in applied linguistics. *Canadian Journal of Learning and Technology/Louisiana Revue canadienne Délelo>t l'apprentissage et Délelo>t Louisiana technologie*, 31(2). <https://doi.org/10.21432/T2G30C>
- Adamopoulou, E., & Moussiades, L. (2020). Chatbots: History, technology, and applications. *Machine Learning with Applications*. Advance online publication,. <https://doi.org/10.1016/j.mlwa.2020.100006>
- Anderson, R. C., Chaparro, E. A., Smolkowski, K., & Cameron, R. (2023). Visual thinking and argumentative writing: A social-cognitive pairing for student writing development. *Assessing Writing*. Advance online publication,. <https://doi.org/10.1016/j.awv.2023.100694>
- Aull, L. L., & Ross, V. (2020). From cow paths to conversation: Rethinking the argumentative essay. *Pedagogy: Critical approaches to teaching Literature, Language, Composition, and Culture*, 20(1), 21–34. <https://doi.org/10.1215/15314200-7878975>
- Bakhtin, M. (1986). *Speech genres and other later essays*. Austin: University of Texas Press.
- Berland, L. K., & McNeill, K. L. (2010). A learning progression for scientific argumentation: Understanding student work and designing supportive instructional contexts. *Science Education*, 94(5), 765–793. <https://doi.org/10.1002/sce.20402>
- Bitchener, J., & Storch, N. (2016). *Written corrective feedback for L2 development* (Vol. 96). Bristol: Multilingual Matters.
- Chin, C., & Osborne, J. (2010). Students' questions and discursive interaction: Their impact on argumentation during collaborative group discussions in science. *Journal of Research in Science Teaching*, 47(7), 883–908. <https://doi.org/10.1002/tea.20385>
- Coirier, P., Andriessen, J., & Chanquoy, L. (1999). From planning to translating: The specificity of argumentative writing. In G. Rijlaarsdam, E. Esperet, J. E. B. Andriessen, & P. Coirier (Eds.), *Foundations of argumentative text processing* (pp. 1–28). Amsterdam, the Netherlands: Amsterdam University Press.
- Dale, R., & Viethen, J. (2021). The automated writing assistance landscape in 2021. *Natural Language Engineering*, 27(4), 511–518. <https://doi.org/10.1017/S1351324921000164>
- Duncan, R. G., Chinn, C. A., & Barzilai, S. (2018). Grasp of evidence: Problematizing and expanding the next generation science standards' conceptualization of evidence. *Journal of Research in Science Teaching*, 55(7), 907–937. <https://doi.org/10.1002/tea.21468>
- Floridi, L. (2023). AI as agency without intelligence: On ChatGPT, large language models, and other generative models. *Philosophy and Technology*. Advance online publication. <https://doi.org/10.2139/ssrn.4358789>
- Fryer, L. K., Nakao, K., & Thompson, A. (2019). Chatbot learning partners: Connecting learning experiences, interest and competence. *Computers in Human Behavior*, 93, 279–289.
- Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology*, 26, 5–24. <http://doi.org/10.1215/73474>.
- Godwin-Jones, R. (2021). Big data and language learning: Opportunities and challenges. *Language Learning & Technology*, 25(1), 4–19. (<http://hdl.handle.net/10125/44747>).

- Graham, S., Gillespie, A., & McKeown, D. (2013). Writing: Importance, development, and instruction. *Reading and Writing*, 26, 1–15. <https://doi.org/10.1007/s11145-012-9395-2>
- Guo, K., Wang, J., & Chu, S. K. W. (2022). *Using chatbots to scaffold EFL students' argumentative writing* (p. 100666). <https://doi.org/10.1016/j.asw.2022.100666>
- Hartwell, K., & Aull, L. (2021). Automated text-matching and writing-assistance tools. *Assessing Writing*. Advance online publication. <https://doi.org/10.1016/j.asw.2021.100562>, 50.
- Hartwell, K., & Aull, L. (2022). Constructs of argumentative writing in assessment tools. *Assessing Writing*. Advance online publication,. <https://doi.org/10.1016/j.asw.2022.100675>, 54.
- Herrington, A., & Moran, C. (2012). Writing to a machine is not writing at all. In N. Elliot, & L. Perelman (Eds.), *Writing assessment in the 21st century: Essays in honor of Edward M. White* (pp. 219–232). New York, NY: Hampton Press.
- Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38(1), 237–257. <https://doi.org/10.1111/jcal.12610>
- Hutson, M. (2021). Robo-writers: The rise and risks of language-generating AI. *Nature*, 591(7848), 22–25. <https://doi.org/10.1038/d41586-021-00530-0>
- 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*. Advance online publication,. <https://doi.org/10.1016/j.lindif.2023.102274>
- Lin, M. P. C., & Chang, D. (2020). Enhancing post-secondary writers' writing skills with a chatbot. *Journal of Educational Technology & Society*, 23(1), 78–92. (<https://www.jstor.org/stable/26915408>)
- Luo, B., Lau, R. Y., Li, C., & Si, Y. W. (2022). A critical review of state-of-the-art chatbot designs and applications. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. Advance online publication,. <https://doi.org/10.1002/widm.1434>
- Malecka, B., Boud, D., & Carless, D. (2022). Eliciting, processing and enacting feedback: mechanisms for embedding student feedback literacy within the curriculum. *Teaching in Higher Education*, 27(7), 908–922. <https://doi.org/10.1080/13562517.2020.1754784>
- Ngo, T. T. N., Chen, H. H. J., & Lai, K. K. W. (2022). The effectiveness of automated writing evaluation in EFL/ESL writing: a three-level meta-analysis. *Interactive Learning Environments*. Advance online publication,. <https://doi.org/10.1080/10494820.2022.2096642>
- OpenAI. (2023a). *ChatGPT* (Feb 13 version) [Large language model]. (<https://chat.openai.com/chat>).
- OpenAI. (2023b). *ChatGPT* (May 3 version) [Large language model]. (<https://chat.openai.com/chat>).
- OpenAI (2023c, March 1). Introducing ChatGPT and Whispers APIs. OpenAI. (<https://openai.com/blog/introducing-chatgpt-and-whisper-apis>).
- OpenAI (2022, November 30). Introducing ChatGPT. OpenAI. (<https://openai.com/blog/chatgpt>).
- Park, S., Thieme, A., Han, J., Lee, S., Rhee, W., & Suh, B. (2021). "I wrote as if I were telling a story to someone I knew.": Designing Chatbot Interactions for Expressive Writing in Mental Health (June) *In Designing Interactive Systems Conference 2021*, 926–941.
- Ranalli, J. (2018). Automated written corrective feedback: How well can students make use of it. *Computer Assisted Language Learning*, 31(7), 653–674. <https://doi.org/10.1080/09588221.2018.1428994>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?. *Journal of Applied Learning and Teaching*. Advance online publication,. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Shi, Z., Liu, F., Lai, C., & Jin, T. (2022). Enhancing the use of evidence in argumentative writing through collaborative processing of content-based automated writing evaluation feedback. *Language Learning & Technology*, 26(2), 106–128. <https://doi.org/10.1257/73481>
- Stokel-Walker, C. (2022). AI bot ChatGPT writes smart essays—should academics worry?. *Nature*. Advance online publication,. <https://doi.org/10.1038/d41586-022-04397-7>
- Su, Y., Liu, K., Lai, C., & Jin, T. (2021). *The progression of collaborative argumentation among English learners: A qualitative study* (p. 102471). <https://doi.org/10.1016/j.system.2021.102471>
- Thorp, H. H. (2023). ChatGPT is fun, but not an author. *Science*, 379(6630), 313–313. <https://doi.org/10.1126/science.adg7879>
- Toulmin, S. E. (2003). *The uses of argument*. Cambridge: Cambridge University Press.
- Vojak, C., Kline, S., Cope, B., McCartery, S., & Kalantzis, M. (2011). New spaces and old places: An analysis of writing assessment software. *Computers and Composition*, 28(2), 97–111. <https://doi.org/10.1016/j.compcom.2011.04.004>
- Wagner, C. J., Parra, M. O., & Proctor, C. P. (2017). The interplay between student-led discussions and argumentative writing. *TESOL Quarterly*, 51(2), 438–449. <https://doi.org/10.1002/tesq.340>
- Williams, C., & Beam, S. (2019). Technology and writing: Review of research. *Computers & Education*, 128, 227–242. <https://doi.org/10.1016/j.compedu.2018.09.024>
- Xu, L., & Zhang, T. (2022). Engaging with multiple sources of feedback in academic writing: postgraduate students' perspectives. *Assessment & Evaluation in Higher Education*. Advance online publication,. <https://doi.org/10.1080/02602938.2022.2161089>
- Yin, Q., & Satar, M. (2020). English as foreign language learner interaction with chatbots: Negotiation for meaning. *International Online Journal of Education and Teaching (IOJET)*, 7(2), 390–410. (<http://iojet.org/index.php/IOJET/article/view/707>).
- Zhang, Z. (2020). Engaging with automated writing evaluation (AWE) feedback on L2 writing: Student perceptions and revisions. *Assessing Writing*, 43, 1–14. <https://doi.org/10.1016/j.asw.2019.100439>
- Zhang, M., & Li, J. (2021). A commentary of GPT-3 in MIT Technology Review 2021. *Fundamental Research*, 1(6), 831–833. <https://doi.org/10.1016/j.fmre.2021.11.011>
- Zhu, W. (2001). Performing argumentative writing in English: Difficulties, processes, and strategies. *TESL Canada Journal*, 19(1), 34–50. <https://doi.org/10.18806/tesl.v19i1.918>

Yanfang Su is a PhD student at the Department of Chinese and Bilingual Studies, The Hong Kong Polytechnic University. Her research interests include blended learning design, computer-assisted language learning, and corpus linguistics.

Yun Lin is a PhD student in the College of Chinese Language and Culture, Jinan University. Her research interests include corpus linguistics and computer assisted language learning.

Chun Lai is an Associate Professor at the Faculty of Education, The University of Hong Kong. Her research interest is technology-enhanced language learning. Her recent research focuses on self-directed language learning with technology beyond the classroom.