The impact of ChatGPT on essay writing performance of students of the Academy of Policy and Development

Pham Thi Hien

Academy of Policy and Development, Vietnam

Abstract: This study examines whether AI-driven corrective feedback delivered via ChatGPT produces genuine improvements in students' academic writing—not just cleaner sentences. We adopted a pre-test / intervention / post-test design with 10 weeks of guided practice. Students interacted with ChatGPT three times per week (≈ 30–60 minutes per session) under explicit guardrails: request feedback, review explanations, revise in their own words, and document changes. Quantitative outcomes compared pre- vs. post-test performance using an analytic rubric emphasizing error identification and correction, coherence, and cohesion. We tested for statistically significant differences and reported effect sizes. Qualitatively, we analyzed students' perspectives on ChatGPT's usefulness for spotting and fixing errors, strengthening logical flow, and aligning drafts with academic writing goals. Conducted with students at the Academy of Policy and Development (APD), the study provides evidence on when AI feedback supports disciplined learning and when it risks superficial polish. Findings aim to guide conservative, integrity-first adoption of AI in writing instruction.

Keyword: ChatGPT; Student; Essay; Academy of Policy and Development.

1. INTRODUCTION

In EFL classrooms, AI-assisted writing has shifted from a niche experiment to a routine part of instruction. Contemporary tools span grammar checkers, drafting assistants, and systems that can generate full pieces of writing (e.g., essays) without direct human composition. Because they are easy to use and efficient, they reduce effort for both students and teachers (Chang et al., 2021; Gayed et al., 2022; Jeanjaroonsri, 2023; Zhao, 2022). They are particularly valuable for learners with limited English proficiency, who benefit from rapid feedback and guidance that accelerate skill development in real time. "AI writing tools have the general purpose of analyzing text content, and providing feedback on various aspects of the text such as grammar, vocabulary, syntax, content, and structure (Hosseini et al., 2023; Strobl et al., 2019; Thorp, 2023)." That feedback is produced by machine-learning models that compare student writing with large repositories of well- and ill-formed language, enabling timely, individualized comments that help writers spot and correct errors quickly. Beyond error repair, this immediacy supports learning the foundations of effective prose organization, clarity, and disciplined revision thereby advancing coherence and cohesion in students' academic writing (Akgun & Greenhow, 2022; Nazari et al., 2021; Rudolph et al., 2023).

2. THEORETICAL OVERVIEW

The recent surge of research centers around the impact of AI writing tools on students' writing skills, which has important implications in the field of Technology-Enhanced Language Learning (TELL). Some studies show that integrating AI-driven writing tools can enhance students' writing abilities (Kurniati & Fithriani, 2022; Wang, 2022; Zhao, 2022), while others Other research expressed concerns about potential adverse effects (Liu et al., 2022; Lund & Wang, 2023; Qadir, 2022). However, the current literature mainly emphasizes the role of AI writing tools in enhancing grammar and syntax, ignoring the broader range of factors important for effective writing. It is worth noting that effective writing includes not only grammar and syntax but also content and organization, both of which play important roles in reader comprehension and engagement (Lee & Yuan, 2021). Content represents the nature of the text, conveys the writer's ideas, thoughts, and messages, and constitutes the content of the story. High-quality content is informative, original, relevant, and meaningful, resonates effectively with audiences, and fulfills its intended purpose, whether to inform, persuade, entertain, or stimulate like to think (Molina et al., 2021).

On the other hand, organization requires the structure of the writing, focusing on how the content is arranged to connect ideas logically and coherently, guiding the reader through the work. Effective organization significantly impacts a reader's comprehension, the persuasiveness of an argument, and the appeal of a story. A well-structured organization ensures a smooth flow of ideas, enhancing the readability and impact of writing (Awada et al., 2020). Therefore, focusing solely on grammar and syntax ignores the important role that content and organization play in effective communication. While using language correctly is fundamental, equal attention should be paid to what the words say (content) and how they are structured organization. Therefore, it is imperative to explore the types of AI writing tools and how they address these important aspects of writing from the perspective of EFL educators.

In the Vietnamese educational context, there is also a similar research gap. As far as the researcher knows, there are very few documents in Vietnamese academic publications that delve deeply into how AI writing tools affect content and organization, considering the perspective of students and teachers. Globally speaking, in the ASEAN region, the authors of this work acknowledge the notable contributions of Miranty and Widiati (2021) and Fahmi and Cahyono (2021), who to some extent have explored the use of Automated Writing Evaluation (AWE) as an AI writing tool for Indonesian EFL students. Miranty and Widiati (2021) investigated students' perceptions of the advantages and disadvantages of AWE, especially Grammarly, while Fahmi and Cahyono (2021) aimed to distinguish differences in students' perceptions of using AWE based on their level of proficiency. Their research highlights the positive benefits of using AWE, a clear consensus in their respective studies. However, further research is needed on the impact of AI writing tools. Although previous research has provided valuable insights into the influence of these tools on students' writing ability, the lack of research in the Vietnamese context calls for further exploration. additional break.

Scholars have engaged in extensive discussions about the pedagogy of academic writing, with various notable figures contributing to this discussion. Prominent names in this field include Cotterall and Cohen (2003), Ferris (2001), Hewings and Hewings (2001), Johns

(1990, 1993, 2003), Johns and Swales (2002), Paltridge (2001), Raimes (1991, 1998), Reid (2001), and Silva (1990). In the field of teaching writing in a second language, authors such as Grabe and Kaplan (2006), Ferris and Hedgcock (2012), Hyland (2019), Hyland (2022), Johns (1997), Kroll (2011), Leki (2008), Silva and Matsuda (2001), Silva and Matsuda (2009), and Wennerstrom (2003) have made significant contributions. Important scholarly discussions regarding responding to writing in a second language are presented in the works of Ferris (2009) and Ferris (2011). Furthermore, academic writing programs around the world are evidenced by the work of Leki (2001). Additionally, academic dialogues address the interrelationship between academic reading and writing as explained by Belcher and Hirvela (2001), Ferris and Hedgecock (2005), and Grabe (2003). However, there is still an ongoing debate surrounding the concept of academic writing.

According to Irvin (2010), academic writing is a form of assessment that requires the writer to demonstrate proficiency and demonstrate knowledge, reflecting disciplinary skills in critical thinking, interpretation, and presentation. Academic writing includes a set of conventions used in the dissertation writing and publication process (Murray 2013). It distinguishes itself as a formal mode of writing, distinct from creative or personal writing (Oshima and Hogue 2007). Irvin (2010) describes academic writing as a literacy task and attaches two important characteristics to it: (1) academic writing is argumentative and (2) academic writing is analytical. The argument is a defining feature of academic writing, cleverly structured to facilitate the presentation of views as a conversational exchange between individuals who may not share identical views but are motivated by a desire to increase understanding of the topic under discussion (Irvin 2010, p. 10). Similarly, according to Irvin (2010), academic writing also requires the analytical presentation of perspectives, emphasizing the investigation of "how and why" questions rather than "what" questions. ". This approach includes activities such as (a) engaging with an open-ended question whose answer is not predetermined, (b) identifying important elements of the topic, and (c) dissecting the individual components while distinguishing their interrelationships.

Swales and Feak (2012) describe academic writing as a meticulously structured product influenced by factors such as audience, flow, organization, purpose, style, and presentation. In their view, these elements are interconnected by the questions "how", "who" and "why". Furthermore, academic writing requires appropriate citations and references, precise punctuation, formal spelling, and a seamless flow of ideas that demonstrates coherence and cohesion. Coherence and cohesion imbue linguistic meaning into discourse or text, facilitating the interpretation of messages and the negotiation of meaning within the text (Poudel 2018).

In this study, academic writing is considered a mode of expression used by writers to delineate intellectual boundaries within their respective fields and areas of expertise. Distinctive features of academic writing include a formal tone, use of the (typical) third-person point of view, uncompromising focus on the research problem at hand, and precise word choice. Similar to the specialized language applied in other professions, such as law or medicine, academic writing is intended to convey general meanings related to complex ideas or concepts in the field. community of academic experts and practitioners.

3. RESULTS OF STUDENTS' ESSAY WRITING USING CHATGPT TOOL

3.1. Academic writing in Vietnam uses chatGPT

The pursuit of English proficiency in Vietnam has gained significant prominence, driven by a growing demand for higher education, employment, international business, and study abroad. Proficiency in English has become a prerequisite for most job positions, with English proficiency considered a key criterion for career advancement (Nunan, 2003). Because English plays a pivotal role in both one's personal life and professional career, more and more individuals are enrolling in courses designed to prepare them for English proficiency exams. Internationally recognized. Writing proficiency, which includes the standard quartet of language skills, remains the preferred measure of English proficiency, especially in academic contexts (White, 1987). However, recent test results of the International English Language Testing System (IELTS a) 2022 have shed light on a significant challenge in writing skills for Vietnamese candidates, marking this as The second most problematic skill after speaking skills. That year, the average score of Vietnamese candidates' writing test was only 6.0, ranking second lowest among the four test sections (IELTS, 2022 a). This concern is evidenced in various studies, such as Phuong's (2021) observation that Vietnamese students often encounter challenges in the area of academic writing in English, which can hinder their writing ability and academic achievement. Among the factors that contribute to this challenge are cultural and linguistic differences between Vietnamese and English, which affect aspects such as grammar, text structure, and coherence in writing. of students (Nguyen & Nguyen, 2022). The overall quality of English teaching in Vietnam, as Hoang (2010) points out, and the specific complexities involved in teaching EFL writing, as Nguyen (2009) discusses, may also be the cause. cause the difficulties students encounter. The causes of these instructional limitations, especially in the area of argumentative essay writing, require a comprehensive examination of current pedagogical approaches to English writing, the roles of teachers and learners in EFL writing classes, and the impact of Vietnamese culture, testing, and assessment. on students' EFL writing proficiency.

It is well known that academic writing requires students to demonstrate proficiency in various aspects, such as organization, grammatical accuracy, vocabulary, and critical thinking. However, language barriers, limited access to academic writing genres, and inadequate pedagogical support can hinder students' progress in this area. Identifying effective strategies to guide Vietnamese college students in academic writing in English is essential to overcoming these barriers. Recent studies have shed light on certain aspects of academic writing difficulties in English as a second language (ESL) students. Kulusakli (2021) suggests that the frequency with which students write in English in their free time is positively correlated with their enthusiasm for learning the language. Therefore, educators are encouraged to incorporate fun and engaging activities into their writing lessons to promote students' interest and motivation in learning academic writing in English. Nguyen (2021) closely examines the role of cultural influences on Vietnamese students' writing, emphasizing the need to apply a learner-centered approach tailored to address specific needs. their body. Another study conducted by Nguyen

and Nguyen (2022) delved into the impact of vocabulary knowledge on writing proficiency, emphasizing the importance of teaching vocabulary in academic writing. Furthermore, students' perception of progress in their English writing ability is significantly influenced by the guidance and feedback provided by teachers (Wahyuni, 2017). According to Wahyuni (2017), students who receive personal and constructive feedback from teachers tend to outperform students who do not receive such feedback, which reinforces the need for educators to Education must provide appropriate criticism and guidance to improve students' writing and strengthen their confidence.

Despite these contributions, the current body of literature still has certain limitations regarding a comprehensive understanding of the challenges faced by Vietnamese university students, especially regarding the impact of feedback for error identification and correction. While recent studies have shed light on specific aspects of the academic writing challenges faced by Vietnamese students, there remains a gap in understanding the relationship between the use of Artificial Intelligence-based writing and their performance in writing argumentative essays.

3.2. The method of data collection

Participants in this study will include approximately 50-60 student volunteers who will be enrolled in an academic writing course focusing on developing skills in writing academic argumentative essays, especially in the writing style. how to IELTS Academic Writing Task 2. All participants will voluntarily enroll in an intensive academic writing course course, the main objective of which is to equip them with the necessary skills to write argumentative essays academically. The course is held three times a week and each session is designed to last 1.5 hours. The entire course lasts for two months. These students were selected to participate in the study to evaluate the impact of corrective feedback from ChatGPT on their academic writing ability. Participants were individuals who were willing to improve their writing skills in an academic context and had agreed to undergo pre-and post-test assessments as part of the study. They come from diverse educational backgrounds and language proficiency levels, all united by a common interest in improving their academic writing skills. These participants, who will apply both quantitative and qualitative data collection methods, will contribute valuable insights to research investigating the effectiveness of ChatGPT's corrective feedback in improving academic writing, emphasizing coherence and cohesion.

The number of students in this study was calculated using the a-priori sample size calculator for student t-tests (Soper, 2023). It should be noted that the type of statistical test is the paired samples t-test. The predicted effect size (Cohen's d) was 0.5; the desired statistical power level (1 - β) is equal to 0.8; and the level of significance (α) is 0.05. The number of participants is very close to the number of requests given by Abramowitz and Stegun (2013) and Cohen (1988).

In terms of demographic characteristics, participants fell within a relatively narrow age range, specifically between 18 and 22 years old. Furthermore, the sample included both sexes, ensuring gender diversity. These students are currently studying at FPT University, representing many regions in Vietnam, thereby promoting regional diversity. Regarding English proficiency,

participants in this study had an English skill level ranked B1. This participant description ensures the inclusion of appropriate demographic information necessary to understand the composition of the sample. It includes factors such as group distribution, age range, gender distribution, organizational affiliation, geographical diversity, and English proficiency, thereby providing a comprehensive overview of the profile of research participants.

3.3. Survey results

In the upcoming mock test, all participants will be provided with a writing prompt taken from the book "IELTS Academic 15" (2020) and will be instructed to write a minimum essay of 250 words within the time frame of 40 minutes and answer the given question. During this exam, candidates will not be allowed access to supplementary materials or allowed to engage in any form of collaborative discussion. Then, two researchers, each with experience in IELTS training and a certificate from the "IELTS Teacher Training Program 2020" issued by IDP Australia, will independently evaluate the essays. Assessment will be carried out according to the IELTS Task 2 writing assessment scale, as described in Appendix 6, whereby a separate score (from 0 to 9.0) will be assigned to each of the four criteria Evaluate. The total score will then be determined, rounded to the nearest whole number or 0.5. After this initial scoring, the two reviewers will convene to review all ratings together, engaging in thorough discussions to reconcile any discrepancies until a consensus is reached. final consensus.

In the upcoming post-test, participants will be tasked with creating an essay consisting of a minimum of 250 words, to be completed within a 40-minute time frame. Essay writing advice will be excerpted from the book "IELTS Academic 17" (2022). Although the thematic basis of this prompt will reflect that of the previous test, the actual content of the question will be different. This intentional difference in question content is intended to ensure that the differences between the two test cases remain moderate, neither too different nor too similar. The evaluation of post-test essays will follow the same procedures as those applied during the pre-test. Initially, individual scoring will be conducted, followed by collaborative deliberation between the two raters to reach a final consensus.

This test is designed to assess students' initial academic writing skills, especially their ability to recognize errors and their level of cohesion and coherence in argumentative essays. Pre-test scoring criteria are aligned to IELTS Academic Writing Task 2 standards, ensuring that the assessment is based on the same criteria used by experienced IELTS instructors. Three experienced IELTS instructors independently evaluate and score pre-essays to maintain objectivity and reliability during the data collection process. Scored pre-test data, including students' academic writing scores related to error identification, error correction, cohesion, and coherence, are meticulously recorded and stored in a database safe. This organized data serves as a baseline measure of students' writing skills before corrective feedback intervention.

After completing the pre-test, students will go through a corrective feedback process, including using ChatGPT to receive guidance on how to recognize and correct errors, as well as to enhance engagement and Coherence in academic writing. Treatment is performed three times a week, with each session lasting 30 minutes to 1 hour, for a total of 10 weeks. After 10 weeks of intervention, students will take a post-test, reflecting the pre-test assessment. The post-

test assesses students' academic writing skills in terms of their ability to recognize errors, correct errors, cohesion, and coherence, again based on IELTS Academic Writing Task 2 criteria.

The scores obtained from the posttest, which reflect the student's academic writing progress, are systematically recorded and stored along with the pretest scores. This storage ensures that post-test data is always available for analysis and comparison. The paired data sets of pre-and post-test scores allowed for subsequent quantitative analysis to assess whether there were statistically significant differences in students' academic writing performance before and after treatment. corrective feedback or not. This analysis contributes to answering basic research questions regarding the effectiveness of feedback interventions.

By following this meticulous process of collecting quantitative data, the study aimed to provide empirical evidence of the impact of corrective feedback on students' academic writing skills. Systematic data collection and future analysis will help address the research objectives, providing valuable insights into the potential benefits of ChatGPT as an educational tool. In the next section, the author will continue to present how to apply paired sample t-tests to test specific research hypotheses. These hypotheses focus on determining whether using corrective feedback from ChatGPT leads to significant improvements in student's ability to identify and correct errors, as well as increased engagement. and coherent in their argumentative essays.

The application of a paired sample t-test will facilitate the testing of specific research hypotheses. Hypotheses will be developed to determine whether the use of corrective feedback from ChatGPT has a statistically significant impact on student's academic writing skills, specifically on their ability to recognize and Correct errors as well as improve cohesion and coherence in argumentative essays. The results of the t-test will be used to evaluate whether these hypotheses can be accepted or rejected. The final stage of quantitative data analysis will focus on interpreting the findings. The results of the paired samples t-test will be examined to determine whether the impact of corrective feedback from ChatGPT on students' academic writing skills is statistically significant. The practical implications of the statistical results will be discussed in the context of the research objectives. This explanation will contribute to the overall understanding of the effects of corrective feedback interventions in quantitative data.

Through the systematic application of these quantitative data analysis steps, the study will be able to provide empirical evidence on the impact of corrective feedback from ChatGPT on students' academic writing skills, especially focusing on identifying and correcting errors, as well as improving the cohesion and coherence of your argumentative essays.

With an initial set of codes, the analysis proceeds to a thematic analysis. This will involve systematically organizing the codes into overarching themes and subthemes. The researcher will engage in a process of constant comparison in which codes will be reviewed and grouped according to their conceptual relevance. Emerging themes encapsulate the core concepts of participants' stories and provide a structured framework for interpreting the data. Once the thematic analysis is complete, the researcher will proceed to interpret the findings. This phase will require a thorough exploration of the identified themes and their relevance to the research questions. Interpretations will be based on the participants' own words and experiences, emphasizing the authenticity of their perspectives. To enhance the validity of the

interpretations, member checking will be performed, allowing participants to review and confirm the accuracy of the findings presented.

The final step in qualitative data analysis will involve synthesizing the findings into a coherent story. The researcher will prepare a comprehensive report detailing the major themes and their implications. In keeping with academic conventions, qualitative results will be presented alongside quantitative results, providing a comprehensive view of the research findings. The report will also include direct quotes from participants to illustrate their voices and experiences. Through rigorous application of these qualitative data analysis steps, the study successfully captures student perspectives and experiences related to the corrective feedback provided by ChatGPT. This qualitative analysis will provide valuable insights into the qualitative aspects of the study, complement the quantitative findings, and enrich the overall understanding of the impact of corrective feedback on students' academic writing skills.

CONCLUSION

The researcher will use a mixed methods approach to address three research questions. The research intervention will involve integrating ChatGPT into Academic Argumentative Writing lessons for a small group of student volunteers at FPT University. Data collection will be primarily based on pre-post survey questionnaires, interviews, and surveys will be provided to participants during the quantitative phase and semi-structured interviews will be conducted with 10 participants in the qualitative phase. The pre-post test will aim to address research question 1, while the insights gained from the interviews will be instrumental in addressing research question 2. The survey questionnaire will play an important role in addressing research question 3. Participants who will experience the ChatGPT intervention in their lives in The Academic Writing Course will be students of the researchers. Data will be stored and analyzed meticulously using SPSS 28 and Microsoft Excel software. Furthermore, the study will strictly adhere to ethical considerations throughout its duration.

REFERENCE

AIContentfy. (2023, January 28). *Chatgpt and the future of human-computer interaction*.

Ajith, N., & Muthumani, S. (2023). Impact on users' perceived effectiveness when using ChatGPT. *International Journal of Advancing Research in Engineering Management and Science*, *3* (9), 20-24. https://doi.org/10.58257/ijprems31972

Akgun, S., & Greenhow, C. (2021). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, *2* (3), 431-440. https://doi.org/10.1007/s43681-021-00096-7

Al-Garaady, J., & Mahyoob, M. (2023). ChatGPT's ability to detect and analyze writing errors experienced by EFL learners. *English Journal of the Arab World, Special Issue on CALL*, *9*, 3-17.

Al-khresheh, M. (2010). Bilingual intervention in Jordanian EFL learners' English word order structure. Proceedings of the European Journal of Social Sciences. *European Journal of Social Sciences*, *16*, 105-116.

Al-Mansour, N. (2015). Teaching academic writing to Saudi university students: Problems and solutions - King Saud University perspective. *Arab World English Journal*, *6* (3), 94-107. https://doi.org/10.24093/awej/vol6no3.6

., & de Winter, J. (2023). Using ChatGPT for human-computer interaction research: A primer. *Royal Society Open Science*, *10* (9). https://doi.org/10.1098/rsos.231053

Taj, I. H., Ali, F., Sipra, M., & Ahmad, W. (2017). The effectiveness of technology-enhanced language learning on EFL learners' vocabulary acquisition. *International Journal of Applied Linguistics & English Literature*, *6* (3), 262-272.

Tambunan, A.R., Andayani, W., Sari, W.S., & Lubis, F.K. (2022). Investigating EFL students' language problems using grammar as automated writing assessment feedback. *Indonesian Journal of Applied Linguistics*, *12* (1), 16-27. https://doi.org/10.17509/ijal.v12i1.46428

Tesfaye, D., & Tsadik, D.G. (2015). Analysis of errors in essays written by graduates concerning teacher training colleges in the Oromia region: A mixed approach. *International Journal of Educational Research and Technology*, *6* (3), 27-40. https://doi.org/ 10.15515/ijert.0976-4089.6.3.2740

Thomas, L. (2023, June 22). *Quasi-experimental designs: Definitions, types, and examples*. Scribbler. https://www.scribbr.com/methodology/quasi-experimental-design/

Thorat, S.A., & Jadhav, V. (2020). Evaluating implementation issues of rule-based chatbot systems. *SSRN electronic magazine*. https://doi.org/10.2139/ssrn.3567047

Thorp, H.H. (2023). FPTPT is fun but not the author. *Science*, *379* (6630), 313-313. https://doi.org/10.1126/science.adg7879

Touchie, H.Y. (1986). Types of errors in learning a second language, causes and solutions. *JALT Journal* , *8* (1).

Trinh, LQ, & Nguyen, TT (2014). Improve the ability to write argumentative essays of Vietnamese learners. *TEF Asian Journal*, *11* (2), 63-91.

Uysal, H.H. (2009). A critical review of the IELTS writing test. *ELT Journal*, *64* (3), 314-320. https://doi.org/10.1093/elt/ccp026

Victor, W. (1973). Analyze mistakes when writing. *ELT Journal*, *27* (2), 177-186. https://doi.org/10.1093/elt/xxvii.2.177

Vygotsky, LS (2012). *Thinking and language*. MIT Press.

Vygotsky, L.S., Cole, M., John-Steiner, V., Scribner, S., & Souberman, E. (1978). *Mind in society: Developing higher psychological processes*. Harvard University Press.

Vygotsky, Lev Semeonovič. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Waddell, T. F., Zhang, B., & Sundar, S. S. (2015). Human-computer interaction. *International Encyclopedia of Interpersonal Communication*, 1-9. https://doi.org/10.1002/9781118540190.wbeic182

Wagner, G., Lukyanenko, R., & Paré, G. (2021). Artificial intelligence and conducting literature reviews. *Journal of Information Technology*, *37* (2), 209-22