rEFLections
Vol 32, No 1, January - April 2025

Improving Academic Writing Proficiency for EFL Students: Leveraging ChatGPT Using Data-Driven Learning Principles

EDWARD DEVERE BACON*
Faculty of Liberal Arts, Thammasat University, Thailand
HARALD KRAUS
Faculty of Liberal Arts, Thammasat University, Thailand
Corresponding author email: superduperreadyeddie@rsu.ac.th

Article information
Article history:
Received: 14 Jun 2024
Accepted: 21 Apr 2025
Available online: 25 Apr 2025

Keywords:
Data-driven strategies
Artificial intelligence
ChatGPT
International students
Academic writing skills

Abstract
This research examines a learner-centered approach to using ChatGPT based on data driven learning principles in essay writing instruction. A quasi-experimental study was adopted with three groups comprising 92 international students enrolled in a pre-sessional foundation writing course at a Thai university. The control group followed the usual textbook-centric curriculum. Experimental Group 1 was taught to compare their own writing with paper-based sample IELTS essays, and Experimental Group 2 used ChatGPT (GPT 3.5) generated paraphrases of their own compositions and guided worksheets for students to compare and analyze. Within group and between group posttest analysis of student compositions found that the participants that used ChatGPT significantly outperformed the other groups (control group posttest writing score $\bar{x}$ = 74.03%; Experimental Group 1 writing posttest $\bar{x}$ = 73.68%; Experimental Group 2 writing posttest score $\bar{x}$ = 94.3% 0.001 at 0.05). A follow-up questionnaire and interview revealed that the participants in Experimental Group 2 appreciated using AI in this way and developed an increased level of confidence. The study concludes that the adapted use of AI powered chatbots is effective for developing short essay writing skills.

INTRODUCTION

The acquisition of academic essay writing skills is crucial for students navigating the challenges of English as a Foreign Language (EFL) in international university programs. While many institutions offer pre-sessional programs, these are often insufficient for many learners to meet required entrance standards, and international students can often face significant challenges (Campbell & Li, 2008; Phakiti & Li, 2011). These challenges often originate from limited exposure to natural language use and lack of practice in analyzing how (academic) language works in everyday situations. As a result, language teachers must balance helping students use the language correctly with guiding them to express their ideas effectively. Additionally, large class sizes and limited time can make it difficult for teachers to give personalized, in-depth feedback that highlights areas where students need to improve.

To address this, two primary aspects of language learning become especially relevant: attention to and awareness of target language and discourse features (Robinson et al., 2019). Facilitating these aspects can be challenging, but one promising approach is data-driven learning (DDL), which involves the 'direct use of corpora by students' (Godwin-Jones, 2021, p. 5) to identify and compare features of language. Indeed, proponents of DDL have an 'undying enthusiasm' (O'Keeffe & Mark, 2022, p. 260) regarding its ability to draw attention to authentic language use, primarily in terms of lexicon and lexicogrammar.

However, there is general acknowledgment that DDL-based instruction has numerous challenges, many of them related to the usual requirement to use concordancing software (Godwin-Jones, 2021). The advent of AI-powered language learning models (LLMs), however, has made the use and analysis of "big data"—in essence, corpora—more accessible and user-friendly. In this paper, we argue that it is possible to use LLM chatbots, in this case specifically ChatGPT, to adapt core principles of data-driven learning and apply them to essay writing instruction in a way that allows learners to independently solicit language data and critically reflect on AI-generated feedback.

This study contributes to the growing body of research in several ways. First, it proposes a student-centered model of ChatGPT integration that aligns with data-driven learning pedagogy. Second, the particular approach examined here—the use of ChatGPT's paraphrasing of learners' own compositions—not only offers personalized feedback, but also supports the development of metalinguistic awareness by prompting students to reflect on alternative lexical choices and syntactic structures. This reflective comparison can encourage and facilitate vocabulary expansion, syntactic restructuring, and a deeper understanding of discourse conventions—cognitive processes central to L2 writing development (Cheng et al., 2025; Khojasteh et al., 2025; Storch, 2005). Third, because ChatGPT is used to paraphrase rather than generate entirely new content, this approach directly addresses ethical concerns regarding the responsible use of AI-powered tools in academic writing instruction.

Three research questions are investigated:

1. Are there significant differences in learner essay compositions between and within groups using ChatGPT, paper-based learning, and the control group?
2. What are the perceptions and attitudes of EFL students toward using ChatGPT as a learning tool compared to student experiences with paper-based learning?
3. What are the challenges faced by learners when using ChatGPT for English writing?

LITERATURE REVIEW

Data-driven learning

Educators and linguists have used corpus analysis to gain insights into language structure, patterns, and use to help them design language courses, identify common learner errors, and create resources that address specific language needs (e.g., Boulton & Cobb, 2017; Boulton &

Pérez-Paredes, 2014; Chang, 2014; Hadley, 2002; O'Donnell, 2013; Vyatkina, 2016). Corpus analysis has been extensively used for purposes such as textbook writing and dictionary compilation (Godwin-Jones, 2021). However, scholars point out that there is a gap between research and classroom practice (Chambers, 2019, Reinders & Lan, 2021), specifically in terms of the learner's use of corpora, and continue to call for wider adoption of corpus-based instruction (e.g., Bennett, 2010; Flowerdew, 2012; Liu, 2023).

An extension of corpus-based instruction is data-driven learning. When DDL is applied in the classroom, learners are typically taught how to search for and identify language patterns and uses such as collocations found in selected corpora. Usually, this is conducted using concordancing software, which automates the procedure by presenting concordance lines, word frequencies and other statistical analyses. For example, in one study, learners examining concordance lines produced fewer errors and used more complex sentence structures compared to control groups relying on traditional dictionaries (Yılmaz, 2017). Additionally, studies by Mao et al. (2018), Sepehri, (2015), and Zhu (2021) have emphasized the positive impact of data-driven learning on students' effective vocabulary usage in academic essays. DDL has also facilitated awareness raising, where students have conducted self-assessment of their compositions, helped them identify their strengths and areas needing improvement, and aided in error-correction (Chambers & O'Sullivan, 2004; O'Sullivan & Chambers, 2006).

Corpus analysis/DDL aligns with constructivist learning theories (O'Keeffe & Mark, 2022) which emphasize activities in which learners are scaffolded via modelling, are provided with and process feedback and break down tasks to reduce cognitive load (Kaufman, 2004). Positioning the teacher as facilitator, social constructivist approaches aim to create conditions that foster inductive learning, and the sharing of thoughts, perspectives, and experiences and 'hypothesis testing', which are seen as crucial for achieving success in learning (Rowe, 2006). Research has found that students who engage in writing exercises with feedback take greater control over their work (Cotterall & Cohen, 2003), and that effective feedback is crucial for a deeper comprehension of the writing endeavor (Storch, 2005). Abasi et al. (2006) and McKinley (2015) further demonstrate that receiving feedback enhances learners' capacity to formulate social and academic discourses in their writing.

Despite these promising applications of DDL, many educators have been deterred from exploring data-driven learning for a number of reasons: the software can be intimidating; users need to become familiar with technical terminology and statistics; users still need to interpret the results of searches, and a relatively high level of language proficiency is necessary for effective corpus consultation and training (Chang, 2014; Yoon & Hirvela, 2004). Students therefore often require the instructor's help to fully utilize the tool (Götz, 2012), undermining its potential for autonomous use (Godwin-Jones, 2021). For example, although learners in Liu's (2023) improved their reading skills and strategies after compiling a corpus of self-selected texts, the researcher noted the need for careful planning and the teacher's central role in facilitating the use of corpora, thus raising the question as to whether DDL based approaches are viable in day-to-day instruction.

Moreover, while in principle DDL could be used in classes of advanced and dedicated language learners, DDL may not be appropriate for lower-level learners whose primary educational

focus is not related to language (e.g., pre-sessional students aiming to enroll in other majors) (Götz, 2012). Likewise, peer or teacher feedback is not always available or necessarily formulated in ways that benefit the learner. Thus, alternative means to acquire feedback and develop an understanding of various aspects of writing may be needed. Finally, another obstacle in adopting concordancing software in EFL is that current versions are primarily tailored for desktop computers and lack support for mobile devices. Consequently, students without computers are unable to benefit from the prevailing data-driven learning platforms (Boulton, 2013). With the introduction of AI however, the principles of DDL can be applied in more user-friendly, 'social' and familiar ways.

Of particular interest in this study then, is to adopt a 'corpus consultation' approach (Chambers & O'Sullivan, 2004) by taking the principles of data driven learning—textual comparison and analysis, feedback, self-assessment, scaffolded and inductive learning—by using LLM-generated corpora. By doing this, it is expected that learners can more independently use the tool to draw attention and awareness to linguistic challenges, encourage problem solving and discovery, and act as a scaffolding mechanism.

ChatGPT

Although chatbots and Artificial Intelligence (AI) have been in use and development for decades, powerful (and often free) LLM chatbots such as ChatGPT, released in November 2022, have only recently enabled language analysis without requiring the use of specialized software. By inputting simple prompts in familiar-looking web-page search bars, LLM chatbots not only provide the requested information, but do so in an interactive and conversational manner, generating human-like text and engaging in ongoing conversations (Dergaa et al., 2023). LLMs have a multitude of uses as a resource for teachers, from lesson planning to generating quizzes, to assessing (Kohnke et al., 2023) and can help learners practice language skills (e.g., Fryer et al., 2019; Kim et al., 2021) provide detailed explanations of linguistic concepts, and assist in generating information for research or assignments (Kostka & Toncelli, 2023). Using ChatGPT for language analysis, learners can, for example, pose questions regarding task achievement, lexical choice, grammatical accuracy, and coherence and cohesion in their academic writing. While established AI based services such as Grammarly.com have long offered this kind of information, they have essentially been oriented to identifying (probable) learner errors and stylistic weaknesses. Newer LLM platforms are able not only to offer this form of feedback, but also foster a more personalized, adaptive, interactive, and tailored learning experience resulting in enhanced motivation (Ayoubi, 2024; Chiu et al., 2023; Fariani et al., 2023; Kohnke et al., 2023).

ChatGPT and (EFL) writing instruction

The use of AI powered chatbots has unsurprisingly attracted attention in foreign language pedagogy and research endeavors are rapidly growing. To date however, much of the research has tended to focus on the development of conversational skills (Belda-Medina & Calvo-Ferrer, 2022; Kim et al., 2021; Lin & Mubarok, 2021), whereas fewer studies have investigated essay writing development in English, with most evidence being anecdotal (Han et al., 2023). But

there have been some interesting recent findings. Özçelik and Ekşi (2024) investigated the use of AI as a tool for acquiring register knowledge. However, this was a small-scale pre-experimental study, and ChatGPT was used only to provide feedback, without requiring the students to compare and reflect on it. Banihashem et al. (2024) compared ChatGPT feedback with peer feedback, concluding that ChatGPT provided more descriptive feedback, whereas peers provided higher quality feedback in terms of areas needing improvement. However, the study was conducted with Dutch graduate students, and not in a foreign language learning context. Finally, in Han et al. (2023) learners consulted ChatGPT for specific language items in their essays, and while participants responded positively, they too were more advanced learners.

Overall, studies have focused on and exploited the interactive and personalized uses of LLMs such as ChatGPT, as well as their ability to foster critical thinking. In a review of 25 empirical studies, Huang et al. (2022) concluded that chatbots offer three affordances: timeliness, ease of use, and personalization, and five pedagogical uses, including helping, transmitting information, recommending, simulating, and serving as interlocutors, which are seen to encourage social behavior and interaction. ChatGPT is not inherently a specialized corpus analysis tool, but it can function in such a manner, providing the features that concordancing software cannot offer—personalized feedback by way clarification, responses to linguistic inquiries, and dialogic engagement which facilitates an engaging, user-friendly approach to corpus analysis.

Challenges using AI chatbots

Given their recent and rapid introduction, LLMs have left many educators in some ways scrambling to catch up. For example, despite increased digital literacy acquired during the COVID19 pandemic, instructors have reported a lack of confidence in incorporating AI in their classes (Kohnke et al., 2023) and it appears they have primarily used ChatGPT (and the like) as a tool for conducting searches: in one study, 44.3% of instructors reported information searching as a primary use, while paraphrasing and content generation was done by 39.4% and 27.2% respectively (Yusuf et al., 2024).

With regard to learners, much of the conversation on AI assisted learning has centered on its potentially damaging impact in education, with concern raised over students' academic (dis) honesty (Hamilton, 2023), learners' possible overreliance on AI, and relinquishing their responsibility to think critically (Seo et al., 2021; Yusuf et al., 2024). Similarly, dependence on AI for feedback could potentially undermine the development of autonomous learning habits, as students might prioritize AI-generated suggestions over their own judgment or peer input. There are also concerns regarding the depth and nuance of AI feedback, particularly in complex areas of writing that require human insight and contextual understanding. While AI can provide feedback on grammar and structure, it may not fully grasp the subtleties of creative expression or the specific requirements of different academic disciplines (Banihashem et al., 2024). Thus, both educators and learners will quickly need to leverage AI in productive, ethical and principled ways.

One way to address many of these issues in writing instruction is to prompt LLM chatbots to paraphrase students' own compositions, rather than prompting them to generate original text.

These corpora of paraphrases can be used by learners to identify how AI 'improves' or fixes their writing holistically, i.e., at the level of grammar, word choice, organization and discourse.

METHODOLOGY

Participants

The participants in this study were students aged 18-22, of equal gender distribution and predominantly from China, enrolled in a foundational English writing course at a Thai private university (further details are shown in Table 1). They had passed entry requirements in an in-house entry exam by achieving a proficiency level equivalent to at least the A2 level in the Common European Framework of Reference (CEFR). The course's objectives were to prepare students for enrollment in one of the international English language programs offered by the same university, the entry requirements for which was satisfactory results (a score of 5.5) in a test modelled on IELTS. The study consisted of 92 participants distributed across three intact English writing classes. The three groups were randomly selected to become one control group (n = 31) and two experimental groups (n = 31 and n = 30 for EG 1 and 2, respectively). This is consistent with accepted group sizes for experimental research (Fraenkel et al., 1993). Students in both experimental groups agreed to participate by signing a consent form.

Table 1

Demographic information of participants

Demographic Information	Category	Frequency (n)	Percentage (%)
Age	18-20	92	100%
Gender	Male	45	48.9%
	Female	47	51.1%
Nationality	Chinese	91	98.9%
	Burmese	1	1.1%

Procedures

Academic writing course

The writing course comprised ten modules over fifteen weeks, and followed a syllabus designed by the instructors. Topics covered included thesis statements, topic sentences, giving evidence and examples, achieving coherence, synthesis of the main points and concluding, among others.

During the study all three groups were exposed to an identical 3-hour class per week, consisting of approximately 1.5 hours of teacher-led instruction, followed by one hour of writing an in-class paragraph or short essay (approximately 250 words) written on paper. Following this, lesson procedure diverged depending on group designation, as shown in Figure 1.

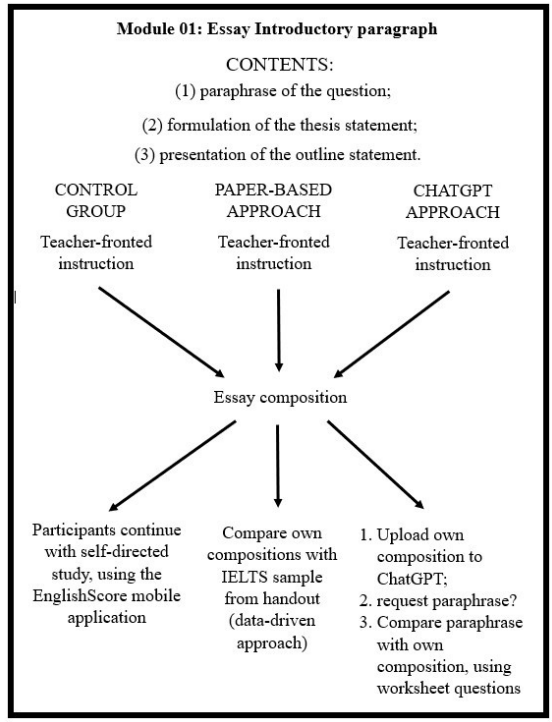

Figure 1 Sample of procedures for module 01

After completing the writing task, students in the control group were asked to spend the rest of the class using the EnglishScore mobile app, which was developed by the British Council. This approach was taken to ensure that all groups had equal time-on-task. EnglishScore was selected because it provides standardized practice in general English skills, such as grammar and vocabulary, which support the overall objectives of the course but are not directly related to academic writing. The aim was to offer a meaningful follow-up activity that would keep students engaged without influencing the specific writing outcomes being studied. The app was chosen for its accessibility, relevance for EFL learners, and its ability to maintain student interest without overlapping with the targeted writing instruction given to the experimental groups.

Although the focus of this research is primarily the use of AI-powered technology, the authors were aware of important provisos and cautions. First, it should not be assumed that all students have access to the technology required (mobile devices and/or internet access), or to use online chatbots, as was the case with some of the participants in this study. We have also noted concerns of the use of chatbots to entirely compose texts for learners. Therefore, the researchers wished to investigate if an analogue, paper-based (i.e., involving only printed materials) version of the corpus consultation approach to learning writing would yield similar results.

The first experimental group followed a paper-based treatment. Following their writing composition, participants in Experimental Group 1 (EG 1) analyzed and compared IELTS

sample essays provided by the instructor. Participants were given a handout of questions to guide them. However, whereas traditional data-driven learning is typically more focused on lexical level analysis, the questions in the handout focused on discourse level features, such as introductory paragraph, body paragraph and concluding paragraph construction, as expected in IELTS-based opinion essays. For example, after reading a sample, participants were asked to consider if the introductory paragraph was an effective one, and if so, why they believed this.

At the beginning of the course, Experimental Group 2 (EG 2) was introduced to ChatGPT through a structured instructional session aimed at familiarizing learners with the platform's capabilities and appropriate uses. The session emphasized that ChatGPT would not be used by learners to compose or simply correct their writing, nor to evaluate or score it, but to paraphrase their original texts in order to encourage comparison and critical reflection. Students were guided on how to input their own writing into the ChatGPT interface and how to craft prompts requesting paraphrased versions of their texts. The instructor further demonstrated how students could refine their queries to elicit more focused or detailed responses. This practice was positioned as a tool for comparison, enabling students to analyze alternative lexical and syntactic choices and reflect on the effectiveness of their original phrasing. To support ethical use and enhance learner autonomy, ChatGPT was framed as a supplementary aid—complementary to traditional tools such as textbooks, grammar references, and writing guides—rather than as a replacement for teacher feedback or peer review. This approach was intended not only to simulate familiar academic support practices but also to minimize potential resistance to integrating AI in their learning process.

Following the introductory session, participants were directed to use ChatGPT after completing each day's classwork task. The instructor provided a link to a worksheet with questions designed to guide them to compare their work with the ChatGPT output. The questions aligned with the IELTS essay writing rubric in the categories of coherence and cohesion, lexical resource, and grammatical range (Pospelova, 2021). For example, for the lesson on introductory paragraphs, the students were given the task below:

1. After copying and pasting your introductory paragraph in ChatGPT (https://chat.openai.com/), list words ChatGPT's THESIS STATEMENT has that your original THESIS STATEMENT does not have. Are the words that ChatGPT used a good decision? Why do you think so?

Further questions explored grammar changes, use of cohesive devices and thesis statements and so on, totaling 12 questions for the introductory paragraph of an essay.

DATA ANALYSIS

Pre- and post-tests

Writing assessment was focused on a rubric resembling the one employed for the IELTS Writing Task 2: task completion, vocabulary, grammar, and coherence and cohesion. Three native

speaker assessors with at least 10 years university teaching experience each were recruited. Scoring ranged from 0 to 100 to allow for greater variation than the IELTS 1-9 band range, and for easier conversion for statistical analysis. Scoring was blinded, i.e., no participant names or identifying information appeared on the essays, and the essays from the groups were mixed.

A writing posttest was administered to all participants two days after the final class session. The pretest data were subjected to an Analysis of Variance (ANOVA) and posttest data were subjected to an Analysis of Covariance (ANCOVA), which allowed for an examination of intra-group and inter-group variances. As the ANOVA and ANCOVA results indicated significant differences among group means, a post-hoc Tukey's Honestly Significant Difference (HSD) test was conducted for pairwise comparison offering a detailed analysis of which specific groups differed from each other.

Questionnaires

Participants' responses from both experimental groups were collected through online questionnaires designed by the researchers. The questionnaire for EG 1 comprised 3 parts totaling 25 statements and the questionnaire for EG 2 (ChatGPT) comprised 39 questions also separated into 3 parts. Both questionnaires were comprised of Likert scale response anchors (1 = Strongly disagree, 2 = Agree, 3 = Not sure, 4 = Agree and 5 = Strongly agree). To ensure validity and reliability both questionnaires were IOC validated by three assessors, all of whom had over ten years of experience in teaching at universities. Despite initial concerns about the questions appearing biased towards positive aspects, the experts confirmed that the questions were balanced and designed to elicit a range of responses, both positive and negative, from participants. The IOC analysis achieved a score of 1.0 indicating a high level of congruence and were subsequently piloted with 39 undergraduate students. The Cronbach's Alpha for EG 1 questionnaire was .949 and the Cronbach's Alpha for EG 2 questionnaire was .953 suggesting a high level of internal reliability.

Semi-structured interviews

Following the treatments, a random sampling of fifteen individuals from EG 1 and EG 2 were selected for a semi-structured interview in English, with consideration given to gender representation and a diverse range of English language proficiency levels. Each interview, based on 10 initial questions, spanned between 10 to 15 minutes. Interview sessions were video recorded, transcribed and analyzed.

RESULT

Pretest and posttest

Table 2 displays the one-way Analysis of Variance (ANOVA), which was used to indicate if there were statistically significant differences between the means of three groups (Ostertagová & Ostertag, 2013). The results for the pretests among all three groups show a p-value under

0.05, indicating that two of the groups were statistically significant from another. To find out which groups, a Tukey HSD was run, the results of which are shown in Table 3.

Table 2

ANOVA test table for the pretest among all groups

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F-Statistic	p-value
Between Groups	19794.546	2	9897.273	36.581	0.000
Within Groups	24079.758	89	270.559		
Total	43874.304	91			

Table 3

Tukey HSD post hoc test for the pretest among all groups

Group 1	Group 2	Mean Difference	Std. Error	p-value
Control group	Experimental group 1	-0.0323	4.18	0.9
Control group	Experimental group 2	-31.3065	4.21	.001
Experimental group 1	Experimental group 2	-31.2742	4.21	.001

As indicated in Table 2, there were no significant differences between the control group and EG 1 (p = 0.9). However, significant differences were noted between the control group and EG 2, (p = 0.001 at 0.05 level) and for EG 1 and EG 2 (p = 0.001 at 0.05 level). The results suggest that participants in EG 2 had lower writing skills when compared to the other two groups before the treatment was implemented, given the negative values in the mean differences. This was attributed to the fact that many students in this group reported having had very little writing experience.

The study utilized Analysis of Covariance (ANCOVA) to analyze post-test scores among the Control Group, Experimental Group 1 (EG 1), and Experimental Group 2 (EG 2). This choice was based on the significant differences observed in the pretest scores among the groups, particularly the notably lower pretest scores in EG 2 compared to the other groups. ANCOVA is justified in this context as it allows for the control of pre-existing differences on the outcome variable, thus providing a clearer picture of the impact of the different teaching methods (Tabachnick & Fidell, 2013). By including pretest scores as a covariate, we aimed to account for these initial differences in writing skills and isolate the effect of the teaching methods on post-test performance. This approach ensures that the observed differences in post-test scores are more accurately attributed to the teaching interventions rather than pre-existing disparities in student performance.

Table 4

ANCOVA posttest analysis

Variable	df	Sum of Squares (SS)	Mean Square (MS)	F-Statistic	p-value
Teaching method	2	7537.20	3768.60	31.58	< 0.00001
Pretest Scores	1	775.18	775.18	6.50	0.01254

The ANCOVA results indicated a significant effect of teaching method on post-test scores, $F(2, 88) = 31.58$, $p < 0.00001$. This finding suggests that the teaching method significantly influenced student learning outcomes. Moreover, pre-test scores were also found to significantly contribute to the variance in post-test scores, $F(1, 88) = 6.50$, $p = 0.01254$, underscoring the importance of accounting for initial student performance levels in educational research. The statistically significant p-value associated with the teaching method indicates a strong effect on post-test scores.

The analysis revealed that while there was no significant difference between the control group and EG 1, both groups showed significant differences in posttest scores when compared to EG 2. This suggests that the treatment applied to EG 2 had a measurable impact on their English posttest performance compared to the other groups.

Cohen's d test results

In order to quantify effect size among the three groups, the researchers conducted a Cohen's d test, which provides a measurement of the magnitude of the difference between two groups or conditions, expressed in standard deviation units. This test helps to better fully understand the practical significance of their findings, beyond mere statistical significance (Cohen, 1988). The results of the Cohen's d test can be seen below in Table 5.

Table 5

Cohen's d test results

	Group	Cohen's d
1	Control Group	0.149
2	EG 1	0.121
3	EG 2	2.561

The analysis of Cohen's d for the three groups revealed varying levels of effectiveness in improving test scores from pretest to posttest. The control group exhibited a small effect size (Cohen's d = 0.15), indicating a modest improvement in scores. Similarly, EG 1 also showed a small effect size (Cohen's d = 0.12), slightly less effective than the control group. In stark contrast, EG 2 demonstrated a large effect size (Cohen's d = 2.56), suggesting a substantial and significant improvement in test scores. These results suggest that the intervention applied in EG 2 was markedly more effective compared to the other groups, which showed only minor improvements.

Questionnaires

Answers to the questionnaire were computed using descriptive analysis.

Experimental Group 1

The responses to the questionnaire, which was made up of three sections: (1) attitudes towards paper-based data-driven learning ($\bar{x} = 4$, SD = 0.91); (2) attitudes towards paper-based data

sources ($\bar{x}$ = 3.5, SD = 1.09); and (3) confidence and attitudes towards outcomes ($\bar{x}$ = 4.5, SD = 1.07), indicated a high and a moderate satisfaction level. All participants (n = 31) in EG 1 completed the questionnaire. The full results are shown in Appendix 1.

The results revealed that 76% of the answers to the questionnaire were classified as 'high' levels of satisfaction, whereas 20% of the answers to the questionnaire were classified as moderate levels of satisfaction and 4% of the answers to the questionnaire were classified as highest levels of satisfaction. One of the lowest mean scores was rated for Item 10, "The quality of data sources in paper-based materials for academic writing is consistently reliable." having a mean of 3, whereas the highest answer was rated for Item 11, "I feel a strong sense of accomplishment when I successfully incorporate data from paper-based sources into my academic papers." with a mean of 5.

Experimental Group 2

As with EG 1, a post-treatment questionnaire was distributed online for students in Experimental Group 2 (n = 30). The 39 responses to each item were analysed separately using descriptive analysis (percentages, means and standard deviation). The scores indicated a high level of satisfaction of the participants, as shown in Appendix 2. Results indicated that the participants' attitude toward the 10-week program was at a high level of satisfaction with a mean of 3.99 and SD of 1.14. The lowest mean score was rated in response to the statement "You are confident in your ability to use artificial intelligence tools to improve your writing skills" with a mean of 3.50. All other questions from the questionnaire had a mean of 4, indicating high within-group agreement.

EG 2's questionnaire was separated into the following three key areas: data-driven approaches, AI and ChatGPT. In terms of data-driven approaches, both groups reported a mean score of 4 for improved writing skills and motivation, with EG 2 showing slightly more consistent satisfaction (SD = 1.03 compared to EG 1's 1.06). For AI, which was only assessed in EG 2, the mean scores were 4 for improved writing skills, motivation, identifying errors, and future use, with a slightly lower score for confidence in using AI ($\bar{x}$ = 3.5). ChatGPT, also assessed only in EG 2, received high mean scores of 4 across all evaluated dimensions, including improved writing skills, generating ideas, ease of use, identifying areas for improvement, confidence and motivation, and personalized feedback. The findings indicate that EG 2, utilizing AI and ChatGPT, experienced better overall satisfaction and specific benefits in their writing pedagogy compared to the traditional paper-based methods used by EG 1, highlighting the enhanced engagement and effectiveness of integrating advanced AI tools in EFL academic writing instruction.

Interview results

A total of 30 students, aged between 18 and 20, volunteered for semi-structured interviews. Their academic backgrounds included Communication Arts (50%), Design (30%), International Business (17%), and Information Communication Technology (3%). The interviews were carried out in English and later analyzed to uncover common themes reflecting the students' experiences in both experimental settings.

Perceived benefits of ChatGPT use

Several students from EG 2 shared that they found ChatGPT useful for building vocabulary, improving grammar, and making their writing clearer. Many appreciated the instant and easy-to-understand feedback it offered. One student remarked, "ChatGPT can help us to improve English writing better than before", while another highlighted how real-time corrections helped them learn new phrases and sentence structures: "It can help us master more phrases and sentences." Others mentioned that the tool made it easier to organize their ideas and write more smoothly. As one participant put it, they used ChatGPT to "change our writings' words and sentence (to become) better than before."

Learner engagement and motivation

Increased motivation and engagement were frequently reported. Students appreciated the accessibility and convenience of the tool, with one stating, "Using ChatGPT improved learners' participation and enthusiasm for studying English." The platform's user-friendly and interactive format appeared to foster a more engaging learning environment compared to traditional approaches.

Challenges and limitations

Despite the benefits, participants identified several challenges in using ChatGPT effectively. Some struggled to understand the advanced vocabulary or formal expressions suggested by the tool. One participant remarked, "Sometimes I don't understand the sentences provided by GTP feedback, so I usually use my phone to translate to help myself understand." Another stated, "ChatGPT sentences are too advanced and official, which may not be at our current level, so we will modify some sentences to improve them."

These comments suggest that while ChatGPT offers valuable support, it may sometimes exceed learners' current proficiency levels. This highlights the importance of supplementing AI use with teacher guidance to scaffold comprehension and ensure appropriate application of feedback.

Role of teacher guidance

The need for instructor support emerged as a critical factor in ensuring the productive use of ChatGPT. While students found the tool helpful, they acknowledged that teacher input was essential in clarifying difficult concepts and in adapting AI-generated suggestions to their own level of language proficiency. This underscores the complementary role of educators when integrating AI tools in writing instruction.

These themes align with the study's broader findings, which suggest that ChatGPT, when used in conjunction with guided reflection and teacher support, can enhance learner autonomy and writing proficiency.

DISCUSSION

Learning outcomes

The integration of ChatGPT with mobile devices played a major role in enhancing the effectiveness of this study. By enabling students to access ChatGPT (during class-time only) conveniently from their mobile devices, the study facilitated greater engagement and flexibility in learning. This accessibility allowed participants to practice and refine their writing skills anytime and anywhere, which is particularly important for international students who may face varying schedules and commitments. The mobile-friendly nature of ChatGPT enabled continuous interaction with the tool, and this may have contributed to significant improvements in writing proficiency, as evidenced by the higher satisfaction and confidence levels reported by the participants in EG 2. The use of mobile devices thus not only supported a more seamless integration of AI technology into the learning process but also underscored the potential of leveraging mobile technology to foster a more adaptive and responsive educational environment.

Research question 1 aimed to establish if ChatGPT, used as a paraphrasing tool for feedback to students' compositions would improve essay writing, engagement and motivation. The findings from the writing posttests indicate that ChatGPT significantly enhanced learning outcomes in English academic writing for EFL students. This may have been the result of many students in EG 2 having previously had very little academic writing experience, meaning that their learning curve was steepest, and this is reflected in the within group results. The results nonetheless align with previous findings from DDL studies finding improved fluency and consistency in writing (e.g., Muftah, 2023). This suggests that ChatGPT is a suitable substitute for corpus analysis software for the purpose of for raising learners' awareness of the organization, style and structural features of essays.

The results from our study demonstrated improvements across all groups; however, we acknowledge the limitations regarding the validity and interpretation of these findings. While the ANCOVA results indicated a significant effect of the teaching method on post-test scores, with EG 2 showing the most substantial improvement, it is crucial to note the variability in initial skill levels among participants. This was due to the need to conduct the research with intact classes where the control group and Experimental Group 1 were more advanced than Experimental Group 2. Despite this limitation, the Cohen's d test results provide valuable insights into the practical significance of the interventions. The Control Group and EG 1 exhibited small effect sizes (0.15 and 0.12, respectively), indicating modest improvements. In contrast, EG 2 displayed a very large effect size (2.56), suggesting a substantial enhancement in test scores. These findings underscore the effectiveness of the treatment applied to EG 2 in improving English academic writing skills, particularly for students with initially lower proficiency levels. Thus, while the results should be interpreted with caution, the evidence supports that the intervention notably benefitted student learning outcomes.

As an example of the improvements, one student for their thesis statement in their pretest stated: "Yes, I agree, because every country have different customs and traditions." Whereas for their post-test, their thesis statement was the following: "This essay agrees with the idea

that the internet is more popular than television." In another example one participant stated the following for their paraphrase question on their pretest: "People change the place when they leave the hometown and go to a new country." For their paraphrase question for their post-test, they stated the following: "It's argued that the internet has a bigger impact on people's lives because it is more popular than television."

In response to Research Question 2, the study revealed that ChatGPT use led to a noticeable boost in student motivation and engagement—areas that are often lacking in traditional EFL writing classes. This finding reinforces earlier research highlighting the value of interactive tools in writing development (e.g., Boulton, 2020), although more investigation is needed to understand exactly how these benefits take shape. A key feature appeared to be the act of asking ChatGPT to paraphrase students' original writing, which encouraged reflection by prompting learners to compare their versions with the AI's suggestions, which is an element central to DDL approaches.

Rather than simply accepting the AI's output, students were guided to examine differences in vocabulary, grammar, and structure. This process aligns with Vyatkina's (2016) claim that DDL fosters critical thinking about language and supports a deeper understanding of linguistic patterns. Working with IELTS-aligned worksheets, students engaged in inductive reasoning by drawing insights from the contrasts between their drafts and ChatGPT's version of their writing. This reflects Rowe's (2006) idea of hypothesis testing as a core aspect of constructivist learning, where learners actively assess and refine their understanding.

The experience also served as a form of scaffolded learning, which supports Storch's (2005) argument that effective feedback deepens students' understanding of the writing process. ChatGPT's interactive features allowed students to follow up with questions, ask for clarification, and adjust the AI's responses—closely mimicking collaborative learning environments where scaffolding helps drive language development. Unlike traditional corpora, which can be intimidating and require teacher support (Boulton & Cobb, 2017), ChatGPT functioned as a more user-friendly, responsive alternative. It helped students identify gaps in their writing and revise their work in accordance with form-focused instruction principles (Cobb & Boulton, 2015).

Importantly, interview data showed that students didn't blindly accept ChatGPT's feedback. Many made conscious edits to the AI's suggestions, showing thoughtful engagement. This supports Boulton's (2020) argument that learner interpretation of feedback fosters not just better writing, but greater autonomy and motivation. The increased engagement seen in EG 2 may therefore be linked not just to the novelty of AI, but to the active thinking and decision-making involved in processing ChatGPT's suggestions.

In addressing the final research question, the study also uncovered several challenges. Some students faced technical issues or lacked experience with prompting strategies, making it harder to use ChatGPT effectively without teacher guidance. While the platform's interface is generally easy to navigate, many learners still needed structured support to make the most of it. On the language side, some participants found ChatGPT's responses difficult to understand,

particularly when the language was too advanced or overly formal. Although they tried to work around these issues using translation tools or peer help, the feedback didn't always match the clarity and structure that a human teacher might provide. These findings highlight that while AI tools have potential, they work best when paired with clear instruction and pedagogical support.

To address limitations in digital access, such as a lack of devices or stable internet, another group used a paper-based DDL approach. These students worked with printed IELTS sample essays and guiding questions to engage in similar comparison-based learning. While this method also encouraged reflective thinking and built awareness of academic writing conventions, it didn't produce statistically significant gains compared to the control group. Still, its inclusion suggests that DDL strategies can be adapted for low-tech environments. With further refinement, such methods could provide a useful alternative or complement to AI tools, especially in contexts where digital resources are limited, helping to ensure more equitable access to language development.

Pedagogical implications

The results of this study point to several practical insights for EFL educators and curriculum planners who are exploring the use of ChatGPT and similar AI tools in teaching academic writing. These insights are organized into four main areas: the evolving role of teachers, curriculum development, fostering learner independence, and adapting assessment practices.

Teacher roles: From instructors to AI literacy coaches

With AI tools becoming more common in educational settings, EFL teachers are now expected to go beyond traditional instruction and take on the role of AI literacy coaches. This means they not only need to be comfortable using ChatGPT themselves, but also skilled in teaching students how to write effective prompts, understand the feedback they receive, and reflect on the differences between their writing and AI-generated suggestions. Equally important is teaching students about ethical usage, helping them learn when and how to use AI support responsibly, without becoming overly dependent on it. Especially in the early stages, teachers should provide clear guidance and structure to ensure ChatGPT is used to enhance, not replace, the learning process.

Curriculum design: Embedding AI strategically

Instead of using ChatGPT as an occasional or optional tool, educators should think about how to meaningfully integrate it into writing instruction. From a data-driven learning (DDL) perspective, this means designing activities that clearly align with lesson goals—such as improving coherence, vocabulary use, or grammatical accuracy. Teachers can use ChatGPT to help students build personal corpora from their own writing or to develop materials like worksheets and reflection prompts that encourage comparison between original drafts and AI-generated revisions. These strategies make DDL principles more accessible and can be tailored to both in-person and online teaching environments.

Student autonomy: Promoting self-regulated learning

One of the strengths of ChatGPT is its ability to offer instant feedback, giving students the chance to revise and improve their work independently. To make the most of this, instruction should aim to cultivate students' metalinguistic awareness and critical thinking—not just their technical skill with AI tools. Students should be encouraged to evaluate AI feedback thoughtfully, reflect on their language choices, and maintain ownership over their writing. This aligns with constructivist teaching approaches and supports the development of self-directed learning habits in writing.

Assessment strategies: Rethinking feedback and evaluation

The introduction of tools like ChatGPT also prompts a rethink of how writing is assessed. Teachers may want to include more process-focused assignments where students document how they used ChatGPT, explain the changes they made, and reflect on their revision choices. Assessment rubrics can be adapted to include criteria such as how well students engage with feedback, the strategies they use for revising, and their ability to justify edits—encouraging transparency and responsibility in how AI is used. In situations where students don't have easy access to technology, traditional DDL-inspired tasks—such as analyzing sample essays and completing guided comparison activities—remain effective and inclusive alternatives.

While this study demonstrates the potential of ChatGPT in supporting writing instruction, it also highlights important challenges. Issues like limited access, the need for training, and concerns about misuse suggest that thoughtful implementation is essential. Teachers should lead by example, showing how ChatGPT can be used ethically and effectively as a learning tool—not a substitute for student thinking. With the right support and guidance, ChatGPT can act as a flexible writing assistant that helps EFL learners better understand academic writing conventions, make meaningful revisions, and grow in both fluency and confidence.

Ethical considerations

Although this study showed that ChatGPT can be used effectively as a paraphrasing support tool to assist with original writing, it also brings to light several important ethical issues that need to be continuously addressed. In this study, ChatGPT was not positioned as a generator of new content, but rather as a reflective aid—helping students rephrase their thoughts while keeping their intended meaning intact. Insights from the semi-structured interviews revealed that many students became more aware of plagiarism risks and language use, with several reporting that using ChatGPT helped them learn how to cite sources more accurately and ethically.

Still, three main ethical concerns stand out and warrant further consideration. The first involves the risk of students becoming too dependent on AI-generated text. While some students engaged thoughtfully with the tool, others, especially those feeling stressed or unsure of their writing skills, might be tempted to rely too heavily on AI suggestions without critically evaluating them. This highlights the importance of teaching students how to use ChatGPT as a supportive

tool rather than a shortcut, and of guiding them to assess, modify, or even disregard AI feedback when necessary.

The second concern relates to the issue of AI hallucination, where language models like ChatGPT sometimes produce incorrect or misleading information. Although this study focused on paraphrasing rather than generating new content, there's still a chance that students might come across awkward phrasing or inaccurate grammar presented as correct usage. To guard against this, instructors should encourage learners to critically review ChatGPT's suggestions and verify any questionable information, especially in academic settings.

The third issue involves transparency and informed consent when using AI in the classroom. In this study, all participants were clearly informed about how ChatGPT would be used, and written consent was obtained. However, as AI tools become more embedded in everyday teaching, it will be crucial for institutions to make sure students are aware of how these tools function, what data might be shared, and how their interactions with AI could influence assessment and learning outcomes.

In the end, the study suggests that when used thoughtfully within a well-structured instructional design, tools like ChatGPT can actually support ethical development in writing, especially in helping students better understand paraphrasing and avoid plagiarism. That being said, to maintain academic integrity, it's essential that educators continue to monitor how AI is being used, provide clear guidelines, and create opportunities for students to reflect critically on their use of such technologies.

LIMITATIONS AND FUTURE RESEARCH

Given that this was a quasi-experimental study with intact groups, there was considerable variance between the groups from the outset, making direct comparison less reliable. Additionally, as the experiment was constructed to allow for individualized approaches to using ChatGPT, it was difficult to gauge every learner's engagement with the procedure. The relatively short duration of the program, spanning only 10 sessions, might also have constrained the depth of skill development achievable within the given timeframe. These factors suggest cautious interpretation when applying findings to different classroom, linguistic or cultural settings. Addressing these limitations may involve further investigation into the factors contributing to result disparities, extending the program duration for more comprehensive skill acquisition, and implementing measures to observe engagement and use of AI support.

As ChatGPT and AI broadly are relatively new in education and the wider world, future research directions are numberless. In ELT and more specific to writing pedagogy, future research could examine various types of integration into different classroom settings: (how) can AI be used during class time for developing writing skills? Can lower-level students make use of it? What other aspects and types of essay writing can be examined? are some of the questions to explore.

CONCLUSION

The results of this study indicate that ChatGPT used on mobile devices was significantly effective in enhancing English essay writing skills among EFL students. By leveraging ChatGPT's capabilities for personalized responses and interactive learning, the study found positive learner engagement, motivation, and comparatively superior development of written language proficiency. Overall, the study points towards the potential of AI-driven tools like ChatGPT in changing language education, suggesting a shift towards more technologically integrated and ethically aware educational practices.

STATEMENTS AND DECLARATIONS

Ethical approval for this study was obtained prior to any data being collected. All participants signed written consent forms to participate in this study, before the study commenced. The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

THE AUTHORS

Edward Devere Bacon is a PhD student at Thammasat University, as well as the Director of the International Language Center at Rangsit University, Thailand. He completed his Master's degree in Bilingual Education at Rangsit University, Thailand. His research areas include peer-assisted learning, academic writing, and ChatGPT in writing pedagogy.
superduperreadyeddie@rsu.ac.th

Harald Kraus is a lecturer from Thammasat University, Thailand. His current research interests include EFL writing pedagogy, intercultural communication and pragmatics.
harald.k@arts.tu.ac.th

REFERENCES

Abasi, A. R., Akbari, N., & Graves, B. (2006). Discourse appropriation, construction of identities, and the complex issue of plagiarism: ESL students writing in graduate school. *Journal of Second Language Writing, 15*(2), 102–117. https://doi.org/10.1016/j.jslw.2006.05.001

Ayoubi, K. (2024). Adopting ChatGPT: Pioneering a new era in learning platforms. *International Journal of Data and Network Science, 8*(2), 1341–1348. https://doi.org/10.5267/j.ijdns.2023.11.001

Banihashem, S. K., Kerman, N. T., Noroozi, O., Moon, J., & Drachsler, H. (2024). Feedback sources in essay writing: Peer-generated or AI-generated feedback? *International Journal of Educational Technology in Higher Education, 21*(1), Article 23. https://doi.org/10.1186/s41239-024-00455-4

Belda-Medina, J., & Calvo-Ferrer, J. R. (2022). Using chatbots as AI conversational partners in language learning. *Applied Sciences, 12*(17), Article 8427. https://doi.org/10.3390/app12178427

Bennett, G. (2010). *Using corpora in the language learning classroom: Corpus linguistics for teachers*. University of Michigan Press. https://doi.org/10.3998/mpub.371534

Boulton, A. (2012, July). *Wanted: Large corpus, simple software. No timewasters* [Paper presentation]. TaLC10: 10th International Conference on Teaching and Language Corpora, Warsaw, Poland. https://halshs.archives-ouvertes.fr/halshs-00938115

Boulton, A. (2019). Data-driven learning for younger learners: Obstacles and optimism. In P. Crosthwaite (Ed.), *Data-driven learning for the next generation: Corpora and DDL for pre-tertiary learners* (pp. 14–20). Routledge. https://hal.archives-ouvertes.fr/hal-02464274

Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning, 67*(2), 348–393. https://doi.org/10.1111/lang.12224

Boulton, A., & Pérez-Paredes, P. (2014). ReCALL special issue: Researching uses of corpora for language teaching and learning [Editorial]. *ReCALL, 26*(2), 121–127. https://doi.org/10.1017/S0958344014000068

Campbell, J., & Li, M. (2008). Asian students' voices: An empirical study of Asian students' learning experiences at a New Zealand university. *Journal of Studies in International Education, 12*(4), 375–396. https://doi.org/10.1177/1028315307299422

Chambers, A. (2019). Towards the corpus revolution? Bridging the research–practice gap. *Language Teaching, 52*(4), 460–475. https://doi.org/10.1017/S0261444819000089

Chambers, A., & O'Sullivan, Í. (2004). Corpus consultation and advanced learners' writing skills in French. *ReCALL, 16*(1), 158–172. https://doi.org/10.1017/S0958344004001211

Chang, J.-Y. (2014). The use of general and specialized corpora as reference sources for academic English writing: A case study. *ReCALL, 26*(2), 243–259. https://doi.org/10.1017/S0958344014000056

Cheng, Y., Fan, Y., Li, X., Chen, G., Gašević, D., & Swiecki, Z. (2025). Asking generative artificial intelligence the right questions improves writing performance. *Computers and Education: Artificial Intelligence, 8*, Article 100374. https://doi.org/10.1016/j.caeai.2025.100374

Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence, 4*, Article 100118. https://doi.org/10.1016/j.caeai.2022.100118

Cobb, T., & Boulton, A. (2015). Classroom applications of corpus analysis. In D. Biber & R. Reppen (Eds.), *The Cambridge handbook of English corpus linguistics* (pp. 478–497). Cambridge University Press. https://doi.org/10.1017/CBO9781139764377.027

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Cotterall, S., & Cohen, R. (2003). Scaffolding for second language writers: Producing an academic essay. *ELT Journal, 57*(2), 158–166. https://doi.org/10.1093/elt/57.2.158

Dergaa, I., Chamari, K., Zmijewski, P., & Saad, H. B. (2023). From human writing to artificial intelligence generated text: Examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport, 40*(2), 615–622. https://doi.org/10.5114/BIOLSPORT.2023.125623

Fariani, R. I., Junus, K., & Santoso, H. B. (2023). A systematic literature review on personalised learning in the higher education context. *Technology, Knowledge and Learning, 28*(2), 449–476. https://doi.org/10.1007/s10758-022-09628-4

Flowerdew, L. (2012). *Corpora and language education*. Palgrave Macmillan. https://doi.org/10.1057/9780230355569

Fraenkel, J., Wallen, N., & Hyun, H. (1993). *How to design and evaluate research in education* (2nd Ed.). McGraw-Hill.

Fryer, L. K., Nakao, K., & Thompson, A. (2019). Chatbot learning partners: Connecting learning experiences, interest and competence. *Computers in Human Behavior, 93*, 279–289. https://doi.org/10.1016/j.chb.2018.12.023

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

Götz, S. (2012). Testing task-types in data-driven learning: Benefits and limitations. In K. Biebighäuser, M. Zibelius, & T. Schmidt (Eds.), *Aufgaben 2.0 – Konzepte, materialien und methoden für das fremdsprachenlehren und -lernen mit digitalen medien* [Tasks 2.0 – Concepts, materials and methods for foreign language teaching and learning with digital media] (pp. 249–276). Narr.

Hadley, G. (2002). An introduction to data-driven learning. *RELC Journal, 33*(2), 99–124. https://doi.org/10.1177/003368820203300205

Hamilton, I. (2023, December 5). *Artificial intelligence in education: Teachers' opinions on AI in the classroom*. Forbes. https://www.forbes.com/advisor/education/it-and-tech/artificial-intelligence-in-school/

Han, J., Yoo, H., Kim, Y., Myung, J., Kim, M., Lim, H., Kim, J., Lee, T. Y., Hong, H., Ahn, S. -Y., & Oh, A. (2023). RECIPE: How to integrate ChatGPT into EFL writing education. *Proceedings of the Tenth ACM Conference on Learning @ Scale, 2023*, 416–420. https://doi.org/10.1145/3573051.3596200

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

Kaufman, D. (2004). Constructivist issues in language learning and teaching. *Annual Review of Applied Linguistics, 24*, 303–319. https://doi.org/10.1017/S0267190504000121

Khojasteh, L., Kafipour, R., Pakdel, F., & Mukundan, J. (2025). Empowering medical students with AI writing co-pilots: Design and validation of AI self-assessment toolkit. *BMC Medical Education, 25*(1), Article 159. https://doi.org/10.1186/s12909-025-06753-3

Kim, H.-S., Cha, Y., & Kim, N. Y. (2021). Effects of AI chatbots on EFL students' communication skills. *Korean Journal of English Language and Linguistics, 21*, 712–724.

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023a). ChatGPT for language teaching and learning. *RELC Journal, 54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023b). Exploring generative artificial intelligence preparedness among university language instructors: A case study. *Computers and Education: Artificial Intelligence, 5*, Article 100156. https://doi.org/10.1016/j.caeai.2023.100156

Kostka, I., & Toncelli, R. (2023). Exploring applications of ChatGPT to English language teaching: Opportunities, challenges, and recommendations. *Teaching English as a Second or Foreign Language--TESL-EJ, 27*(3). https://doi.org/10.55593/ej.27107int

Lin, C.-J., & Mubarok, H. (2021). Learning analytics for investigating the mind map-guided AI chatbot approach in an EFL flipped speaking classroom. *Educational Technology & Society, 24*(4), 16–35. https://www.jstor.org/stable/48629242

Liu, M. (2023). Data-driven tertiary English learning: A study of the application of self-built corpus for teaching EFL reading. In S. Jia & H. Dong (Eds.), *Sixth international conference on computer information science and application technology (CISAT 2023)*. SPIE. https://doi.org/10.1117/12.3003816

Mao, L., Liu, Y., & Zhang, M. (2018). Research on the effectiveness of college student English writing teaching based on data-driven learning. *Educational Sciences: Theory & Practice, 18*(5), 1160–1169. https://doi.org/10.12738/estp.2018.5.017

McKinley, J. (2015). Critical argument and writer identity: Social constructivism as a theoretical framework for EFL academic writing. *Critical Inquiry in Language Studies, 12*(3), 184–207. https://doi.org/10.1080/15427587.2015.1060558

Muftah, M. (2023). Data-driven learning (DDL) activities: Do they truly promote EFL students' writing skills development? *Education and Information Technologies, 28*(10), 13179–13205. https://doi.org/10.1007/s10639-023-11620-z

O'Donnell, M. (2013). From learner corpora to curriculum design: An empirical approach to staging the teaching of grammatical concepts. *Procedia – Social and Behavioral Sciences, 95*, 571–580. https://doi.org/10.1016/j.sbspro.2013.10.684

O'Keeffe, A., & Mark, G. (2022). Principled pattern curation to guide data-driven learning design. *Applied Corpus Linguistics, 2*(3), Article 100028. https://doi.org/10.1016/j.acorp.2022.100028

Ostertagová, E., & Ostertag, O. (2013). Methodology and application of one-way ANOVA. *American Journal of Mechanical Engineering, 1*(7), 256–261. https://doi.org/10.12691/ajme-1-7-21

O'Sullivan, Í., & Chambers, A. (2006). Learners' writing skills in French: Corpus consultation and learner evaluation. *Journal of Second Language Writing, 15*(1), 49–68. https://doi.org/10.1016/j.jslw.2006.01.002

Özçelik, N. P., & Ekşi, G. Y. (2024). Cultivating writing skills: The role of ChatGPT as a learning assistant—A case study. *Smart Learning Environments, 11*(1), Article 10. https://doi.org/10.1186/s40561-024-00296-8

Phakiti, A., & Li, L. (2011). General academic difficulties and reading and writing difficulties among Asian ESL postgraduate students in TESOL at an Australian university. *RELC Journal, 42*(3), 227–264. https://doi.org/10.1177/0033688211421417

Pospelova, T. (2021). The collaborative discussion model: Developing writing skills through prewriting discussion. *Journal of Language and Education, 7*(1), 156–170. https://doi.org/10.17323/jle.2021.10748

Robinson, J., Dusenberry, L., Hutter, L., Lawrence, H., Frazee, A., & Burnett, R. E. (2019). State of the field: Teaching with digital tools in the writing and communication classroom. *Computers and Composition, 54*, Article 102511. https://doi.org/10.1016/j.compcom.2019.102511

Rowe, K. (2006). *Effective teaching practices for students with and without learning difficulties: Constructivism as a legitimate theory of learning and teaching*. Australian Council for Educational Research. https://research.acer.edu.au/learning_processes/10

Seo, K., Tang, J., Roll, I., Fels, S., & Yoon, D. (2021). The impact of artificial intelligence on learner–instructor interaction in online learning. *International Journal of Educational Technology in Higher Education, 18*(1), Article 54. https://doi.org/10.1186/s41239-021-00292-9

Sepehri, M. (2015). The effects of data driven learning on Iranian EFL learners' writing skills development. University of Birmingham.

Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing, 14*(3), 153–173. https://doi.org/10.1016/j.jslw.2005.05.002

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th Ed.). Pearson.

Vyatkina, N. (2016). Data-driven learning for beginners: The case of German verb-preposition collocations. *ReCALL, 28*(2), 207–226. https://doi.org/10.1017/S0958344015000269

Yılmaz, M. (2017). The effect of data-driven learning on EFL students' acquisition of lexico-grammatical patterns in EFL writing. *Eurasian Journal of Applied Linguistics, 3*(2), 75–88. https://doi.org/10.32601/ejal.460966

Yoon, H., & Hirvela, A. (2004). ESL student attitudes toward corpus use in L2 writing. *Journal of Second Language Writing, 13*(4), 257–283. https://doi.org/10.1016/j.jslw.2004.06.002

Yusuf, A., Pervin, N., & Román-González, M. (2024). Generative AI and the future of higher education: A threat to academic integrity or reformation? Evidence from multicultural perspectives. *International Journal of Educational Technology in Higher Education, 21*(1), Article 21. https://doi.org/10.1186/s41239-024-00453-6

Zhu, F. (2021). Supporting EFL writing during the pandemic: The effectiveness of data-driven learning in error correction. *English as a Foreign Language International Journal, 1*(2), 8–28. https://doi.org/10.56498/66122021

Appendix 1

Results of questionnaire items of experimental group 1

Attitudes towards paper-based data-driven learning

Std. No	Items	Mean	SD	Level of Satisfaction
1	The use of paper-based data-driven learning has significantly improved my academic writing skills.	4	0.88	High
2	Paper-based data-driven learning enhances my ability to structure and organize academic papers effectively.	4	0.78	High
3	I believe that paper-based data-driven learning is an essential tool for improving academic writing.	4	0.78	High
4	The feedback I receive on my academic writing using paper-based data is consistently helpful.	4	0.84	High
5	I am highly motivated to engage with paper-based data-driven learning for academic writing.	4	1.06	High
6	Paper-based data-driven learning has substantially enhanced my ability to make evidence-based arguments in my writing.	4	0.89	High
7	Paper-based data-driven learning helps me identify and select relevant sources for my research with ease.	4	1.37	High
8	I believe that paper-based data-driven learning is a practical and valuable approach to improving academic writing.	4	1.02	High
9	Paper-based data-driven learning enhances my ability to critically evaluate sources for academic writing effectively.	4	0.96	High
10	I am highly satisfied with the guidance and support provided for using paper-based data-driven learning materials for academic writing.	4	0.94	High
11	Paper-based data-driven learning is the most effective way to apply theoretical knowledge in academic writing.	3	1.01	Moderate
12	I believe that my overall academic writing skills have improved significantly through paper-based data-driven learning.	4	0.96	High
13	The variety of paper-based data sources available greatly contributes to my academic writing skills development.	4	0.90	High
14	Paper-based data-driven learning encourages me to think creatively when integrating data into my writing.	4	0.82	High
15	The hands-on nature of paper-based data-driven learning significantly enhances my learning experience.	4	0.85	High
16	Paper-based data-driven learning allows me to explore different perspectives and angles in my academic writing.	4	0.63	High
17	Overall, paper-based data-driven learning has been instrumental in honing my academic writing skills to a high standard.	4	0.80	High
	Mean	4	0.91	High

Attitudes towards paper-based corpus consultation

Std. No	Items	Mean	SD	Level of Satisfaction
1	I find it easy to incorporate data from paper-based sources into my academic writing.	4	0.79	High
2	I prefer using paper-based materials for data-driven learning over digital resources for academic writing.	3	1.17	Moderate
3	I actively seek out paper-based sources when conducting research for my academic writing.	4	1.26	High
4	I feel that paper-based data-driven learning is more reliable than digital sources when it comes to academic writing.	3	1.13	Moderate
	Mean	3.5	1.09	Moderate

Affective factors

Std. No	Items	Mean	SD	Level of Satisfaction
1	I feel a strong sense of accomplishment when I successfully incorporate data from paper-based sources into my academic papers.	5	1.13	Highest
2	I feel more confident in my ability to effectively communicate complex ideas through my academic writing due to paper-based data-driven learning.	4	1.01	High
	Mean	4.5	1.07	High

Appendix 2

Results of questionnaire items of experimental group 2

Data-driven approaches

Std. No	Items	Mean	SD	Level of Satisfaction
1	Data-driven learning can improve student writing skills.	4	1.10	High
2	You are confident in using data-driven learning to make decisions about writing instructions.	4	1.06	High
3	You enjoy using data-driven learning to guide your instructional planning for writing.	4	1.19	High
4	Data-driven learning can help participants become more motivated and engaged in writing activities.	4	1.03	High
5	You are motivated to improve your writing skills through data-driven learning.	4	1.19	High
6	It is easy to access and interpret the data related to your writing skills.	4	0.99	High
7	Data-driven learning helps you identify areas of improvement in your writing.	4	0.96	High
8	Data-driven learning is important in achieving your writing goals.	4	1.19	High
9	You are likely to continue using data-driven learning techniques to improve your writing skills.	4	1.10	High
10	You encourage your future university lecturers to integrate data-driven learning into future course curriculums.	4	1.07	High
	Mean	4	1.09	High

Artificial intelligence

Std. No	Items	Mean	SD	Level of Satisfaction
1	Artificial intelligence improves your writing skills.	4	1.33	High
2	You are confident in your ability to use artificial intelligence tools to improve your writing skills.	3.5	1.23	High
3	You feel motivated to improve your writing skills through artificial intelligence.	4	1.11	High
4	Artificial intelligence tools are helpful in identifying grammar and spelling errors in your writing.	4	1.06	High
5	It is easy to access and use artificial intelligence tools for improving your writing skills.	4	1.20	High
6	Artificial intelligence helps you identify areas of improvement in your writing.	4	1.13	High
7	You are likely to continue using artificial intelligence tools to improve your writing skills.	4	1.01	High
	Mean	3.9	1.15	High

ChatGPT

Std. No	Items	Mean	SD	Level of Satisfaction
1	Artificial intelligence, by way of ChatGPT, improves your writing skills.	4	1.11	High
2	Artificial intelligence is important in achieving your writing goals.	4	1.23	High
3	You encourage other university lecturers to integrate artificial intelligence into future course curriculums.	4	1.19	High
4	You are confident in your ability to use ChatGPT to improve your writing skills.	4	1.16	High
5	You feel motivated to improve your writing skills through ChatGPT.	4	1.01	High
6	ChatGPT is helpful in generating ideas for your writing.	4	1.06	High
7	It is easy for you to use ChatGPT to improve your writing skills.	4	1.11	High
8	ChatGPT helps you identify areas of improvement in your writing.	4	1.20	High
9	You are likely to continue using ChatGPT to improve your writing skills.	4	1.16	High
10	You are comfortable with using ChatGPT to generate content for your writing.	4	1.06	High
11	ChatGPT provides you with personalized feedback on your writing.	4	1.16	High
12	You encourage other university lecturers to integrate ChatGPT into future course curriculums.	4	1.20	High
13	ChatGPT has positively influenced my confidence in my English writing abilities.	4	1.13	High
14	I feel more motivated to write in English as a result of using ChatGPT.	4	1.33	High
15	ChatGPT has helped me believe in my capacity to write effectively in English.	4	1.12	High
16	I have become more self-assured in my English writing skills because of ChatGPT.	4	1.26	High
17	My academic English writing has improved since I started using ChatGPT.	4	1.13	High
18	ChatGPT has enhanced the clarity and coherence of my academic English writing.	4	1.17	High
19	I believe ChatGPT has positively impacted the organization and structure of my academic English essays.	4	1.19	High
20	The use of ChatGPT has helped me produce more grammatically accurate academic English writing.	4	1.20	High
21	Overall, I am satisfied with the effects of ChatGPT on my writing self-efficacy and academic English writing skills.	4	1.14	High
22	I would recommend ChatGPT to other EFL students looking to improve their English writing skills.	4	1.26	High
	Mean	4	1.16	High