International Journal of Learning, Teaching and Educational Research
Vol. 24, No. 5, pp. 644-661, May 2025
https://doi.org/10.26803/ijlter.24.5.33
Received Mar 4, 2025; Revised Apr 27, 2025; Accepted May 15, 2025

Effects of the Use of Generative AI Tools on EAP Writing Development: A Case Study with Medicine Undergraduates

Abdullah Alshakhi*
English Language Institute, King Abdulaziz University
Jeddah, Saudi Arabia

Abstract. The present empirical study examines the impact AI tools intervention may have on the writing development of undergraduate EFL learners. This mixed-methods research was conducted at a Saudi university with sixty students as participants to teach writing short essays focused on medical issues. The participants were taught to plan, organize, draft, revise and prepare the final drafts of the essays on given topics using AI tools, such as ResearchRabbit for referencing, Acrobat Chat with PDFs for summary, and Otio for grammar check and planning, drafting and revising the essays. The study involved experimental teaching, conducted for 4 weeks. The participants were undergraduate students majoring in Medicine and learning English for one year as a pre-requisite for university study. After a pre-test, two groups- experimental and control- were formed dividing the research subjects randomly. The research subjects in the experimental group were taught to use AI tools to search relevant medical terminology, search reference materials and prepare bibliography, checking English grammar and usage, organizing, planning, drafting and revising their essays. On the other hand, research subjects in control group were taught to write essays on the same topics using a traditional approach, that is, finding relevant terms using online or offline dictionary, using their knowledge of grammar to write error-free essays, finding relevant ideas from online/offline sources, planning, drafting and writing the essays. The control group participants were strictly forbidden to take AI help in any way. After the experimental teaching, the groups were given a post-test. The marks obtained by participants were compared. The mean of marks obtained by experimental group participants was higher by 2.95 points, with t-test value 8.83 (statistically significant at p .05). The t-test value obtained on a comparison of the pre-assessment and post-assessment scores of the experimental group participants was 4.743 (statistically significant at p .05). Thus, the research findings show positive effects of AI tools intervention on participants' writing development.

*Corresponding author: Abdullah Alshakhi; *amalshaikhi1@kau.edu.sa*

Keywords: EFL research; AI in education; writing development; use of AI in writing; EAP

1. Introduction

Artificial Intelligence (henceforth AI) has been harnessed as an aid to natural human intelligence for quite some time now, however, once the Open AI model ChatGPT got launched in 2022, AI tools started being used as educational aids on an unprecedented scale. This intervention of the AI model in education caused happiness as well as concerns in the academic circles. The cause for happiness is that the AI model is loaded with unbelievable function features and is useful in so many daily life applications, thus saving time and effort. But at the same time, this AI application is so richly self-sufficient to provide answers to queries and presenting solutions at a high speed that it leaves everyone worried about the future of human endeavours, particularly concerning the fields of human activity, such as academics, where human intelligence, efforts, and hard work are valued and rewarded. ChatGPT, and other AI models, have placed in human hands the power to claim value and reward without putting in efforts, hard work, or even intelligence.

In the academic field, used as support instruments, AI models can enhance teaching and learning in a big way, positively affecting cognitive/linguistic development and helping teachers to become more creative in class. Whereas, used as replacements for human efforts, the same models have the power to blur the thin line between human endeavour and machine output, between realism and verisimilitude. For instance, learners can become better and faster writers using AI tools to learn writing, while the same tools can write for learners whatever they need, such as essays, reports, papers, and so on. The learner need not learn any writing; the AI tools will produce all kinds of writing on demand, while the user being undetected of the fraud.

However, the thrust of the present research was to study the positive aspects of the changes the intervention of expert systems has brought about in education. Motivation for the present study lay in self-reported success stories of teachers and researchers who used expert systems in classrooms to enhance their teaching experience.

1.1 Research Background

Saudi Arabian education system teaches English as a foreign language (EFL) and since university courses for all educational streams are offered in English, learning English is mandatory for all students enrolled in undergraduate courses. All university students learn English for a year. The focus of teaching is English for academic Purposes (EAP) with emphasis on the major course of study of the students. For instance, students enrolled in Health Sciences discipline are taught English with emphasis on medical English. In general, on joining the course, the students possess background knowledge in writing in English but not sufficient knowledge to write long, focused essays with supportive research on any given topic (Al-Mohanna, 2024). Teachers encourage students' learning autonomy and

guide them to self-learning, yet most often the strategy falls short of expectations, for various reasons. The rise of AI tools helpful in teaching and learning has given some hope to teachers as well as learners in this regard (Son et al., 2023). Teachers using AI tools as teaching aids can save a lot of their time spent in preparing and delivering the lessons, conducting formative assessments, giving feedback, marking, and preparing exam results. Similarly, learners can learn a few things faster with the help of AI tools, such as preparing summaries of long articles, making a reference list, planning and organizing an essay and checking the writing for grammatical and structural errors (Wilson, 2022; Yuan & Liu, 2025).

1.2 Research Problem

In his class, teaching medical English focused on writing skills development, the researcher noted that students commonly struggled with writing short essays. On closer scrutiny, it was found that the main issues in students' writing were spelling errors, syntax errors, punctuation errors, organizational errors (missing coherence, missing unity), and absence of research on the given topic. Additionally, the students had no idea how to prepare a references list, and summarizing a long article was very difficult for them since their reading skills were also weak (Alotaibi, 2022). The number of students in each class was generally high, so, it was difficult for the teacher to devote personal attention to each learner. The teacher/researcher had to essentially find some ways to supplement teachers' class instructions with self-learning device accessible to students to enhance their learning at their own pace. Of late, several AI tools have come up usable as aids to teach writing that allow teachers to "multiply" themselves in many ways.

Learners are aware of the use of AI in various smart phone applications and AI-powered chat-based services aimed at enhancing productivity, creativity and information understanding, available in the latest versions of mobile phone operating systems and computer operating systems. But learners may not be using these services for language learning and skills development. They needed to be taught to use freely available machine learning models for the development of expertise in writing, and the strategy effectiveness was to be investigated in an empirical, quasi-experimental study.

However, prior to teaching his students the ways to avail the services of machine learning tools, the investigator reviewed numerous research studies on the probable positive effects of these tools on enhancing the development of English writing skills in adult EFL teaching environments. But most of the success stories are based on teachers' use of AI tools as aids to teaching and the perceived enhancement in their students' learning development, reported without any formal research on the exact cause and effect relationship between the input and the output. The studies are either theoretical inputs, or present a scenario where only known benefits of AI tools in developing writing skills have been highlighted. They do not present any results documented after their research studies were conducted to teach students the use of machine learning to sharpen their writing skills and then reporting their findings. Thus, there is a lack of focused research literature on the topic. Therefore, the present research was an

attempt to understand what positive effects expert systems can have on the learning curve of students. The study contributes in its modest way to the growth of research literature on the subject.

1.3 Research Aims

The current study was primarily conducted to examine whether expert systems intervention can positively influence undergraduate EFL learners' expertise in writing skills who learn English in a non-native environment. As a secondary objective, the current study was taken up to add to the growth of reference materials citing the effects of expert systems intervention tools on foreign language learners' writing skill development.

2. Literature Review

2.1 GenAI

Artificial Intelligence (AI), also known as "expert systems," is the expertise of machines, such as computer systems, in executing commands using software that enable the machines to perceive their surroundings, learn from it, and take action. Generative Artificial Intelligence (GenAI) is a subsystem of expert systems which is capable of producing a variety of output in response to prompts. GenAI is a step forward since it is capable of learning the underlying patterns of its training data and then producing new data not fed into the system (Farrelly & Baker, 2023). Thus, GenAI can produce a variety of texts, images, audios, videos, text-to-images, text-to-videos, designs, and other forms of data. At present the most talked about GenAI model is ChatGPT which is an AI chatbot developed by Open AI. The model is designed to respond to users' prompts and questions in a humanlike fashion. This is possible as the model is fundamentally a computational method built upon a huge data corpus. Májovský et al. (2023), for instance, offers the following explanation:

> "*At its core, ChatGPT is a large neural network trained on a massive corpus of text data, such as books, articles, and web-based content. The model is a multi-layered network capable of self-attention and feedforward, which facilitate it to recognize and imitate complex relationships between morphological and syntactical elements in natural language texts*" (e46924).

The neural network model is inspired by neuronic structures found in the brain of man and it is made up of three layers - the outer layer for input, a layer or two hidden in the middle, and a layer meant for output. As far as writing tasks are concerned, the model provides the users with plenty of ways to improve their writing as well as a mechanism to receive feedback, boosting their expertise in writing and performance. The consolidated use of this technology as a teaching aid in learners' writing development is, thus, founded on the principles of cognitive constructivism and sociocultural theory that derive on Vygotsky's (1978, p. 86) ideas of the "zone of proximal development and scaffolding" (Jingxin & Razali, 2020; Piaget, 1973).

2.2 Generative AI Tools and Writing Development

As described in the previous section, a few GenAI models have been particularly designed to help users in learning the specifics of writing and perform better.

Learners, especially in non-native English-speaking countries, at times find it hard to acquire writing skill in English, and since teachers in those countries spend a lot of their time in preparing lesson plans and evaluating students' formative and summative assessment papers, they are hardly left with any time to provide the much-needed individual attention to their students' writing skill development. Writing skill development for adult learners, such as undergraduate EFL learners, may be understood as involving five major steps: brainstorming for concepts on the chosen subject of writing, collecting data on the subject from relevant sources, planning the preliminary write-up, revision, and writing the last version of the essay. Whereas major hindrances in writing skill development are cognitive barriers, lack of ideas, lack of research on the topic, linguistic hindrances, and lack of planning. The GenAI tools available at present are helpful for teachers as well as learners to deal with each of the steps in writing good essays on any given topic.

Taking help from programs like CustomGPT, teachers can create their own tailored ChatGPT-style chatbots trained on data that is specific to their class (Clay, 2025). For instance, Gayed et al. (2022) created a web application based on AI called "AI KAKU" which is useful for adult foreign language learners to tackle the difficulties they face in writing short/long passages in English. The researchers evaluated how AI KAKU can positively influence the writing expertise of the subject population and reported that, compared to traditional word processors, the AI application was potentially a more useful tool for EFL learners since they needed more structured assistance.

In addition to helping shape ideas, plan and execute the writing, GenAI tools support learners' writing skill development in many other ways, such as providing personalized learning and instant feedback, providing constructive feedback, enhancing creativity and learner engagement, supporting learner autonomy, addressing specific needs of adult EFL learners, and overcoming learners' writing anxiety by building confidence (Giglio & Costa, 2023; Golan et al., 2023; Guan et al., 2024; Kung et al., 2023; Xu & Wang, 2024). Adult EFL learners often make a heterogenous group differing from each other in many ways, such as learning background, English proficiency level, and objectives of learning English.

Generative AI can provide personalized learning experiences to them by tailoring writing prompts, exercises, and feedback for them to meet individual needs, offering targeted practice and support (Baidoo-Anu & Ansah, 2023; Dergaa et al., 2023; Marzuki et al., 2023). For instance, Generative AI-based platforms can analyse learners' writing, identify areas for improvement, and provide specific suggestions to enhance grammar, vocabulary, and coherence in their writings. This personalized approach helps learners focus on their weaknesses and track their progress over time.

As has been discussed above, it is a common observation that EFL teachers in conventional classroom settings might not be in a position to address feedback instantly on writing assignments since they have to spend a large amount of their time in other pedagogy related activities, such as lesson plan, evaluation, and

administrative work, etc. Generative AI tools integrated into teachers' teaching plan can offer immediate and supportive evaluation of learners' writings so that they correct their errors and enhance their writing skill (Jingxin & Razali, 2020). Learners can use GenAI tools, such as Grammarly or ProWritingAid, to detect grammatical errors in their writings, get vocabulary enhancement suggestions, and resolve stylistic issues (Golan et al., 2023; Jarrah et al., 2023; Kacena et al., 2024; Kung et al., 2023). This instant feedback helps adult learners to continuously refine their writing skills, fostering greater confidence and autonomy in their language use (Herft, 2023).

Similarly, the use of GenAI tools can help stimulate creativity and engagement among adult EFL learners (Dwivedi et al., 2023; Gayed et al., 2022). Generative AI can encourage learners to explore different genres and writing styles by generating diverse and contextually relevant writing prompts (Giglio & Costa, 2023; Guan et al., 2024). For instance, teachers and learners can get assistance from AI-based writing assistants to get creative story starter ideas, debate topics, and persuasive essay prompts, which will enhance their critical thinking and creativity. Moreover, learning may become more interesting and motivating with the use of interactive writing platforms or virtual writing communities, promoting sustained learner engagement (Wale & Kassahun, 2024).

Learner autonomy has gained more currency in recent years as GenAI language learning tools support learner autonomy in a big way. The GenAI writing tools, for instance, provide access to AI-based resources, instant feedback and progress tracking, and thus offer flexible and self-paced learning opportunities, allowing learners to practice writing at their convenience (Herft, 2023; Jingxin & Razali, 2020). Adult learners often juggle multiple responsibilities making it difficult to attend regular classes, so, the flexibility to learn without relying on fixed schedules or physical classroom environments proves to be a great help to them.

The learning needs of Adult EFL learners are often more distinct and complex compared to the needs of younger learners, such as professional communication, academic writing, research publications, or specific industry-related language skills (Golan et al., 2023; Jarrah et al., 2023; Kacena et al., 2024). It is not possible to learn everything in classrooms. GenAI tools cater to such specific and complex requirements as they provide specialized writing resources and practice materials, such as business email templates, academic paper outlines, industry-specific vocabulary lists, and so on (Kung et al., 2023). The readily available assistance also helps learners overcome writing anxiety by building confidence, especially in adult learners who may feel self-conscious about their [low] proficiency in writing. The learning and practice atmosphere created by GenAI tools is such that learners feel supported and free from prejudices. They also provide learners opportunities to experiment with different writing styles, to receive constructive feedback, and thus build their confidence gradually (Guan et al., 2024; Khalifa & Albadawy, 2024). Learners can even take risks and be more expressive with the reassurance of having a virtual assistant to guide and support them.

Research studies highlighting the positive side of integrating GenAI tools in EFL writing pedagogy are aplenty, most often underlining the desired effects of the tools on both teaching and learning. However, most of these studies are either reviews of the theoretical underpinnings of the technology used in education or not based on empirical, classroom-based enquiries. For example, the study by Giglio and Costa (2023) is a comprehensive mapping of recent studies that survey the use of expert systems to write scientific articles, particularly to help non-native English speakers. The researchers used the relevant search terms like 'artificial intelligence' and 'scientific writing' for the purpose. They state in their findings from the reviewed studies that for non-native speakers of English researchers and scientists for whom there exist linguistic hindrances in publication, AI tools can be very useful to improve their scientific writing. GenAI tools, such as Elicit, ResearchRabbit, Scispace, and Copilot are useful to look for relevant science research studies, prepare summarized documents in pdf format, and prepare a good references list.

Learners can also get help in writing various sections of their manuscripts. Syntactic and spelling errors are taken care of by handy software tools, such as ChatGPT. The review study by Xu and Wang (2024) indirectly supports the findings of Giglio and Costa (2023). Xu and Wang (2024) found that AI tools are very effective in supporting writing in English. Students who integrated AI in learning English always outperformed their tradition-bound counterparts. The study by Golan et al. (2023) is an evaluation of expert systems useful in writing, and the researchers recommend that AI-based tools should widely be adopted by writers and researchers, especially for scientific writing.

They specifically list the tools helpful in scientific writing: Semantic Scholar, Penelope.ai, and Elicit for literature review, Writeful, CoSchedule Headline Analyzer, Quillbot, Wordtune, and ChatGPT for writing, and DALL-E 2 for figures. On similar lines, the study by Guan et al. (2024) is a meta-representational synthesis of research on how GenAI affects foreign language learning, which concluded that in the selected research GenAI draws favourable results on foreign language learning, though it did not show any significant effects on learning motivation. Overall, GenAI is found to be versatile in enhancing language learning outcomes. The study by Kung et al. (2023) too is not a classroom-based investigation, yet it shows the power of GenAI to perform classroom and examination related tasks seamlessly in a humanlike fashion. GenAI tools can easily come up with answers related to medical and engineering examination questions.

2.3 Impact of Expert Systems on Writing Instruction

Once again, research works available on the probability of GenAI influencing writing instruction [if integrated into teaching plans] are mostly review studies, and quantitative, classroom-based real-time research investigations are just scanty. For instance, Baidoo-Anu and Ansah's (2023) article reviews the literature available on the relevance of ChatGPT in education and lists some potential benefits of the tool they found in the literature. The researchers say that ChatGPT has its own limitations too. For example, it can give false information, and its

training data may be biased. Dergaa et al. (2023) also reviewed the literature available on the expected advantages/disadvantages of ChatGPT to be used as teaching aid, particularly highlighting the ethical considerations and impact of the GenAI on the genuineness and originality of students' essays and other writings. In the opinion of the researchers, AI tools are good to enhance students' writings, but they put the genuineness and originality in students' essays at risk. The researchers emphasize that there is a need for an encompassing exchange of views among scholars concerning the threat of misuse of AI tools in the academia.

Khalifa and Albadawy's (2024) article is also a review of the existing literature on the topic. The researchers reviewed 24 studies that investigated particular study fields supporting academic writing and influenced by AI, such as facilitating generation of ideas and research design, content and structure improvement, literature review support, and so on. Khalifa and Albadawy (2024) say that expert systems have largely revolutionized students' essay-writing and research work in all the fields of study.

However, there does exist some literature throwing light on the impacts of GenAI tools - Writerly and Google Docs - integrated into EFL teaching, such as the study by Wale and Kassahun (2024). This study collected data by common data collection methods from a large sample of participants. The report of the researchers is that teaching/learning of English is favourably affected by the integration of GenAI tools into teaching plans.

2.4 Criticism of, and Apprehensions about, GenAI

The studies reviewed in the forgone sub-sections underscore the probable favourable impact of expert systems integrated into foreign language instruction. However, despite all the potential and proven benefits of GenAI in education, researchers have begun raising concerns about the flip side of the AI tools in teaching and learning, especially after the launch of ChatGPT, an artificial intelligence-based program that can perform all kinds of writing tasks at the command of the user (Williams, 2023). Leuenberger (2024), for example, says that if used regularly to perform tasks that need human creativity, artificial intelligence bears the risk of regression in human capacity to create and innovate. If we assume total dependence on AI to perform tasks that we ourselves should, we risk forgetting to do the tasks ourselves. There exists a danger that total reliance on algorithms will reshape the entire human identity without them realizing it.

The essays in the book *AI Morality*, edited by Edmonds (2024), are explorations in how AI is all set to revolutionize human life and the moral dilemmas it will trigger, such as the issues around privacy, bias, transparency, accountability, and autonomy. Edmonds' write-up, entitled "Should You Let AI Tell You Who You Are and What You Should Do?" makes a point that the algorithms used in expert systems, which is the foundation of today's digital platforms, may know more about humans than humans know themselves. The question the researcher poses is: "Can we trust them to make the best decisions for us, and what does that mean for our agency?" Regarding issues concerning authenticity of human skill and academic integrity, Májovský et al. (2023) demonstrated that artificial intelligence

is capable of writing very strong, scholarly articles on any issue, such as a medical issue, in a matter of hours, which would be completely made up but highly persuasive. It requires very little human labor and expertise. Equally true is the disturbing fact that expert systems can be exploited to generate many types of counterfeit academic texts.

On similar lines, Morris (2018) draws readers' attention to the common problem of contract cheating. Students in higher education have been using ghost writing services (Elsen-Rooney, 2023) customized for their study streams. The researcher warns higher education institutions to be vary of the practice and find a way to deal with the menace. The study by Werdiningsih et al. (2024) also underlines the significance of expert systems tools in teaching/learning but with a crucial balance of AI tools and human judgment for authenticity. In the words of Werdiningsih et al. (2024), owing to the rise of expert systems, authenticity and integrity of academic work is at risk. Academic communities are in need of ethical guidelines, and they should encourage critical thinking.

However, expert systems are also bound by certain limitations. It is reported that at times they provide extremely complex guidelines to the user, and that their training lacks cultural sensitivity. In a story published by *The Associated Press* (October 26, 2024), Burke and Schellmann (2024) draw readers' attention to a major inadequacy in Whisper, a transcription tool developed by OpenAI: The program quite often creates made up text chunks or even full sentences. The experts in the field commented that some of the texts made up by the program (called "hallucinations") are found to be racial commentary, aggressive arguments and even treatment plans that are out of the world. While the creators of Whisper claim that the program is as powerful and accurate as human intelligence.

To sum up, the research studies reviewed above can be cited to support the argument that empirical, classroom-based investigations into the impacts of integrating GenAI tools on adult EFL teaching and learning are very scarce at present. There is a high concentration of systematic as well as non-systematic review studies and research studies on the theoretical aspects of the benefits of GenAI in education. Thus, there exists a research gap in this significant area of academic research, which justifies the relevance of the present study.

3. Research Questions

Keeping in view the stated research problem and to fulfil the research aims, the following research questions were set to commence the current research:

RQ 1: What is the effect of AI tools intervention on undergraduate EFL learners' writing skill development?
RQ 2: What specific areas of writing skill development are impacted by AI intervention in learning writing?

4. Research Methodology

The present study was conducted employing mixed-methods research methodology. Mixed-methods methodology was chosen since the study involved dealing with numerical data which made sense only through quantitative analysis. Whereas qualitative method was employed to make meaning of the results obtained from quantitative investigation and to report the research outcomes in narrative format. The methods were also mixed wherever required to obtain a clear picture emerging from data analysis.

4.1 Research Participants

The research participants were undergraduate students majoring in the discipline of Health Sciences. They were learning English for one year as a pre-requisite for university study. The participants, sixty in number, all male, ranged in ages between 22 and 24. The participants were proficient in English as they were taught English in schools for 12 years. Their level of proficiency in English was intermediate (B1). In the beginning of the research, they could write only short essays with spelling, vocabulary, structure, and organization errors. The participants had the basic knowledge of AI tools as used in smartphone applications, but they hardly used the tools as aids in education, such as learning writing. Moreover, they had very little opportunities to write anything in English except in classroom activities. Table 1, given below, presents the vital statistics on research subjects.

Table 1: Participants' demographic statistics

Data Collection Instrument	Participants	Gender	N	Average Age	Proficiency level in English	Knowledge of AI tools
Pre-test	Experimental Group	Male	30	23	B1	Basic
	Control Group	Male	30	23	B1	Basic
Post-test	Experimental Group	Male	30	23	B1	Basic
	Control Group	Male	30	23	B1	Basic

4.2 Data Collection

4.2.1 Instruments

The instruments used to collect research related statistics were pre- and post-writing exams. For both exams, the research subjects were provided prompts to write one long essay on a given medical topic (minimum 300 words) and one data-based report. Each writing task carried 10 marks, 2 marks allotted to each element of writing - ideas, vocabulary, spelling, structure, and coherence. Both the groups were given the same topic prompts. The tests were piloted before being administered for validity and reliability. Expert suggestions were incorporated in the final drafts of tests. The writing tests administered to both experimental and control group participants were the same. In fact, both groups took the tests at the

same time and place. To ensure further test transparency, the grading rubrics were consistent across the groups.

4.3 Research Design

The present research was designed as a quasi-experimental study by establishing experimental and control groups to conduct the experiment to measure the effect of one variable upon the other, following pre- and post-exam measurements. The effect of independent variable on the dependent variable was measured by observing the difference in marks obtained by research subjects in pre- and post-exams. The subjects in the experimental batch were trained to use expert systems to hone their essay writing skill, while participants in the control group were used as base scale against which the cumulative effects were measured. The study variables were as follows:

4.3.1 Independent Variables

AI tools intervention in undergraduate writing development. This involved teaching students the ways to access and use AI tools helpful in their writing skill development, such as ResearchRabbit for referencing, Acrobat Chat with PDFs, and Grammarly for grammar check and planning, drafting and revising the essays. These AI tools were chosen especially for being easily accessible to all learners as well as for their ease of use.

4.3.2 Dependent Variables

Writing development, which was measurable in terms of appropriate referencing, summary skills, use of relevant medical terminology, correct syntax and punctuation, appropriate textual organization (introduction, discussion, conclusion), and correct word spellings.

The variance in the scores the experimental batch subjects obtained in the post-test, calculated by comparing their marks with their pre-test marks and with the scores the control group participants obtained, was taken to be an indicator of the progress in their learning impacted by AI tools intervention in their writing skill development. The significance of the difference was established by statistical analysis.

4.4 Research Procedure

The present research involved experimental teaching conducted for 4 weeks. The research participants constituted 2 full classes of undergraduate students, 30 students in each class, taught English by the researcher. After the pre-test, one class of 30 students was designated as 'experimental group' participants, while the other class of equal number of students was called 'control group' participants. The training of experimental batch subjects involved using AI tools to search relevant medical terminology, reference materials for ideas and to prepare a short bibliography, to check correct grammar and usage, and organize, plan, draft and revise their essays. For instance, the participants were taught to prepare a list of relevant references using ResearchRabbit, a screenshot of which is given below as Figure 1.

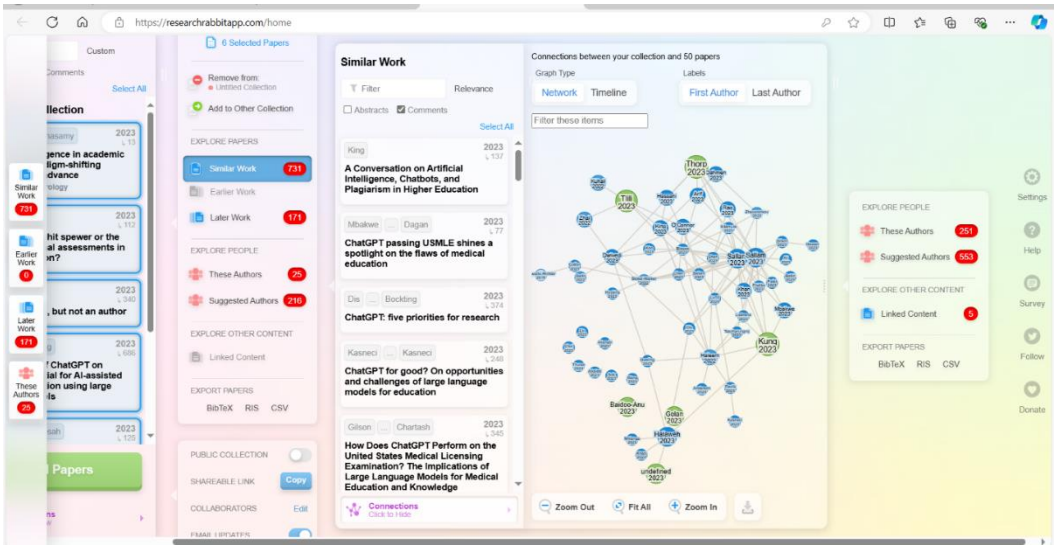

Figure 1: A screenshot of researchrabbit.ai page

Whereas the control group participants were taught to write essays on similar topics using traditional approach, that is, finding the terms using a dictionary (online or offline), correcting grammar errors on their own, finding relevant ideas from online/offline sources, and planning, drafting and writing the essay. After the experimental teaching, the two groups were given a post-test. The scores obtained by research subjects were compared: (i) post-test scores of experimental and control groups with their pre-test scores, (ii) post-test scores of the groups with each other. The obtained results were statistically analysed to measure the significance of difference.

5. Results

Table 2, given below, presents the raw scores (out of a total of 20 marks) obtained by research subjects in two tests.

Table 2: Scores obtained by research subjects in the tests

Participant	Pre-test		Post-test	
	Experimental Batch	Control Batch	Experimental Batch	Control Batch
1	12	10	18	12
2	10	12	16	14
3	8	13	16	16
4	10	12	17	14
5	12	12	18	14
6	13	8	17	13
7	12	8	17	14
8	14	9	18	13
9	14	10	18	14
10	8	11	16	15
11	9	12	15	14
12	11	11	17	14
13	11	13	18	15
14	12	14	17	16
15	12	9	19	12

16	11	10	18	12
17	14	15	16	16
18	15	14	18	17
19	8	12	16	14
20	12	10	18	12
21	9	12	15	14
22	10	15	17	17
23	10	13	18	15
24	12	12	19	14
25	14	13	16	15
26	8	11	16	14
27	12	12	18	14
28	14	14	18	16
29	13	9	17	12
30	12	10	18	14
Descriptive Statistics	Mean: 11.06 SD: 2.911 SE: 0.531 Var.: 4.04	Mean: 11.53 SD: 1.94 SE: 0.354 Var.: 3.64	Mean: 17.16 SD: 1.085 SE: 0.1980 Var.: 1.13	Mean: 14.2 SD: 1.42 SE: 0.259 Var.: 1.96

N = 30 + 30

The descriptive statistics of participants' scores obtained in two tests can be graphically represented as illustrated by Figures 2 and 3 below:

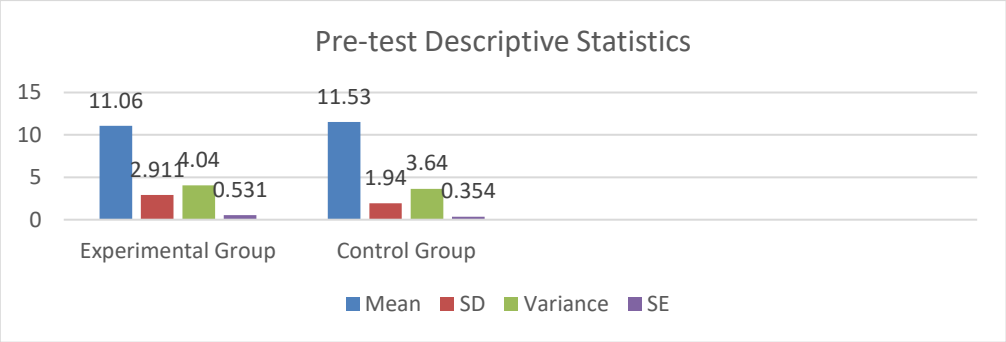

Figure 2: Graphical representation of descriptive statistics of participants' pre-test scores

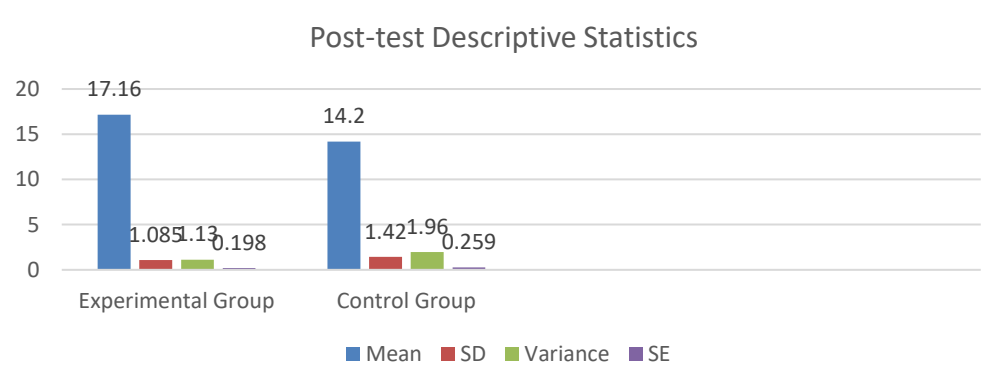

Figure 3: Graphical representation of descriptive statistics of participants' post-test scores

To calculate the means of participants' scores in two tests, the raw scores were subjected to statistical analysis. To find the significance of difference in the means of participants' scores, the obtained values (Mean, Standard Deviance, Variance, Standard Error) were used to calculate the t-test value. Table 3, given below, presents the paired-sample t-test values for experimental and control group participants.

Table 3: Paired-Sample t-test values for experimental and control groups

S. No.	Group	N	Pre-Test			Post-Test			t-test value	df (n-2)
			Mean	SD	Variance	Mean	SD	Variance		
1.	Experimental Group	30	11.06	2.911	4.04	17.16	1.085	1.13	4.743*	28
2.	Control Group	30	11.53	1.941	3.64	14.21	1.420	1.96	5.13**	28
3.	Experimental vs. Control Group (pre-test)	60							0.484†	58
4.	Experimental vs. Control Group (post-test)	60							8.83‡	58

* Significant at p\.05
** Significant at p\.05
† Not significant at p\.05
‡ Significant at p\.05

6. Discussion

A glimpse at Table 3, above, shows that after the experimental teaching research subjects in the experimental batch have outperformed their control batch peers. Prior to the experimental teaching, all the participants stood at the same level of proficiency in writing in English. The means of pre-test marks of participants calculated after the groups were formed show no significant difference from each other. The mean of pre-test marks obtained by control group participants stands at 11.53; while that of the experimental group participants, it is 11.06, a difference of merely 0.47 points. Subsequently, a comparative analysis of the mean pre-assessment scores of the groups obtains t-test value 0.484 (statistically not significant at p .05).

Whereas the difference between the means of their post-assessment scores is statistically significant. The mean of the marks obtained by the experimental group participants is 17.16; while that of the control group participants, it is 14.21, a difference of 2.95 points. Consequently, the t-test value in this case is 8.83 (statistically significant at p .05). The t-test value obtained on a comparison of the pre-assessment and post-assessment scores of the experimental group participants is 4.743 (statistically significant at p .05). The t-test value obtained on

a comparison of the pre-assessment/post-assessment scores of the control group participants is 5.13 (which is also statistically significant at p .05).

The numerical values cited above can be interpreted to mean that participants in both groups have made progress in developing their writing skill, but the experimental batch subjects have made greater progress in comparison to what control batch subjects have achieved. This difference in the progress made by the experimental group participants can be safely attributed to the fact that expert systems tools were integrated into their writing instruction materials since all other study variables have been the same for both groups and there was no other factor that could affect and enhance their writing skill otherwise. These findings are significant since, as more future studies on the topic are expected to corroborate the present findings, a generalizable theoretical foundation on the impact of AI tools intervention on EFL learners' writing development can be formulated which will pave way for the preparation and institutionalization of further guidelines.

7. Conclusion

The present research was designed to investigate whether AI tools intervention affects the undergraduate EFL learners' writing skill development. The statistical analysis carried out on the research data showed that the intervention does affect learners' writing skill development in a positive way, making them better English essay writers. The research was also focused on knowing what specific areas of writing skill development are impacted by AI intervention. The results indicate that GenAI tools help learners understand the topics of essays, gather ideas on the topic from online/offline resources, prepare a list of relevant references and plan, organize and prepare the final drafts of their essays, apart from assisting learners in grammar and spell-check. Thus, the primary as well as the secondary objectives of the study have been achieved since the current research findings have made significant addition to the growth of reference material on the effects of expert systems tools on EFL learners' writing skill development.

The results and findings from the present study could not be contrasted with findings from previous studies for a comparative insight at a larger scale since at present empirical, classroom-based investigatory studies on the topic are very scanty. Alneyadi and Wardat (2023) conducted a study with eleventh graders in a UAE school in the field of electronic magnetism. Their findings square well with the results obtained from the current study as regards the use of ChatGPT as the researchers report that the GenAI model positively influenced student achievement and perception of learning. However, since the study was not focused on teaching writing in EFL, the insights from the study are not highly relevant to the present research. Alshater's (2022) findings concerning the favourable effect of ChatGPT on students' institutional performance in economics and finance also corroborate the findings from the present study. However, the most relevant findings are reported by Wale and Kassahun (2024) who studied the impacts of GenAI tools - Writerly and Google Docs - integrated into EFL teaching. Their findings support the results obtained from the current study, that

is, expert systems tools integrated into study plans significantly improved EFL writing instruction.

8. Limitations of the Present Study

For various reasons, the present study has its own limitations despite the researcher's best efforts to achieve the stated aims of the research. The first research limitation was that there were not enough empirical research studies on the topic to see the findings of the present study from a comparative and contrastive perspective. The second research limitation was that the investigator could not have any mechanism in place to check whether the control group participants also took help from AI tools to learn writing since GenAI tools are freely and readily available on mobile devices and the present generation of students are aware of their existence and know what can be achieved with the help of those tools. The possibility of the control group participants using AI tools to learn English writing would affect the results. Moreover, the findings from the present study may not be applicable to other contexts as the sample size in the present study was rather small, and more research on the topic is required before the findings can be generalized.

9. Further Recommendations

Based on the limitations of the present study and the present-day relevance of the research area, future researchers may focus on (i) examining the effects of AI tools intervention on other aspects of EFL learning, such as reading and listening, (ii) checking whether the GenAI intervention affects the writing development among male and female students equally since the present study was conducted using only male students as research participants, and (iii) investigating whether AI tools intervention brings about a long-lasting cognitive development among learners.

10. References

Al-Mohanna, A. D. (2024). Difficulties and challenges encountered by Saudi EFL learners: A diagnostic study. *Scholars International Journal of Linguistics and Literature*, *7*(10), 288-299. https://doi.org/10.36348/sijll.2024.v07i10.002

Alneyadi, S., & Wardat, Y. (2023). ChatGPT: Revolutionizing student achievement in the electronic magnetism unit for eleventh-grade students in Emirates schools. *Contemporary Educational Technology, 15*(4), ep448. https://doi.org/10.30935/cedtech/13417

Alotaibi, K. M. (2022). *Reading skills difficulties among EFL learners in Saudi Arabia* (ED622161). ERIC. https://files.eric.ed.gov/fulltext/ED622161.pdf

Alshater, M. (2022). Exploring the role of artificial intelligence in enhancing academic performance: A case study of ChatGPT. *SSRN*. http://dx.doi.org/10.2139/ssrn.4312358

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of Generative Artificial Intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI, 7*(1), 52-62. https://doi.org/10.61969/jai.1337500

Burke, G., & Schellmann, H. (2024, October 26). *Researchers say an AI-powered transcription tool used in hospitals invents things no one ever said*. The Associated Press. https://apnews.com/article/ai-artificial-intelligence-health-business-90020cdf5fa16c79ca2e5b6c4c9bbb14

Clay, G. (2025). AutomatED: Teaching Better with Tech https://automatedteach.com

Dergaa, I., Chamari, K., Zmijewski, P., & Ben Saad, H. (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

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., … Wright, R. (2023). Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, *71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Edmonds, D. (Ed.) (2024). *AI Morality*. Oxford University Press. https://doi.org/10.1093/oso/9780198876434.001.0001

Elsen-Rooney, M. (2023). *NYC education department blocks ChatGPT on school devices, networks*. Chalkbeat. https://www.chalkbeat.org/newyork/2023/1/3/23537987/nyc-schools-ban-chatgpt-writing-artificial-intelligence/

Farrelly, T., & Baker, N. (2023). Generative artificial intelligence: Implications and considerations for higher education practice. *Education Sciences, 13*(11), 1109. https://doi.org/10.3390/educsci13111109

Gayed, J. M., Carlon, M. K. J., Oriola, A. M., & Cross, J. S. (2022). Exploring an AI-based writing Assistant's impact on English language learners. *Computers and Education: Artificial Intelligence*, *3*, 100055. https://doi.org/10.1016/j.caeai.2022.100055

Giglio, A. D., & Costa, M. (2023). The use of artificial intelligence to improve the scientific writing of non-native English speakers. *Revista da Associacao Medica Brasileira*, *69* (9), e20230560. https://doi.org/10.1590/1806-9282.20230560

Golan, R., Reddy, R., Muthigi, A., & Ramasamy, R. (2023). Artificial intelligence in academic writing: A paradigm-shifting technological advance. *Nature Reviews Urology*, *20*, 327–328. https://doi.org/10.1038/s41585-023-00746-x

Guan, L., Li, S., & Gu, M. M. (2024). AI in informal digital English learning: A meta-analysis of its effectiveness on proficiency, motivation, and self-regulation. *Computers and Education: Artificial Intelligence*, *7*, 100323. https://doi.org/10.1016/j.caeai.2024.100323

Herft, A. (2023). *A teacher's prompt guide to ChatGPT*. https://acesse.one/herfteducator

Jarrah, A. M., Wardat, Y., & Fidalgo, P. (2023). Using ChatGPT in academic writing is (not) a form of plagiarism: What does the literature say?. *Online Journal of Communication and Media Technologies, 13*(4), e202346. https://doi.org/10.30935/ojcmt/13572

Jingxin, G., & Razali, A. B. (2020). Tapping the potential of Pigai automated writing evaluation (AWE) program to give feedback on EFL writing. *Universal Journal of Educational Research*, *8*(12B), 8334–8343. https://doi.org/10.13189/ujer.2020.082638

Kacena, M. A., Plotkin, L. I., & Fehrenbacher, J. C. (2024). The use of artificial intelligence in writing scientific review articles. *Current Osteoporosis Reports, 22*, 115–121. https://doi.org/10.1007/s11914-023-00852-0

Khalifa, M., & Albadawy, M. (2024). Using artificial intelligence in academic writing and research: An essential productivity tool. *Computer Methods and Programs in Biomedicine Update*, *5*, 100145. https://doi.org/10.1016/j.cmpbup.2024.100145

Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., Maningo, J., & Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large

language models. *PLOS Digit Health*, *2*(2), e0000198. https://doi.org/10.1371/journal.pdig.0000198

Leuenberger, M. (2024). *AI 'can stunt the skills necessary for independent self-creation': Relying on algorithms could reshape your entire identity without you realizing*. LIVESCIENCE. https://www.livescience.com/technology/artificial-intelligence/ai-can-stunt-the-skills-necessary-for-independent-self-creation-relying-on-algorithms-could-reshape-your-entire-identity-without-you-realizing

Májovský, M., Černý, M., Kasal, M., Komarc, M., & Netuka, D. (2023). Artificial intelligence can generate fraudulent but authentic-looking scientific medical articles: Pandora's box has been opened. *Journal of Medical Internet Research*, *25*, e46924. http://doi.org/10.2196/46924

Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*, *10*(2), 2236469. https://doi.org/10.1080/2331186X.2023.2236469

Morris, E. J. (2018). Academic integrity matters: Five considerations for addressing contract cheating. *International Journal for Educational Integrity*, *14*, 15. https://doi.org/10.1007/s40979-018-0038-5

Piaget, J. (1973). *To understand is to invent: The future of education*. Grossman Publishers. https://unesdoc.unesco.org/ark:/48223/pf0000006133

Son, J.-B., Ružić, N. K., & Philpott, A. (2023). Artificial intelligence technologies and applications for language learning and teaching. *Journal of China Computer-Assisted Language Learning*. https://doi.org/10.1515/jccall-2023-0015

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press. https://home.fau.edu/musgrove/web/vygotsky1978.pdf

Wale, B. D., & Kassahun, Y. F. (2024). The transformative power of AI writing technologies: Enhancing EFL writing instruction through the integrative use of writerly and google docs. *Human Behavior and Emerging Technologies*, 9221377. https://doi.org/10.1155/2024/9221377

Werdiningsih, I., Marzuki, & Rusdin, D. (2024). Balancing AI and authenticity: EFL students' experiences with ChatGPT in academic writing. *Cogent Arts & Humanities*, *11*(1), 2392388. https://doi.org/10.1080/23311983.2024.2392388

Williams, C. (2023). Hype, or the future of learning and teaching? 3 Limits to AI's ability to write student essays. *London School of Economics internet blog*. https://kar.kent.ac.uk/99505/

Wilson, J. R. (2022). *Academic Writing*. https://wilson.fas.harvard.edu/AcademicWriting

Xu, T., & Wang, H. (2024). The effectiveness of artificial intelligence on English language learning achievement. *System*, *125*, 103428. https://doi.org/10.1016/j.system.2024.103428

Yuan, L., & Liu, X. (2025). The effect of artificial intelligence tools on EFL learners' engagement, enjoyment, and motivation. *Computers in Human Behavior*, *162*, 108474. https://doi.org/10.1016/j.chb.2024.108474