

<https://doi.org/10.69639/arandu.v13i3.2485>

The Impact of LuzIA on Academic Writing in English as Foreign Language Learners

El impacto de LuzIA en la escritura académica de estudiantes de inglés como lengua extranjera

Eleyvi Castellón Vladés

ecastellonv@ube.edu.ec

<https://orcid.org/0009-0003-0481-6111>

Universidad Bolivariana del Ecuador
Durán – Ecuador

Jaime David Sáenz Herdoiza

jdsaenzh@ube.edu.ec

<https://orcid.org/0009-0003-7317-8563>

Universidad Bolivariana del Ecuador
Durán – Ecuador

Josué Reinaldo Bonilla Tenesaca

jrbonillat@ube.edu.ec

<https://orcid.org/0000-0002-6748-2345>

Universidad Bolivariana del Ecuador
Durán – Ecuador

Jacqueline Elizabeth López López

jlopezl@ube.edu.ec

<https://orcid.org/0000-0002-1765-8103>

Universidad Bolivariana del Ecuador
Durán – Ecuador

Artículo recibido: 10 julio 2026- Aceptado para publicación: 16 agosto 2026
Conflictos de intereses: Ninguno que declarar.

ABSTRACT

Academic writing is one of the most difficult skills for English as a Foreign Language (EFL) learners to develop because it requires not only linguistic competence but also the ability to organize ideas and achieve coherence. This challenge is particularly relevant in multilingual environments, where learners must manage these demands simultaneously. This study analyzes the influence of LuzIA, a GPT-powered conversational artificial intelligence assistant integrated into WhatsApp, on academic writing, writing anxiety, and learner autonomy among EFL students at a bilingual secondary school in Ecuador. A sequential explanatory mixed-methods design was employed with 145 students from six intact classroom cohorts in Grades 8–12, whose expected English proficiency ranged from CEFR B2 to C1 according to their curricular placement. LuzIA was used for 12 weeks as a writing support tool in three English courses within the curriculum. Writing performance was assessed using the analytic rubric developed by Jacobs et al. (1981), which includes the higher-order components of Content and Organization and the lower-order

components of Vocabulary, Grammar, and Mechanics. Writing anxiety (SLWAI), learner autonomy, and the Technology Acceptance Model (TAM) were also assessed through an 18-item questionnaire. The results showed a statistically significant increase in writing performance ($\Delta M = 16.4$ points; $t(144) = 14.23, p < .001, d = 0.90$). The largest gains were observed in Grammar ($d = 1.12$) and Vocabulary ($d = 1.05$), while significant but smaller gains were found in Content ($d = 0.72$) and Organization ($d = 0.68$). Writing anxiety decreased significantly ($\Delta M = -1.24; d = 1.35$), whereas learner autonomy increased ($\Delta M = +0.87; d = 0.92$). TAM results were also positive, with Perceived Ease of Use reaching 4.47/5 and Perceived Usefulness reaching 4.12/5. The study concludes that the use of LuzIA may provide benefits for EFL academic writing when it is used as a support tool rather than as a replacement for learners' own writing. The strongest changes were observed in language-level aspects of writing. In addition, accessing the tool directly through WhatsApp may facilitate its implementation.

Keywords: LuzIA, EFL academic writing, GPT conversational AI, writing anxiety, learner autonomy

RESUMEN

La escritura académica es la habilidad más difícil de desarrollar para los estudiantes de Inglés como Lengua Extranjera (EFL) debido a la competencia lingüística requerida pero también por la necesidad de organizar los pensamientos y lograr la coherencia. Este problema se acentúa especialmente en un entorno multilingüe donde se juntan todos estos requerimientos. Este trabajo analiza la influencia de LuzIA, un asistente de inteligencia artificial conversacional basado en la tecnología GPT integrado en la aplicación WhatsApp en la escritura académica, la ansiedad por la escritura y la autonomía del estudiante de EFL de una escuela secundaria bilingüe de Ecuador. En esta investigación se aplicó un diseño explicativo secuencial mixto y se seleccionaron 145 estudiantes de seis grupos completos de 8° a 12° grado que supuestamente tienen un nivel de inglés entre B2 y C1 del MCER dependiendo de su ubicación curricular. LuzIA se utilizó durante 12 semanas como un ayudante en el proceso de escritura en tres cursos de inglés del plan de estudios. Los resultados de la escritura fueron evaluados mediante la rúbrica analítica de Jacobs et al. (1981) la cual incluye componentes de alto nivel de contenido y organización así como elementos de bajo nivel de vocabulario, gramática y mecánica. Además se evaluó la ansiedad por la escritura (SLWAI), la autonomía del estudiante y el Modelo de Aceptación de la Tecnología (TAM) mediante un cuestionario de 18 ítems. Como resultado se encontró un aumento estadísticamente significativo en el desempeño de la escritura ($\Delta M = 16.4$ puntos; $t(144) = 14.23, p < .001, d = 0.90$). Los mayores aumentos se observaron en la gramática ($d = 1.12$) y el vocabulario ($d = 1.05$). Hubo ganancias significativas pero algo menores en el contenido ($d = 0.72$) y la organización ($d = 0.68$). Aunque hubo tal mejora la ansiedad por la escritura disminuyó significativamente ($\Delta M = -1.24; d = 1.35$) y la autonomía del estudiante aumentó positivamente

($\Delta M = +0.87$; $d = 0.92$). Los resultados del TAM también fueron positivos: la puntuación de usabilidad fue 4.47/5 y la puntuación de utilidad percibida fue 4.12/5. Concluyendo este estudio se puede afirmar que el uso de LuzIA puede proporcionar ciertos beneficios a la escritura académica en EFL pero solo si no se utiliza como un reemplazo sino como un ayudante en la búsqueda de algún objetivo. Los cambios más significativos se hicieron en la parte lingüística de la escritura. Además el hecho de usar el programa directamente en la aplicación WhatsApp hace más fácil su implementación.

Palabras Clave: LuzIA, escritura académica EFL, IA conversacional GPT, ansiedad por la escritura, autonomía del estudiante

Todo el contenido de la Revista Científica Internacional Arandu UTIC publicado en este sitio está disponible bajo licencia Creative Commons Attribution 4.0 International. 

INTRODUCTION

English academic writing acts as the gateway for university entrance and worldwide certification, and the learning of the skill among students whose native tongue is not English (English as a Foreign Language students) becomes extremely complicated. Academic writing requires learners to manage multiple demands simultaneously, including grammatical accuracy, appropriate vocabulary, logical organization of ideas, genre conventions, and academic style. For EFL learners, these demands may be particularly challenging because they are developing academic writing skills while continuing to develop proficiency in the target language (Ghafar & Raheem, 2025). There is one more challenge in the learning process of English academic writing. Providing continuous individualized feedback in conventional classroom settings may be difficult because of practical constraints such as limited instructional time (Sanchez, 2024).

Fortunately, one more tool for language development appeared due to the rapid progress in artificial intelligence. AI-powered LLMs are capable of giving instant feedback on grammar, vocabulary, organization, and other writing-related issues without any necessity of teacher intervention during all steps. Recent research has provided growing evidence that AI-assisted tools can support EFL writing development. Experimental and comparative studies have reported improvements in different dimensions of writing following AI-assisted instruction or revision, while systematic reviews have identified broader patterns of linguistic and cognitive scaffolding across EFL contexts (Song & Song, 2023; Tsai et al., 2024; Li & Wilson, 2025). In particular, AI-supported writing appears to offer useful assistance with language-related aspects such as grammar and vocabulary, while also providing opportunities for revision, idea development, and organization.

There are many different AI-assistive tools. One of the unique features of the LuzIA tool is the fact that it can work in WhatsApp. It becomes important because of the fact that WhatsApp is integrated in daily practices of many Latin American teenagers. As a result, it will be possible to receive AI help using the application with which they are already familiar. LuzIA may reduce some barriers to AI access because it can be used through familiar digital environments such as WhatsApp, potentially reducing the need for students to adopt additional specialized platforms. The tool also provides language-related functions, including writing assistance, translation, grammatical correction, and explanations, which may support its use in EFL learning contexts.

However, accessibility does not automatically lead to educational effectiveness. Recent studies have reported promising applications of AI-based conversational tools in EFL speaking development, including improvements in oral performance, pronunciation, confidence, willingness to communicate, and reduced speaking anxiety (Isabella et al., 2026; Siagian et al., 2026). Nevertheless, these studies examined DOLA AI rather than LuzIA and focused primarily on oral language development. To date, there appears to be limited empirical evidence regarding

the use of LuzIA to support EFL academic writing. Consequently, it remains unclear whether LuzIA can support not only lower-order concerns, such as grammar and vocabulary, but also higher-order concerns, including content development and the organization of ideas.

To address this research gap, the present study examines the use of LuzIA in a multilingual secondary-school EFL context in Ecuador. The study evaluates multiple dimensions of the writing process, such as grammar, vocabulary, organization and ideation of the texts along with writing anxiety and learner autonomy as important writing-related factors.

Thus, there are three contributions that this study makes to science. First, it focuses precisely on the use of LuzIA as the tool for improving academic writing. Second, it includes the evaluation of AI assistance in Latin American secondary education, which is the context that is relatively neglected in most of the existing research. Finally, the study evaluates the effect of AI assistance on higher order and lower order concerns in order to understand what type of help can be provided efficiently.

Research Objective

The main objective of this study was to evaluate the influence of LuzIA-assisted writing support on EFL learners' academic writing performance, writing anxiety, and learner autonomy in a multilingual secondary-school context in Ecuador, with particular attention to changes in Higher-Order Concerns (HOCs) and Lower-Order Concerns (LOCs) in students' academic writing.

Literature review

Learning English as a Foreign Language (EFL)

English as a Foreign Language (EFL) generally refers to the teaching and learning of English in contexts where English is not the primary language of everyday communication and is mainly learned through formal educational settings. It creates specific challenges for the learners because the vast majority of all the experience acquired by EFL learners happens inside the classroom. Thus, the amount of natural exposure that they may receive outside the classroom setting might be minimal, and, as a result, the opportunity to get the necessary meaningful input decreases significantly (Krashen, 1982). The same concerns writing because without getting the exposure to academic English outside the classroom writing development becomes even harder.

However, in the case of Ecuador, things are even a little bit more complicated. English is a mandatory foreign language throughout primary and secondary education, while Spanish serves as a prevalent national language there. German, Kichwa, and other tongues are spoken in the specific communities. In the private educational establishment where this research will be conducted students learn English along with German, and therefore, the multi-language environment is formed. For the learners, it requires managing several different systems simultaneously, what increases the metacognitive and cognitive load for them during the learning process (Ayala-Pazmiño & Alvarado-Lucas, 2023).

The complexity of the described situation becomes most prominent in secondary education because by the eleventh grade students prepare for Cambridge B2 First, with high-performing candidates potentially demonstrating performance at C1 on the Cambridge English Scale. In addition, writing plays a crucial role during such examinations since learners need to not only create grammatically correct sentences but also develop ideas, structure them logically, use the necessary vocabulary, and adapt the created texts depending on the purpose. These competencies are just the one that it may be quite difficult to develop in the framework of conventional class learning.

However, it seems that similar issues with the development of writing happen among EFL students in other Latin American countries. Research has identified academic writing as a particularly challenging area for EFL learners. These difficulties are associated with the need to manage several aspects of writing simultaneously, including grammatical accuracy, vocabulary use, organization, coherence, and the development of ideas. Limited opportunities for writing practice and individualized feedback may further increase these challenges (Ghafar & Raheem, 2025).

Thus, from this perspective, it makes sense to use AI to develop learners' writing skills in the classroom. The main benefit of such a tool is that it allows getting an immediate and repeated feedback during the text creation process. Using it in this way can add additional opportunities for revising grammatical structures, selecting different vocabulary and developing ideas since it is not possible to do so before the teacher's feedback is provided.

Academic Writing in the EFL Context

In contrast, academic writing entails skills beyond generating correct sentences. As EFL writers think about grammar and vocabulary, they also organize their thoughts, make connections, and meet requirements characteristic of argumentative or expository styles. Coordinating all these processes places considerable demands on working memory, as students are still far from acquiring a high level of proficiency in English (Ghafar & Raheem, 2025). The issues associated with writing can be divided into two groups called Lower-Order Concerns (LOCs) and Higher-Order Concerns (HOCs). While LOCs relate to grammar, spelling, punctuation, and vocabulary selection, HOCs entail such things as formulating a thesis statement, organizing one's arguments, and choosing language depending on purpose and audience (Jacobs et al., 1981).

This distinction is particularly relevant to understanding the role of AI in EFL writing instruction. Recent research suggests that AI-assisted writing can be particularly effective in supporting lower-order concerns (LOCs), including grammar and vocabulary, by providing learners with immediate language-focused feedback and opportunities for revision (Tsai et al., 2024). When it comes to higher-order concerns, on the other hand, such judgments require certain understanding of the text being written – something that is often lacking in computer programs.

For this reason, AI should be regarded as writing assistance rather than substitute of a person. It will allow helping students to identify and eliminate language problems, so instructors would have more time to focus on other important things such as arguments, organization of ideas, purposes of writing, and meanings conveyed through the text. In this way, two means of help might complement each other.

Another aspect of writing is emotional. Writing essays in a second language might cause feelings of insecurity, mainly due to the fact that every mistake will be taken into consideration. As identified by Cheng (2004) via Second Language Writing Anxiety Inventory, there are three kinds of it, namely somatic, cognitive, and avoidance anxiety. All these types of anxiety might make students unwilling to do any writing at all, avoid it, delay writing, etc.

The use of AI in learning environments, on the contrary, gives an opportunity to experiment with writing, to generate questions and get answers to them, revise the text, get feedback – without making it public at first. That way, the pressure caused by evaluating is somewhat reduced. Still, this reduction alone will not make students write better; it might motivate them to write more, so they would have more chances to practice.

Defining LuzIA: Design, Accessibility, and Educational Potential

LuzIA is a conversational artificial intelligence assistant designed to provide accessible AI-based support through different digital environments. The platform allows users to interact with the assistant through text or voice and can respond in the language used by the user. Importantly for the context of the present study, LuzIA can be accessed through WhatsApp, allowing students to interact with an AI assistant within a digital environment that may already be familiar to them (LuzIA, n.d.-a).

LuzIA is a general-purpose AI assistant rather than a tool developed specifically for EFL writing. Its documented functionalities include answering questions, assisting with writing and study-related tasks, translating texts, transcribing audio, reading documents, explaining concepts, and supporting language-learning activities (LuzIA, n.d.-a, n.d.-b). These functions make the platform potentially relevant to EFL learning; however, the availability of these features does not in itself demonstrate their effectiveness for developing academic writing.

In the present study, LuzIA was therefore used as a conversational scaffold rather than as a tool for generating complete texts on behalf of learners. Its use was organized around three areas related to the study's HOC–LOC framework. At the lower-order level, students used the assistant to obtain support with grammar, vocabulary, spelling, punctuation, and alternative wording. At the higher-order level, students could request guidance related to paragraph organization, idea development, thesis statements, and cohesion. Students remained responsible for evaluating the suggestions provided by the AI and deciding whether and how to incorporate them into their own writing.

This distinction is central to the present research because it allows LuzIA to be examined not simply as a text-generation tool but as a form of writing support. The study therefore evaluates both the extent to which AI-assisted interaction is associated with improvements in HOCs and LOCs and the areas in which learners may continue to require teacher guidance.

LuzIA AI in EFL Academic Writing

Nevertheless, despite some preliminary efforts to explore the effectiveness of using LuzIA in writing English as a Foreign Language (EFL), the relevant literature has yet to develop substantially. However, one can infer the potential benefits of the app from a number of empirical findings related to AI writing support in general. Some of them seem especially relevant to the theoretical framework of this research.

Song and Song (2023) examined the use of ChatGPT to support academic writing among Chinese EFL university students. In their mixed-methods study, 50 learners were assigned to either an AI-assisted group using ChatGPT or a control group receiving traditional writing instruction. The AI-assisted group demonstrated significantly stronger writing performance, including improvements in content, organization, grammar, and vocabulary. Their findings suggest that immediate AI-supported feedback and opportunities for revision may contribute to the development of EFL writing skills.

Marzuki et al. (2023) provide complementary qualitative evidence from EFL teachers' perspectives. Based on semi-structured interviews with four EFL teachers from three Indonesian universities, the study examined how teachers perceived the influence of AI writing tools on students' writing. The findings indicated perceived benefits for content development and organization, including the generation and structuring of ideas. However, because the study relied on teachers' perceptions rather than direct experimental measurement of students' writing gains, its findings should be interpreted as qualitative evidence of the perceived educational potential of AI writing tools.

Tsai et al. (2024) provide further evidence regarding the potential of ChatGPT-assisted revision in EFL writing. In their paired-comparison study, 44 EFL English majors submitted original essays and ChatGPT-assisted revised versions, which were independently evaluated by two graders. The ChatGPT-assisted versions received significantly higher writing scores, suggesting that AI-supported revision can improve assessed writing performance. However, the authors also raised concerns about the implications of AI-assisted revision for fairness and conventional approaches to writing assessment. This fact, obviously, makes this result especially relevant for the current study, as if AI is capable of quickly identifying issues in language, then similar effects may be expected from LuzIA when it is used as a writing scaffold.

Further support for the educational potential of AI-assisted writing can be found in Li and Wilson (2025), who conducted a systematic review of AI-integrated scaffolding among K–12 English language learners. Their findings suggest that AI tools can provide linguistic, cognitive,

and creative scaffolding by supporting processes such as grammar correction, vocabulary development, idea generation, and content organization. However, the authors also emphasize the importance of maintaining learner agency and avoiding excessive dependence on AI-generated assistance. These findings are particularly relevant to the present study because they suggest that AI can support different stages of the writing process while students remain responsible for evaluating and applying the assistance they receive.

Thus, taking into account the results discussed above, one can come to the conclusion that the usage of AI-assistant in writing usually provides useful, instant assistance during drafting and revision stages, leading to substantial improvements in language-level aspects.

Theoretical framework

In total, three theories complement each other and address different aspects of EFL writing under AI support – the development of learners who have guidance, gradual emergence of learner autonomy and reasons behind the continued usage of certain technologies.

One way to start analyzing the intervention and understand its underlying processes is using Vygotsky's Sociocultural Theory (1978) and specifically its aspect known as Zone of Proximal Development (ZPD). The ZPD concept means the difference between what learners are able to do independently and what they are able to do having the right help. Traditionally, this help comes from a teacher, a peer who is better at something or generally More Knowledgeable Other. In this situation, the role of a mediating tool may partially belong to LuzIA as it can analyze texts of students, find mistakes and give suggestions that may allow moving beyond the limits of their own capability.

It is very important to note the limit that should be taken into account in this case. The idea is not that the machine does the job, and the student takes the results. On the contrary, it is important to note that even when the AI provides certain help, the student remains responsible for the decision whether to accept its suggestions or reject them.

The theory by Holec (1981) of learner autonomy provides another point of view for this issue. In general, the theory claims that the learner takes responsibility for the process of his/her own learning. For example, the learner sets the aims, chooses strategies, monitors himself/herself, evaluates the outcome of the process. This is especially important in cases when learners deal with AI since the process of getting help only starts with the suggestion. Next, students need to decide what to do with it.

As a result, the usage of LuzIA may go beyond grammar/vocabulary corrections in certain moments since the learner compares options, analyzes suggestions, applies the changes and evaluates the options – all of this is metacognitive strategies of writing. Some of them may become an integral part of the strategy of students' revision. Initially, the process happens between a student and a machine but gradually some of those actions become a part of learners'

autonomous practice – identification of a problem, consideration of alternatives and evaluation of the final option.

Such an approach is natural in terms of sociocultural theory since the support that comes initially externally can gradually become internal over time. The idea is not the permanent dependency on a particular machine but rather learning the strategies to work independently.

Finally, TAM of Davis (1989) deals with another issue related to learners. It is about motivation behind the consistent use of the same technology in order to continue learning. Technology Acceptance Model focuses on the following aspects – Perceived Usefulness (PU) and Perceived Ease of Use (PEOU). The first one describes learners' opinion about the extent to which the technology helps to perform meaningful tasks. The latter one refers to the ease of use of the technology in question.

It is important to mention the aspect related to PEOU that the tool appears in the context of an already used application. It means that the learners do not have to spend additional time on learning how to use a new technology in order to be able to communicate with the machine. PEOU concerns learners' perceptions of how easy the technology is to access and use, whereas PU concerns their beliefs about its usefulness for improving writing performance.

MATERIALS AND METHODS

Research Design

The study employed a Sequential Explanatory Mixed-Methods Design consisting of an initial quantitative phase followed by a qualitative explanatory phase (Maforah & Leburu-Masigo, 2018). This design was selected considering that numbers alone cannot suffice in explanation of processes occurring as students incorporate AI into writing practice. Thus, the investigation takes place through two interrelated phases, with the qualitative phase assisting in the interpretation of the first phase quantitative findings.

During the first phase, a quantitative analysis is conducted. Writing scores before and after the intervention were compared using an analysis rubric alongside with composite Likert-scale questionnaire scores. The data was used to reveal whether any changes take place after the implementation of the LuzIA intervention and, most importantly, in what way these changes were statistically significant and in which direction.

At the second stage of the study, the experiences associated with the numeric data were analyzed in greater detail using semi-structured interviews and reflective journals as data collection instruments. This stage does not aim at verification of the previous findings but at more profound interpretation. More precisely, the aim is to explore students' experience related to changes that have taken place during the intervention.

The quasi-experimental design is used rather than a completely randomized one. As the research is performed with intact existing groups of students in one school, there is no possibility

of randomization of individual participants to distinct groups. Hence, the focus of statistical comparisons is within-participant comparisons between pre and post test. The initial difference in writing ability across grade levels is controlled by means of ANCOVA.

Thus, the quasi-experimental nature of the design limits the strength of the causal inferences that can be drawn. Nonetheless, it reflects the actual conditions under which educational intervention research is conducted, with pre-existing classes and student schedules already established. This approach is consistent with conducting educational research in authentic settings (Creswell & Creswell, 2023) and provides evidence of how AI-assisted writing support operates when implemented within an actual EFL instructional context.

Context and Participants

The study was carried out in a private school located in Quito within a bilingual secondary-school context offering a multilingual program that includes Spanish, German, English, and French. As part of its educational program, the school combines Ecuadorian and German educational systems, including the German Abitur program, resulting in substantial demands related to linguistic accuracy, critical thinking, and academic writing.

English language instruction begins at approximately age ten and becomes progressively more advanced throughout secondary education. By the upper secondary grades, students prepare for international English certification, with expected proficiency generally ranging from B2 to C1 according to the Common European Framework of Reference for Languages (CEFR). This setting is particularly relevant to the present study because academic writing forms part of both students' language development and their preparation for international certification.

The study included 145 EFL learners enrolled in six intact classroom cohorts ranging from 8th to 12th grade. All eligible students from these cohorts who met the study's participation requirements and for whom the required consent and assent had been obtained were included in the study. Therefore, rather than selecting a subset of students through stratified random sampling, the study included the accessible population across the participating cohorts. This approach ensured representation across grade levels and gender while preserving the naturally occurring classroom groups in which the intervention was implemented.

An a priori power analysis was used to assess whether the available number of participants would provide adequate statistical power for the planned analyses (Cohen, 1988). Assuming $\alpha = .05$ and a medium effect size ($d = 0.50$), the available sample of 145 participants exceeded the minimum requirement for 80% statistical power. Participants represented different stages of secondary education, with expected English proficiency ranging primarily from B2 to C1. Table 1 presents the distribution of participants across the six classroom cohorts.

Inclusion and Exclusion Criteria

Students were eligible for inclusion if they were enrolled in one of the six participating classroom cohorts, were actively attending the English courses involved in the study, and had

obtained written informed consent from a parent or legal guardian and provided their own assent to participate. Participants were also required to complete the pre-test and post-test writing assessments and participate in the intervention activities.

Students were excluded from the final analysis if parental or guardian consent or student assent was not obtained, if they did not complete either the pre-test or post-test assessment, or if they did not participate sufficiently in the intervention to provide complete data for analysis. No students were excluded on the basis of gender or expected CEFR proficiency level.

Table 1

Sample Distribution by Grade Level, Gender, and English-Medium Subjects

Grade / Level	Males	Females	n	Subjects in English	Expected CEFR Level
8th Grade EGB A	20	12	32	English; Social Studies	B2
9th Grade EGB A	8	10	18	English; Social Studies	B2
10th Grade EGB A	10	11	21	English; Social Studies; Reading	B2
1st Year BGU (Sciences)	12	7	19	English; Reading	B2–C1
3rd Year BGU A	14	16	30	English; Reading; Intl. Cert. Prep.	B2–C1
3rd Year BGU B	9	16	25	English; Reading; Intl. Cert. Prep.	B2–C1
Total	73	72	145*		

Note. N = 145. Participants represented six intact classroom cohorts and ranged in age from approximately 12 to 18 years. CEFR classifications represent expected curricular proficiency levels established by the school and were not independently assessed as part of the present study.

Instruments

Data collection was carried out using five devices that catered for both qualitative and quantitative aspects investigated in the study.

Pre-test Prompt

The pre-test consisted of the 60-minute argumentative writing test and followed the structure of the writing that is expected from students when doing Cambridge FCE. The task included writing of 140 to 190 words, discussing two different opinions about a sociological-educational problem, presenting one's point of view and organizing the writing in accordance with typical academic essay outline: introduction, body paragraphs and conclusion. The prompt specified: “Some people think that technology makes students independent learners. Others argue that technology makes students dependent on the external assistance. Discuss both sides and give

your opinion”. The selected topic was considered appropriate for the range of proficiency expected across the participating cohorts and allowed more advanced students to develop increasingly complex arguments. Students had no possibility to use LuzIA and any other AI tool.

Post-test Prompt

12 weeks later, the participants performed another similar argumentative writing task, also in the conditions: 60 minutes, between 140 to 190 words, same writing requirements as pre-test and a different topic to minimize influence of familiarity with the first prompt. As in the pre-test, participants completed the post-test without access to LuzIA or any other AI-assisted writing tool. The second prompt stated: “There are some educators who are in favor of using AI tools in language learning to improve students’ results. There are others who say that AI hinders the development of independent language skills. Discuss both sides and give your opinion.” Just like before, this test included considering different viewpoints, stating and defending one's opinion on the issue. However, unlike the pre-test prompt, this task was specifically connected to their experiences during the previous 12 weeks, during which they engaged in using LuzIA tool, so they had a tangible starting point for discussion of the advantages and limitations of the tool in language learning.

Pilot testing

Prior to conducting data collection, both prompts underwent pilot testing during Week 1 among 12 students of similar proficiency level who did not take part in the current study. The aim of the pilot testing was to ensure the comparability of the two tasks, their performability in the allotted time frame and their interpretability for scoring.

Analytic Scoring Rubric

The analysis of the written work of students was carried out through an analytic rubric based on the ESL Composition Profile created by Jacobs et al. (1981). As opposed to assigning one rating based on one’s general impression of the writing performance, the rubric offered a possibility to analyze each essay according to five specific dimensions based on HOC and LOC framework used in this study.

Content and Organization were the HOC dimensions, for which 30 and 20 points were assigned, respectively. Vocabulary, Language Use/Grammar and Mechanics were LOC dimensions, with 20, 25 and 5 points. Thus, the total writing score could reach the value of 100 points when combining the results for all the five dimensions.

Each of the dimensions was rated for four performance bands – Excellent, Very Good, Good and Very Poor. The behavioral descriptors that corresponded to each band were formulated and served to guide rater during the evaluation.

In order to evaluate reliability of rater judgments, two trained raters analyzed a stratified random sample including 30% of pre- and post-test essays. Inter-rater reliability was computed using Cohen’s weighted kappa (κ), with the pre-established minimum acceptable level $\kappa = .70$.

The disagreements in case of the difference in at least one performance band for any particular dimension between raters were discussed until consensus was reached.

Finally, the inter-rater reliability reached the value of $\kappa = .83$ and exceeded the required level.

Table 2

Analytic Scoring Rubric: HOC and LOC Domains, Maximum Points, and Constructs Measured

Category	Domain	Max. Points	Construct Measured
HOCs (50 pts)	Content	30	Thesis development, topic relevance, depth of argumentation
	Organization	20	Logical sequencing, cohesion, transition deployment
LOCs (50 pts)	Vocabulary	20	Lexical range, word choice accuracy, academic register
	Language Use / Grammar	25	Morphosyntactic accuracy, tense, agreement, complexity
	Mechanics	5	Spelling, punctuation, capitalization, paragraphing
Total Score		100	Overall academic writing proficiency

Notes. HOCs: Higher Order Concerns; LOCs: Lower Order Concerns. All domains were rated on four performance bands with behavioral descriptions.

Student Perception Questionnaire

An 18-item questionnaire using a five-point Likert scale (1 = Strongly Disagree to 5 = Strongly Agree) was used to examine three variables related to students' experiences throughout the intervention; these are writing anxiety, learner autonomy, and technology acceptance.

The first part of the questionnaire consisted of items (1-5), which focused on writing anxiety, based on the Second Language Writing Anxiety Inventory (SLWAI) by Cheng (2004). The second section (items 6-10) examined learner autonomy based on Holec (1981)'s conceptualization, focusing on observable learning actions rather than seeing autonomy only as an abstract construct (for example, conducting an individual brainstorm, conducting revision independently, actively searching for support).

Third section consisted of items 11-18 and focused on TAM developed by Davis (1989); this model examines two factors, namely Perceived Usefulness (PU) and Perceived Ease of Use (PEOU).

The writing anxiety and learner autonomy sections of the questionnaire were administered twice, at pre-intervention and post-intervention, whereas the TAM section assessing Perceived

Usefulness (PU) and Perceived Ease of Use (PEOU) was administered only at post-intervention, after students had experienced using LuzIA. Reliability was measured using Cronbach's Alpha, where value above $\alpha \geq .70$ was considered to be adequate (Nunnally, 1978). Reliability of all sections was satisfactory – SLWAI $\alpha = .82$, Learner Autonomy $\alpha = .79$, PU $\alpha = .88$, and PEOU $\alpha = .91$.

Technology Acceptance Model (TAM) Perceptions of LuzIA

The questions included in the TAM portion were designed to measure students' views regarding LuzIA as a support for their writing tasks. In particular, four questions measured PU, measuring whether LuzIA helped improve writing, identify grammar mistakes, expand their vocabulary, and help improve their organizational skills.

Four more questions focused on PEOU and evaluated whether students considered LuzIA easy to reach through WhatsApp, user-friendly, not complicated by unnecessary technicalities, and quickly attainable. This adaptation is consistent with the Technology Acceptance Model (Davis, 1989) and with recent research on AI-supported language learning through chatbot-based environments (Behforouz & Al Ghaithi, 2024).

Specifically, WhatsApp seems particularly relevant in the study as students did not have to master another platform before starting working with AI. Familiar environment seemed to be expected to result in improved PEOU.

EFA demonstrated the expected two-factor model of PU and PEOU. All factors were loaded on the appropriate scale at $\lambda \geq .62$, with no cross-loading being equal to or exceeding .30.

Data Collection Procedure

For each stage, data were collected for the entire twelve-weeks period. In Week 1-2 (Stage 1), the pretest was taken with all the essays done under uniform test conditions at the school computer lab without LuzIA or any other AI access. Immediately after that, participants took the baseline questionnaire. Participants then took part in two guided workshops on the use of LuzIA where topics of formulating purposeful prompts, evaluating AI recommendations, deciding whether to accept or reject AI suggestions, recognizing possible errors or hallucination by AI, as well as the limits of academic integrity were addressed. Right from the very beginning, it was made clear that the aim of the use of LuzIA is to support the process of students' writing but not to do the writing for them. Participants were given an alphanumeric code and were made aware of the voluntary and confidential nature of their participation in the research.

In Weeks 3-10 (Stage 2), the tool started to be incorporated into students' writing process in three subjects which were Academic Writing, Grammar & Composition and Advanced Writing. In each of these disciplines, it had a different purpose – developing arguments in the essay, grammatical correction, vocabulary choosing and more sophisticated writing tasks. Students used LuzIA both during in-class writing activities and outside school in individual writing practice. The frequency, duration of the work with the tool and the primary task performed

(either grammatical correction, finding vocabulary, structural support or translation) were recorded in WhatsApp interaction logs. Besides, students kept weekly reflective journals containing information on the use of LuzIA and attempts at writing.

Finally, in Week 11-12 (Stage 3), the post-test was taken under the same standardized conditions and then followed by the post intervention questionnaire. Selected 20 students were interviewed in a semi-structured manner. A maximum variation sampling approach made sure that interviews included participants of different grades, genders, and proficiency bands.

Data Analysis

Quantitative data analysis entailed five stages. First, descriptive statistics (means, SD, medians, ranges, etc.) for overall writing scores, all rubric dimensions, and all questionnaire constructs were calculated. When sample sizes allowed it, comparisons between subgroups were performed in accordance with their grade level and gender. Tests for normality (Shapiro-Wilk tests) were run before selecting the inferential methods.

Second, paired-samples t-tests or Wilcoxon signed-rank test was selected for comparing pretest/posttest performance based on the rubric total score as well as five writing dimensions. As five different writing rubrics were compared, the significance threshold was adjusted (Bonferroni procedure, $\alpha=.01$). Significant p-value only tells about the presence or absence of difference. Therefore, Cohen's d effect size was calculated as M_{diff}/SD_{diff} for all of them according to the criteria suggested by Plonsky and Oswald (2014) for EFL studies: small ($d = .40$), medium ($d = .70$), and large ($d = 1.00$).

Third, in order to establish whether post-test writing performance differed among the six classroom cohorts, a one-way ANCOVA was performed while controlling for initial writing performance. Post-test total score was entered as the dependent variable, classroom cohort as the between-subjects factor, and pre-test total score as the covariate. Homogeneity of regression slopes and normality of residuals were checked in advance.

Fourth, Pearson correlation analyses were conducted to examine the relationships between students' perceptions of LuzIA and the intervention outcomes. Specifically, the association between Perceived Usefulness (PU) and writing-score improvement was examined, as well as the relationship between Perceived Ease of Use (PEOU) and PU.

Fifth, changes in learners' writing anxiety and autonomy were estimated using paired-samples t-tests or Wilcoxon tests depending on the distributions. The TAM questionnaire items were analyzed using exploratory factor analysis. Cronbach's alpha was calculated for each questionnaire construct.

As for the qualitative data analysis, it followed Braun and Clarke's (2022) framework of Reflexive Thematic Analysis. Journals and interviews were analyzed in six stages: familiarizing with data, generating initial codes, searching for themes, reviewing themes, defining them and producing report.

Auditing trail was compiled to record researchers' codes and reflections.

The joint display framework made it possible to integrate qualitative and quantitative strands of data through comparing their findings, explaining one type of results with the other and finding out the discrepancies in the experiences of the respondents and statistical data.

Ethical Considerations

Given that a substantial proportion of the participants were between 12 and 18 years of age, the protection of minors was an integral component of the research design. The study followed internationally recognized principles for research involving human participants, with particular attention to informed consent and assent, voluntary participation, confidentiality, data protection, and participants' right to withdraw from the study without penalty.

Prior to data collection, authorization to conduct the study was obtained through the institutional procedures established by the participating school. Because the study involved minors, a two-level consent procedure was implemented. Written informed consent was obtained from parents or legal guardians, while students provided their own assent to participate. The consent and assent forms were provided in Spanish and explained the purpose and procedures of the study, the use and protection of the collected data, the voluntary nature of participation, and the right to withdraw without academic penalty. Only students for whom parental or guardian consent and student assent had been obtained were included in the research procedures.

To protect participants' identities, alphanumeric codes were assigned to all research records. Personally identifying information was stored separately from the research data in a protected file accessible only to the research team. Research data were securely stored and managed in accordance with the data-protection and retention requirements established by the participating institution.

RESULTS

Descriptive Analysis of the Pre-test and Post-test

Table 3 shows the summary statistics for the pre-test and post-test among the overall sample ($n = 145$), comprising the scores on all of the five dimensions of writing. Before the intervention, students scored an average Total Score of 51.4 ($SD = 12.3$) from 100. At the end of the 12-week period, the average score increased to 67.8 ($SD = 10.7$) after the intervention, which is indicative of an increase of 16.4 in their scores in the pre-test and post-test comparison. It provides a preliminary picture of the effects of the intervention. Further statistical analysis seeks to examine whether there is a significant effect, what the size of the effect is, and whether the effect can be observed at the Lower and Higher-Order dimensions level separately.

Table 3*Descriptive Statistics for Rubric Domain Scores: Pre-test and Post-test (n = 145)*

Domain	Max.	Pre M (SD)	Post M (SD)	ΔM	t	df	p	Cohen's d
Content	30	18.2 (4.1)	22.1 (3.4)	+3.9	10.24	144	< .001	0.72
Organization	20	12.4 (3.2)	15.8 (2.7)	+3.4	9.87	144	< .001	0.68
Vocabulary	20	10.3 (3.5)	14.2 (2.9)	+3.9	13.45	144	< .001	1.05
Grammar	25	8.1 (3.8)	12.4 (3.2)	+4.3	14.71	144	< .001	1.12
Mechanics	5	2.4 (0.9)	3.3 (0.8)	+0.9	11.33	144	< .001	0.88
Total Score	100	51.4 (12.3)	67.8 (10.7)	+16.4	14.23	144	< .001	0.90

Note. M = mean; SD = standard deviation; ΔM = post-test minus pre-test mean gain; t = paired-samples t-statistic; df = degrees of freedom; p = two-tailed significance value; Cohen's d interpreted against Plonsky & Oswald (2014) EFL benchmarks: small = 0.40, medium = 0.70, large = 1.00. Bonferroni-corrected $\alpha = .01$ applied to domain-level comparisons.

It becomes clear that all dimensions of the individual writing have similar dynamics. It should be noted that there were the most impressive improvements of Lower-Order Concerns (LOCs): Grammar was improved to $d = 1.12$, while Vocabulary gained a result of $d = 1.05$. Additionally, Mechanics showed improvement to the extent of $d = 0.88$. Higher-Order Concerns (HOCs) also became better but in comparison with LOCs, this progress is less obvious. Content showed a result of $d = 0.72$, and Organization improved to $d = 0.68$.

It should be mentioned that the discrepancy is especially noticeable in the case of LOCs vs. HOCs: The largest effects were observed for Grammar ($d = 1.12$) and Vocabulary ($d = 1.05$), followed by Mechanics ($d = 0.88$). However, students improved both surface-level and deeper dimensions of their writing: both Content and Organization dimensions showed some progress despite the difference in d values. In general, all five dimensions, as well as the Total Score value became significantly higher after the intervention under the Bonferroni-adjusted criterion of significance ($\alpha = .01$).

The distribution of scores was normally distributed almost for all groups. One exception should be discussed here. There were no normal distributions in the 9th Grade EGB group ($n = 18$) because the Shapiro–Wilk test reported deviation from normality ($p = .038$). Thus, Wilcoxon signed-rank test was used in the analysis for this group. The significant difference in the Total Score ($Z = 3.68$, $p < .001$) means that the same pattern of improvement can be traced even with the use of the non-parametric statistical procedure.

ANCOVA has helped to get more details about differences between various classroom cohorts. While controlling for the pretest writing scores, the classroom cohort had a significant impact on post-test performance of students, $F(5, 138) = 8.74$, $p < .001$, partial $\eta^2 = .24$. The result shows that there are still differences between various classroom cohorts of students even after taking into account writing performance of participants.

The Bonferroni-adjusted post-hoc test showed that students in 3rd Year BGU scored significantly higher adjusted post-test scores compared to those from 8th Grade EGB. The difference in mean values between these two groups was 11.3 points (95% CI [7.2, 15.4]).

It should be highlighted that the improvement was observed not only among students in the upper-grade cohorts. All six groups of students were at different stages of English acquisition at the beginning and showed diverse writing skills at the end of intervention. However, all students made significant progress during the experiment. Table 4 presents descriptive statistics for the three questionnaire constructs measured at pre- and post-intervention.

Table 4
Questionnaire Construct Composite Scores: Pre- and Post-Intervention (n = 145)

Construct	Items	α	Pre (SD)	M	Post (SD)	M	ΔM	p	Cohen's d
Writing Anxiety (SLWAI)	5	.82	3.42 (0.73)		2.18 (0.61)		-1.24	< .001	1.35
Learner Autonomy	5	.79	2.89 (0.68)		3.76 (0.57)		+0.87	< .001	0.92
Perceived Usefulness (PU)	4	.88	—		4.12 (0.64)		—	—	—
Perceived Ease of Use (PEOU)	4	.91	—		4.47 (0.52)		—	—	—

Note. α = Cronbach's alpha internal consistency coefficient; M = mean composite score (5-point Likert); SD = standard deviation; ΔM = post-minus-pre mean change. SLWAI: higher scores = greater anxiety; post-intervention reduction indicates anxiety amelioration. Writing anxiety and learner autonomy items were administered at both pre-intervention and post-intervention, whereas the TAM constructs (PU and PEOU) were administered only after the intervention because they assessed students' perceptions of LuzIA following actual use of the tool.

Correlation analyses showed a moderate positive association between Perceived Usefulness and improvement in writing scores ($r = .48$, $p < .001$), indicating that students who perceived LuzIA as more useful tended to show greater gains in writing performance. Perceived Ease of Use was also positively associated with Perceived Usefulness ($r = .39$, $p < .001$), indicating that students who perceived LuzIA as easier to use also tended to report greater perceived usefulness.

DISCUSSION

As shown in the table of results above, there is a definite trend regarding writing scores, questionnaire outcomes, and students' reports about experiences with the AI-based tool during the intervention. These results indicate where the support provided by LuzIA worked best and where it proved less useful.

Positive effects occurred most strongly in the Lower-Order Concerns (LOCs), particularly Grammar ($d = 1.12$), Vocabulary ($d = 1.05$), and Mechanics ($d = 0.88$). This pattern suggests that LuzIA may have been particularly useful in providing immediate language-level support. Previous research has similarly shown that LLM-based writing tools can provide immediate feedback on grammar, vocabulary use, sentence structure, and alternative phrasing, creating

opportunities for learners to revise their writing (Song & Song, 2023). More broadly, the stronger post-revision writing performance reported by Tsai et al. (2024) provides additional evidence that ChatGPT-assisted revision can contribute to improved assessed writing performance, although their study did not establish the same LOC-specific effect pattern observed in the present research.

This trend was also illustrated by qualitative findings as learners stated grammar correction and vocabulary substitution as primary motives for using LuzIA in reflective journals and interviews. Requesting assistance related to argumentation and essay organization appeared in students' responses too, yet not consistently.

In Higher-Level Concerns category the improvement is meaningful yet smaller, with $d = 0.72$ in terms of Content and $d = 0.68$ with regards to Organization. However, this aspect of the results is crucial since developing an argument is a completely different matter compared to correcting a sentence. Developing a powerful thesis, a counter-argument, and good paragraphs require context, purpose, audience, and rhetorical awareness – the areas where the current LLMs appear to provide less consistent support for higher-order rhetorical decisions.

Qualitative data also confirmed this aspect – while some students found the suggestions concerning organization helpful (it came in handy when learners did not know how to structure their arguments and connect their ideas), the others viewed them as too generic or unsuited to the academic tone.

Thus, the experience with regards to higher-level concerns can be considered unpredictable, confirming the idea of applying the tool in combination with the teacher, but not as a replacement. AI can provide immediate language-level support while the teacher plays the role of helping students develop argumentation and rhetorical skills needed for writing good essays.

Perhaps, the most impressive result of this study was not related to grammar or rhetoric, but to emotion. The writing anxiety decreased with the largest measured effect ($d = 1.35$).

Qualitative findings offer two plausible explanations for this pattern. Firstly, it is possible that learners could communicate with the tool without feeling immediate criticism of their mistakes since they had private interaction with it. Secondly, it is possible to correct errors in numerous revisions – revise a sentence, submit the next version, reflect on feedback and continue working without exposing failed attempts to public judgment.

The opportunity to make mistakes in a private and non-judgmental environment may be particularly important in high-stakes writing contexts. This interpretation is consistent with Siagian et al. (2026), who reported that EFL learners using DOLA AI for speaking practice experienced reduced fear of negative evaluation and greater emotional comfort during AI-mediated interaction. A similar mechanism may help explain the reduction in writing anxiety observed in the present study, as learners were able to experiment with language and receive

support without immediate peer or teacher evaluation. In certification-oriented writing contexts, this private space may be particularly relevant because written errors remain visible and may subsequently be evaluated.

Autonomy of students also improved significantly ($d = 0.92$). The reason why it is an important point is the ultimate goal of introducing an AI into writing course, which is making the students less dependent on it in the long run.

According to reflective journals, students began to move toward increased autonomy in the later stages of the intervention. During the last two weeks, students started to describe themselves doing grammar check, vocabulary search and organization review independently without immediately resorting to LuzIA.

This finding is consistent with the developmental path proposed by Vygotsky's ZPD and learner autonomy concept of Holec (1981). An external support can trigger this path and some strategies which have been developed under such conditions can slowly become part of students' practices. The conclusion is not about discontinuing using LuzIA after 12 weeks but rather about the evolution of students' practices: students begin to utilize the tool not automatically looking up the answers there but using it among the other tools available to them.

It is important to mention results of TAM questionnaire. Perceived ease of use turned out to be high ($M = 4.47/5$), which confirms the idea of decreasing the effort required for adopting the educational tool through its access through already familiar platform (WhatsApp).

Perceived usefulness was also high ($M = 4.12/5$) and was positively associated with writing improvement ($r = .48, p < .001$). Students reporting greater perceived usefulness tended to show larger gains in their writing scores. Perceived ease of use was also positively associated with perceived usefulness ($r = .39, p < .001$), which is consistent with Davis's (1989) TAM.

Thus, it turns out that usefulness of LuzIA does not consist in excelling in all components of academic writing, but consists in the fact that it provides the language-level support along with reduction of anxiety in students and their increased independence in the later stages of revision practices. Teacher remains irreplaceable in helping with more complex rhetorical questions and decision-making.

CONCLUSIONS

This research analyzes the effect of using LuzIA, an artificial intelligence assistant based on the use of GPT and accessible on WhatsApp for improving academic writing, writing anxiety, and learner autonomy among EFL learners in a multilingual secondary-school environment in Ecuador. Combining quantitative results with the learners' experiences leads to three key conclusions.

To begin with, the use of LuzIA is related to statistically significant improvements in the area of academic writing skills. In this case, the strongest increases were noticed in Lower-Order

Concerns (LOCs), especially in such areas as Grammar ($d = 1.12$) and Vocabulary ($d = 1.05$). Moreover, Content ($d = 0.72$) and Organization ($d = 0.68$) also become better. These effect sizes can be compared positively to the benchmarks for EFL interventions provided by Plonsky & Oswald (2014).

It is also important that different improvements for LOCs and HOCs have particular pedagogical meaning. It appears that LuzIA is beneficial to deal with writing issues that many learners find difficult to monitor while developing an argument – grammar, vocabulary, spelling, and other linguistic issues. In this way, with the help of some of such tasks being delegated to AI, students can develop the ability to pay more attention to the intended messages and structures of texts. However, the data do not imply that the role of AI will allow replacing the entire writing process as it still needs contextual decision-making in which teacher intervention is indispensable. Therefore, the most beneficial role of AI in the case of LuzIA is to serve as assistance for teachers.

Secondly, it is worth noticing that using LuzIA causes the reduction of writing anxiety ($d = 1.35$). In this case, it is essential to consider the fact that it does not merely reflect the negative feeling but also the decision about writing. Moreover, the anxiety can affect the extent to which a student will exercise his/her writing abilities. Lack of exercises will lead to poor improvement, as it was mentioned before.

In fact, LuzIA allows giving students a personal space, which is very difficult to create in a crowded classroom, where they can try, ask, correct mistakes, etc. In this way, the opportunity to deal with a text multiple times seems to cause the relief related to the process of academic writing. So, it means that not only the corrections can be valuable in the case of this tool but also the positive feeling related to writing and correcting the text can encourage more practice.

Finally, one more important thing to mention is the increase in learner autonomy ($d = 0.92$). In this case, qualitative analysis shows a specific dynamic in terms of the development of the use of strategies: in the second half of the intervention, students start reporting their use of independently performed strategy that they initially practiced with the help of LuzIA (like checking grammatical issues, searching for alternatives of vocabulary and the organization structure).

This does not mean the automatic elimination of AI dependence. The dependence depends on how the tool is introduced and used in the class. In the current research, the approach implies that students needed to question the answer of AI, decide which suggestion is suitable for him/her, etc. So, it means that simple acceptance of AI advice can decrease thinking required in writing. Otherwise, analyzing and implementing the feedback is able to be included in the process of learning revision strategies.

What can be considered the distinctive feature of the current study is that LuzIA operates via WhatsApp, a tool familiar for all students. In this way, there is no need to introduce any new interface and platform as there is no special software that requires installation. Moreover, there is

no additional platform that should be integrated in the class before usage. This aspect can be very helpful to use the technology as there are cases in which schools are unable to afford it.

Thus, the main finding of the current research is that technology can provide effective support of EFL learners in improving language-level accuracy of the academic texts while teachers still bear full responsibility for developing arguments and rhetorical issues of students. In fact, the goal of this use is not to train students to write properly only when they use AI but make the process such as they will become better writers through AI support.

Many important issues need further investigation. Future research is supposed to investigate whether such effects will last if students have reduced access to LuzIA, to compare LuzIA and alternative AI writing assistants under similar conditions, and to conduct the randomized controlled design of the current intervention where possible in such an environment. In addition, longitudinal studies will be extremely useful to check if the learned strategies become autonomous.

REFERENCES

- Ayala-Pazmiño, M. F., & Alvarado-Lucas, K. I. (2023). Integración de la inteligencia artificial en la educación del idioma inglés en Ecuador: Un camino para mejorar los resultados del aprendizaje. *593 Digital Publisher CEIT*, 8(3-1), 679–687. <https://doi.org/10.33386/593dp.2023.3-1.1862>
- Behforouz, B., & Al Ghaithi, A. (2024). The impact of using interactive chatbots on self-directed learning. *Studies in Self-Access Learning Journal*, 15(3), 317–344. <https://doi.org/10.37237/150302>
- Braun, V., & Clarke, V. (2022). *Thematic analysis: A practical guide*. SAGE Publications.
- Cheng, Y.-S. (2004). A measure of second language writing anxiety: Scale development and preliminary validation. *Journal of Second Language Writing*, 13(4), 313–335. <https://doi.org/10.1016/j.jslw.2004.07.001>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). SAGE Publications.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340. <https://doi.org/10.2307/249008>
- Ghafar, Z. N., & Raheem, B. R. (2025). Understanding EFL learners' academic writing problems and the benefits of artificial intelligence tools: A descriptive review. *Journal of Language and Pragmatics Studies*, 4(3), 187–193. <https://doi.org/10.58881/jlps.v4i3.141>
- Holec, H. (1981). *Autonomy and foreign language learning*. Pergamon Press.
- Isabella, P., Ferdiyanto, F., & Zainuddin, Z. (2026). Implementation of Dola Artificial Intelligence as a learning tool to improve students' English speaking skills. *Journal of English Language and Education*, 11(2), 640–650. <https://doi.org/10.31004/jele.v11i2.2264>
- Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Newbury House.
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press.
- Li, M., & Wilson, J. (2025). AI-integrated scaffolding to enhance agency and creativity in K-12 English language learners: A systematic review. *Information*, 16(7), 519. <https://doi.org/10.3390/info16070519>
- LuzIA. (n.d.-a). *LuzIA: Your intelligent assistant at a click*. <https://www.luzia.com/en>
- LuzIA. (n.d.-b). *Help: Frequently asked questions*. <https://www.luzia.com/ayuda>

- Maforah, N., & Leburu-Masigo, G. (2018). Application of the mixed methods research using sequential explanatory design. In *ICERI2018 Proceedings* (pp. 9710–9715). IATED. <https://doi.org/10.21125/iceri.2018.0797>
- Marzuki, M., Widiati, U., Rusdin, D., Darwin, D., & 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>
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). McGraw-Hill.
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878–912. <https://doi.org/10.1111/lang.12079>
- Sanchez, H. S. (2024). Teacher-written feedback in L2 writing education: Insights into student tensions and responses. *Feedback Research in Second Language*, 2, 58–80. <https://doi.org/10.32038/frsl.2024.02.04>
- Siagian, L. P., Kharisma, A. J., & Syahputri, D. (2026). Mitigating EFL speaking anxiety with Dola AI: Enhancing oral proficiency. *Journal of Social, Humanity, and Education*, 6(2), 255–266. <https://doi.org/10.35912/jshe.v6i2.3926>
- Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14, 1260843. <https://doi.org/10.3389/fpsyg.2023.1260843>
- Tsai, C.-Y., Lin, Y.-T., & Brown, I. K. (2024). Impacts of ChatGPT-assisted writing for EFL English majors: Feasibility and challenges. *Education and Information Technologies*, 29(17), 22427–22445. <https://doi.org/10.1007/s10639-024-12722-y>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.