https://doi.org/10.69639/arandu.v13i3.2599

The artificial intelligence tool ChatGPT and its impact on the development of the integrated writing skill in students at the International Learning' Institute

La herramienta de inteligencia artificial ChatGPT y su impacto en el desarrollo de la habilidad de escritura integrada en los estudiantes del Instituto International Learning

Paul Sebastián Montenegro Rosero
sebastian.montenegro@upec.edu.ec
https://orcid.org/0009-0006-3203-4717
Universidad Politécnica Estatal del Carchi
Ecuador – Tulcán

Elena Valeria Flores Borja
elenaflores@upec.edu.ec
https://orcid.org/0009-0005-4800-8938
Universidad Politécnica Estatal del Carchi
Ecuador - Tulcán

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

ABSTRACT

The present research analyzed the impact of the pedagogical use of ChatGPT on the development of English written production in students at the International Learning Institute, located in the city of Tulcán, Ecuador. The study arose from the difficulties identified in grammatical precision, spelling, syntactic organization, coherence, cohesion, and lexical variety of the texts produced by the students. The study was conducted with the entire population, consisting of 45 students aged between 14 and 22 years, with English proficiency levels ranging from A1 to B2. The research was developed using a quantitative approach, with a descriptive scope and a single-group exploratory design with a pretest and posttest. A questionnaire and a written production evaluation rubric were used for data collection. During the intervention, ChatGPT was used as a complementary tool to provide immediate feedback, identify grammatical and spelling errors, suggest improvements in textual structure, and promote vocabulary expansion. The results obtained through the diagnostic survey evidenced difficulties in sentence structuring, the use of varied vocabulary, punctuation, and grammatical accuracy, as well as differences in the frequency with which students used ChatGPT to support their writing. It is concluded that the planned, ethical, and supervised integration of ChatGPT can constitute a complementary strategy to strengthen English written production, if its application is accompanied by teacher guidance and the student's critical participation.

Keywords: chatGPT, english written production, artificial intelligence, language learning, linguistic feedback

RESUMEN

La presente investigación analizó la incidencia del uso pedagógico de ChatGPT en el desarrollo de la producción escrita en inglés de los estudiantes del Instituto International Learning, ubicado en la ciudad de Tulcán, Ecuador. El estudio surgió ante las dificultades identificadas en la precisión gramatical, la ortografía, la organización sintáctica, la coherencia, la cohesión y la variedad léxica de los textos producidos por los estudiantes. Se trabajó con la totalidad de la población, conformada por 45 estudiantes con edades comprendidas entre 14 y 22 años y niveles de dominio del inglés de A1 a B2. La investigación se desarrolló mediante un enfoque cuantitativo, con alcance descriptivo y un diseño exploratorio de un solo grupo con pretest y postest. Para la recolección de información se emplearon un cuestionario y una rúbrica de evaluación de la producción escrita. Durante la intervención, ChatGPT se utiliza como complementario para proporcionar retroalimentación inmediata, identificar errores gramaticales y ortográficos, sugerir mejoras en la estructura textual y favorecer la ampliación del vocabulario. Los resultados obtenidos mediante la encuesta diagnóstica evidenciaron dificultades en la estructuración de oraciones, el uso de vocabulario variado, la puntuación y la corrección gramatical, así como diferencias en la frecuencia con la que los estudiantes utilizaban ChatGPT para apoyar su escritura. Se concluye que la integración planificada, ética y supervisada de ChatGPT puede constituir una estrategia complementaria para fortalecer la producción escrita en inglés, siempre que su aplicación esté acompañada por la orientación docente y la participación crítica del estudiante.

Palabras clave: chatGPT, producción escrita en inglés, inteligencia artificial, aprendizaje de lenguas, retroalimentación lingüística

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

INTRODUCTION

Writing in English as a foreign language faces recurring barriers, such as the lack of timely feedback and deficiencies in cohesion and vocabulary. Although generative artificial intelligence, particularly ChatGPT, emerges as a resource to personalize learning, its unregulated adoption generates a critical debate: does it act as a pedagogical guide or as a substitute for cognitive effort? In contexts like the Ecuadorian one, where traditional approaches still predominate, there is an empirical gap regarding how to structure teacher mediation to guarantee an ethical and effective use of these tools. The present study addresses this gap by evaluating the impact of a supervised pedagogical intervention with ChatGPT on 45 students at the "International Learning" Institute in Tulcán, Ecuador. Unlike passive use approaches, this research implements a structured model in four phases (planning, drafting, assisted revision, and editing), where the student must contrast and justify the algorithmic suggestions. The results provide quantitative evidence that the planned integration of AI significantly improves linguistic precision and lexical diversity, fostering metacognitive processes and offering a practical framework for digital literacy in language teaching.

Background of the Problem

Written production in English constitutes a fundamental competence in 21st-century academic and professional training, as it allows students to articulate ideas, argue rigorously, and participate in international disciplinary communities. However, those who learn this language as a second language face recurring difficulties: scarcity of specialized vocabulary, problems with cohesion and textual structuring, and, above all, the absence of timely feedback that allows them to identify and correct their errors autonomously. Faced with these limitations, generative artificial intelligence tools such as ChatGPT have emerged as a possible pedagogical resource to accompany the writing process. Dou et al. (2025) and Román et al. (2026) highlights that intelligent environments reduce the barriers associated with the lack of immediate feedback, particularly among second language learners, and facilitate the creation of personalized learning experiences that accelerate the development of written competence.

Likewise, Ossa and Willatt (2023) implemented generative AI to provide descriptive feedback on pedagogical responses elaborated by pre-service teachers, highlighting the tool's capacity to support formative and large-scale assessment of students' written work. Baldrich and Domínguez (2024), in their case study on the use of ChatGPT in academic writing, observed improvements in written tasks regarding coherence, cohesion, and academic language; they concluded that ChatGPT can serve as a complementary tool for academic work as long as students already possess fundamental skills in critical thinking, ethics, and argumentation.

The development of written competence in a foreign language requires systematic guidance, continuous feedback, and interactive learning environments that favor meaningful

learning. However, traditional methods of teaching English present considerable limitations: as Cordero et al. (2025) point out in their study conducted in Ecuador, numerous educational institutions still depend on approaches focused on grammar and memorization, which restrict opportunities for communicative practice and hinder the effective learning of writing. Faced with this lack of pedagogical support, students autonomously resort to artificial intelligence tools. These findings make it evident that the unregulated use of AI, far from supplying training deficiencies, can aggravate them, and reinforce the need for a pedagogically mediated integration that guides the student in the critical and responsible use of these tools.

The incorporation of artificial intelligence applications such as ChatGPT has transformed the teaching of English through immediate feedback, content personalization, and autonomous practice. Farrokhnia et al. (2024) argue that ChatGPT functions as a dynamic educational resource by increasing access to information, facilitating personalized learning, and providing coherent responses in real time. Adel et al. (2024) point out that ChatGPT can significantly enrich education by offering dynamic and personalized learning experiences, as well as real-time feedback, thus boosting student engagement and teaching efficiency when integrated into responsible ethical frameworks.

Lee and Ko (2026) found that AI-based assistants provide automated feedback on writing, improving cohesion, accuracy, and fluency in English as a Foreign Language (EFL) teaching context. Likewise, Zhai and Nezakatgoo (2025) determined that AI facilitates personalized progress tracking, fostering metacognitive processes that reinforce self-assessment and independent revision skills. Kohnke et al. (2023) highlight that the pedagogical relevance and quality of AI suggestions depend largely on effective prompt engineering, which requires users to iteratively refine their input instructions. Nevertheless, the literature also points out certain challenges related to the integration of artificial intelligence in educational settings. Yan et al. (2025) warn that excessive reliance on AI tools can lead to critical problems such as algorithmic bias, data privacy vulnerability, the weakening of human relationships, and the deterioration of critical thinking if there is no adequate pedagogical mediation. Baidoo-Anu and Owusu Ansah (2023) point out that its adoption must be carefully guided to avoid problems such as the generation of inaccurate information, privacy vulnerabilities, and the amplification of pre-existing biases in training data.

Specifically in the context of EFL writing, Aliakbar (2024) demonstrates that ChatGPT improves linguistic precision and lexical variety in students' compositions. However, research suggests that its pedagogical potential is maximized when used as a mediation tool, rather than as a total substitute for cognitive effort and writing processes. Liu (2024) documents performance differences between groups of students assisted by AI and those who received traditional instruction, generally supporting the positive impact of automated writing evaluation on writing accuracy.

One of the challenges facing the teaching of English is written composition; learning takes place in an educational environment where traditional methods based on repetition and memorization usually predominate (Párraga Solórzano et al., 2024). Academic writing is a crucial skill in education worldwide, as it allows students to communicate ideas clearly, coherently, and effectively in diverse academic disciplines. Despite its importance, many students lack proficiency in academic writing, which significantly hinders their school performance and limits their opportunities for academic and professional success.

In the Ecuadorian context, educational institutions continue to present deficiencies in the teaching and development of writing skills. Students often have difficulties organizing ideas, maintaining textual coherence, and correctly applying grammar and syntax. Traditional pedagogical approaches still predominate in classrooms and frequently fail to offer personalized and effective feedback that fosters writing competence. Furthermore, the implementation of technological tools, particularly those based on artificial intelligence, remains limited or non-existent in many classrooms. Faced with this situation, the following research questions arise: How does the application of the artificial intelligence tool ChatGPT influence the development of the integrated skill of written production in English? Likewise, the following guiding questions have been formulated for the study: a) What are the main difficulties faced by students at the "International Learning" institute in the academic writing process before using artificial intelligence tools? b) What characteristics does ChatGPT possess that can effectively support written production? c) How is ChatGPT implemented in the teaching-learning process of academic English written production?

Research Objectives

In this sense, the research proposes the following general objective: to determine the impact of the digital tool ChatGPT on the development of the integrated skill of written production in students at the "International Learning" Institute, located in the city of Tulcán, Carchi province. To achieve this, the following specific objectives were proposed: a) to identify the main difficulties in students' written production before the implementation of ChatGPT; b) to apply the artificial intelligence tool ChatGPT as a pedagogical resource in the teaching-learning process of written production; and c) to evaluate the changes in the level of students' written production by comparing the results obtained in the pretest and posttest, in order to formulate pedagogical guidelines for the effective integration of this tool in the classroom.

Research Hypothesis

According to the research objectives, the following null hypothesis was formulated: H_0: There is no statistically significant difference between students' writing scores in the pretest and posttest after the pedagogical use of ChatGPT. H_1: There is a statistically significant difference between students' writing scores in the pretest and posttest after the pedagogical use of ChatGPT.

Justification of the Research

Considering the reviewed evidence, ChatGPT offers significant opportunities to overcome the limitations of traditional English teaching methods that still predominate in the classrooms of the Ecuadorian educational system. Educational inclusion, in this sense, is not reduced to access, but demands the implementation of pedagogical strategies that eliminate barriers and guarantee quality training, a principle supported by the Constitution of the Republic of Ecuador and the Organic Law of Intercultural Education (National Assembly of Ecuador, 2008). However, despite the regulatory framework that promotes innovation and educational equity, research on the use of generative artificial intelligence tools in the development of English writing remains virtually unexplored in local contexts such as "International Learning". This absence of empirical studies that characterize the practices, perceptions, and needs of students in said locality constitutes a gap that limits informed pedagogical decision-making and, therefore, justifies the realization of the present research.

At the local level, a possible contextual limitation at the "International Learning" institute would lie in the accentuation of the difficulties described previously. It is likely that teachers do not employ technology tools based on artificial intelligence in the teaching of writing. If so, students might be producing texts with deficiencies in structure, clarity, and communicative purpose, which would limit both their academic development and their capacity for personal expression. If this scenario persists without pedagogical interventions aimed at modifying it, there would be a risk that students continue to face difficulties in written production, show low confidence when expressing themselves in writing, and develop an excessive dependence on grammatical rules to the detriment of comprehensive communicative competence.

Significance of the Research

This research is justified from a theoretical perspective by contributing to the understanding of how generative artificial intelligence, particularly ChatGPT, can transform the teaching of English writing beyond approaches focused on grammar and memorization. Its social relevance lies in the direct benefit for the students and teachers of the "International Learning" Institute, who access more functional and motivating learning experiences. This purpose is articulated with Sustainable Development Goal 4 of the United Nations 2030 Agenda, aimed at guaranteeing inclusive and quality education through innovative methods (UNESCO, 2024), and with the National Development Plan 2025-2029, which promotes technological innovation and digital transformation (National Planning Secretariat, 2025).

MATERIALS AND METHODS

Research Paradigm and Approach

This study is framed within the positivist paradigm, as it adopts a quantitative approach based on the collection and analysis of numerical data to identify trends and measure the scope

of the impact of an intervention, in this case, the use of the artificial intelligence tool ChatGPT in the development of English written production (Dávila, 2022).

This study adopts a quantitative approach. According to Vizcaíno et al. (2023), it is based on the collection and analysis of numerical data to identify trends and measure the scope of the impact of an intervention. The reasoning is deductive, starting from the theory on the use of AI in writing to test its effect through the statistical measurement of pretest and post-test results. Medina et al. (2023) mention that the use of data collection instruments is determined based on the characteristics of the required information and the goals of the research.

Research Design

Specifically, a pre-experimental single-group design with pre- and post-test evaluation was used. According to Mills and Jordan (2022), the pretest-post-test structure is valuable in educational and behavioral research because it allows describing and quantifying the change within the group between both measurement moments. However, since it is a single-group design without a comparison group, this structure does not allow isolating the effects of the intervention against threats to internal validity such as the natural maturation of the participants, history, the effect of the test itself, or regression to the mean. Consequently, these factors constitute inherent limitations of the single-group pre-experimental design that are not controlled by the simple fact of applying a pretest and a post-test but would require the inclusion of a control group or quasi-experimental designs to be adequately addressed. Despite these limitations, this type of design is adequate when the objective is to explore and describe the magnitude of the change within a group after an intervention, without having comparison groups, obtaining evidence of improvement or intra-group progress, although not causal evidence. Since a control group was not included, the results are strictly interpreted as changes associated with the intervention and not as causal effects attributable exclusively to it.

Intervention

The intervention had an effective duration of 12 weeks of instruction, comprising from January 12 to April 10, 2026, and consisted of 36 sessions of 60 minutes each. ChatGPT (GPT-4o model, OpenAI) was used as a supervised support tool. Throughout the 12 weeks, students were guided through the four phases of the writing process: planning, drafting, revising, and editing.

Table 1
Phases of the writing process

Phase	Activity	Teacher's role	ChatGPT's role
Planning (10 min)	Brainstorming, organizational outline, and topic selection	Guides idea generation and validates the outline	Generates guiding questions and suggests possible thematic approaches
Drafting (20 min)	Production of the first draft by the student	Supervises individual work and addresses any questions.	Does not intervene; the student writes autonomously
Assisted revision (20 min)	The student enters their draft into ChatGPT using the structured prompt (see below) and analyzes the suggestions	Verifies that the student contrasts each suggestion with their prior knowledge and justifies their decisions	Provides feedback in five dimensions (mechanics, structure and syntax, coherence, cohesion, and lexicon)
Editing and closing (10 min)	The student produces the final version incorporating the accepted improvements and reflects on the changes	Collect the final version and the student's decision log	Does not intervene

Scope of Research

The scope of this research is quantitative and pre-experimental, with a single-group design and pre- and post-test evaluation, aimed at exploring and describing the magnitude of intra-group change in English written production after a pedagogical intervention with the artificial intelligence tool ChatGPT (GPT-4o model). The study is delimited to a total population of 45 students. Since a control group was not included, the scope of the conclusions is limited to evidencing the progress or improvement associated with the intervention through the analysis of intra-subject differences (post-test − pretest) using descriptive statistics and the non-parametric Wilcoxon signed-rank test. Consequently, the results are strictly interpreted as changes associated with the intervention and not as causal effects attributable exclusively to it, recognizing as inherent limitations the impossibility of isolating threats to internal validity such as natural maturation, history, the effect of the test itself, or regression to the mean.

Population and Sample

La The accessible population for this study consisted of 45 students enrolled at the "International Learning" English Institute, located in the city of Tulcán, Carchi province, Ecuador. The participants were aged between 14 and 22 years and had English proficiency levels ranging

from A1 to B2, according to the Common European Framework of Reference for Languages (CEFR), a classification determined through the institutional placement test applied at the time of enrollment. The inclusion criteria established that students had to be enrolled in the course under study during the intervention period, have a proficiency level between A1 and B2 according to the CEFR, and participate voluntarily. On the other hand, the exclusion criteria contemplated irregular attendance (less than 80% of the sessions) and dropping out of the program before its completion.

Given the reduced and finite nature of the target population (N < 50), probabilistic or random sampling methods were not used; instead, a total population non-probabilistic sampling approach was used (Cohen et al., 2018). In practice, all 45 students who met the inclusion criteria completed the process without incurring the exclusion criteria, maintaining an attendance equal to or greater than 80% of the sessions. Consequently, no dropout rate, loss of participants, or exclusions due to absenteeism were recorded, so the analyzed sample coincided exactly with the entire accessible population (n = 45) in both the pretest and posttest and in all subsequent statistical analyses. It should be noted that, to mitigate the possible effect of the heterogeneity of levels (A1–B2) as a confounding variable, the score obtained in the pretest was considered as an individual baseline, so that the analysis of the change focused on the intra-subject difference (post-test − pretest) and not exclusively on the absolute score of the posttest.

Table 2

Population

Population= 45 students		
Sample	Frequency	Percentage
A1	9	20%
A2	13	29%
B1	15	33%
B2	8	185
Total	45	100%

Source: Data obtained by the academic institution. **"***International Learning"*

Data Collection

Data collection with the 45 students was structured in three phases. The diagnostic phase included the application of a 15-item questionnaire on AI use and a written production pretest. During the intervention phase, students carried out writing activities assisted by ChatGPT under teacher mediation. Finally, in the evaluation phase, a post-test graded with the same analytical rubric as the pretest (5-20 points) was applied to measure progress. Prior to informed consent, all data were anonymized and safeguarded for exclusively academic purposes.

Validity and Reliability of the Data Collection Instrument

The instruments were subjected to validation through expert judgment. According to Venegas et al. (2025), ensuring validity implies controlling extraneous factors to establish internal validity and confirm that the instrument measures the intended construct, thus ensuring the accuracy of the conclusions about students' writing skills. Three specialists in English teaching, educational research, and instrument design evaluated the pertinence, clarity, coherence, and correspondence of each item with the variables and dimensions of the study. After the review, the experts issued their conformity, approving the definitive structure of the questionnaire and the analytical rubric without major observations, which certified the content validity of both instruments for their application in the selected sample.

Table 3

Reliability statistics

Dimensión	Preguntas (Ítems)	Valor Alfa de Cronbach
1. Teacher's use of AI	1, 3 and 10	0.794
2. Student's use of AI (Correction/Consultation)	2, 4, 5, 6, 7, 8 and 9	0.879
3. Written production skills	11, 12, 13 and 14	0.768

Note: Value calculated using Excel.

According to the scaling criteria of George and Mallery (2016), all values exceed the minimum threshold of acceptability (>0.70). This guarantees that the instrument has adequate internal consistency and is statistically reliable to measure the study variables in the evaluated population. Statistical processing was carried out excluding the nominal and single-item variables (item 15), thus ensuring compliance with the covariance assumptions required by the Cronbach model.

RESULTS AND DISCUSSION

Results related to the General Objective: determining the impact of the digital tool ChatGPT on the development of the integrated written production skill

The comprehensive analysis of the data confirms that the pedagogical integration of ChatGPT had a positive, statistically significant, and large-magnitude impact on the students' writing competence. Quantitatively, the overall mean of grades increased from 12.96 to 16.51, completely eradicating the "Insufficient" performance level (from 20% to 0%) and exponentially increasing the "Outstanding" level (from 2.2% to 37.8%). The non-parametric Wilcoxon test ($p < 0.001$) and a large effect size ($r = .87$) validate that this improvement was not a product of chance, but a direct and robust consequence of the intervention.

Qualitatively, this impact did not derive from the passive or autonomous use of AI, but from a supervised pedagogical mediation model structured in four phases (planning, drafting, assisted revision, and editing). By requiring students to contrast and justify AI suggestions, ChatGPT ceased to function as a mere automatic corrector. This strategy made it possible to overcome initial deficiencies in grammar, cohesion, and lexicon, fostering autonomy, metacognition, and critical thinking.

The numerical results obtained are based on the administration of a structured 15-item questionnaire with a four-level Likert scale, as well as the application of a pre-test and post-test for writing evaluated through an analytical rubric (score from 5 to 20 points). The items addressed two main dimensions: on the one hand, the use and perception of artificial intelligence in the English classroom (questions 1 to 10); and on the other, the students' self-perception regarding their competence and difficulties in integrated writing skills (questions 11 to 14). Question 15 evaluated the preference for the tool. To interpret the survey results, the numerical values corresponding to the response options were recorded into two groups to facilitate comparative analysis. The first group included the options: “Always (A) + Almost Always (AA)”, which grouped the favorable or higher frequency responses. The second group included the options: “Sometimes (S) + Never (N)”, which grouped the lower frequency or unfavorable responses.

Results related to Specific Objective 1: Identification of difficulties and perceptions prior to the intervention

To identify the main difficulties in the students' writing process and their perceptions of AI tools before the implementation of ChatGPT, the results corresponding to the 15 questions of the instrument were analyzed. Table 4 presents the consolidated frequency summary.

Table 4

Consolidated summary of survey response frequencies (N = 45)

Nro.	Question / Item Focus	Favourable (A + AA)	Unfavorable (S + N)
1	Teacher uses online AI tools as teaching material	13	32
2	Students use ChatGPT to correct writing errors	6	39
3	Teacher uses ChatGPT for vocabulary exercises	19	26
4	Perception: AI facilitates English learning	34	11
5	Students use ChatGPT to correct grammatical errors	8	37
6	Students use ChatGPT to correct spelling errors	3	42
7	Students use ChatGPT to correct punctuation errors	4	41
8	Students use ChatGPT to correct text structure	10	35
9	Students use ChatGPT as a reference to expand vocabulary	22	23

10	Teacher uses ChatGPT to provide writing feedback	10	35
11	Self-assessment: Correct spelling in writing	22	23
12	Self-assessment: Correct punctuation in writing	29	16
13	Self-assessment: Correct sentence structure	10	35
14	Self-assessment: Use of diverse vocabulary	9	36
15	Preferred AI tool for writing (ChatGPT)	40	5 (Other tools)

Note: The Always and + Almost Always options were grouped in the (Favorable) column, while Sometimes and Never were combined in the (Unfavorable) column.

The data show that, regarding practical implementation in the classroom, most students considered the systematic use of AI applications to be limited. In the questions about teacher integration (Q1, Q3, Q10) and autonomous student use for grammatical and structural correction (Q2, Q5, Q8), unfavorable responses clearly predominated. For example, 39 students indicated that they rarely or never use ChatGPT to correct writing errors (Q2), and 35 reported that teachers do not use it to provide writing feedback (Q10).

However, the perception changes when analyzing the potential usefulness of these tools and specific superficial corrections. In Q4, a significant majority (34 students) recognized that AI facilitates English learning. Furthermore, students reported greater use for discrete rule-based tasks, such as spelling (Q6) and punctuation (Q7), although the overall frequency remained low. Regarding self-perceived writing difficulties, the greatest weaknesses were identified in sentence structure (Q8, Q13) and lexical diversity (Q9, Q14), where unfavorable responses exceeded 35 students. Finally, Q15 revealed a clear preference for ChatGPT (40 students) over other tools such as Writing Aid or Grammarly, which underscores its perceived versatility for writing development.

Results related to Specific Objectives 2 and 3: Didactic process and application of ChatGPT

Following the initial diagnosis, a 12-week didactic intervention (36 sessions) was implemented. Students were guided through the four phases of the writing process (planning, drafting, revision, and editing) using ChatGPT (GPT-4o) as a supervised mediation tool. Strict protocols were applied to ensure the critical evaluation of AI suggestions and avoid direct copying.

To evaluate the initial level of integrated competence in writing, a pre-test was administered. Students wrote a composition of between 150 and 200 words about a previous trip. The compositions were evaluated using an analytical rubric focused on grammar, structure, coherence, cohesion, and lexical competence. Table 5 presents the results of the initial test.

Table 5

Initial test results: Writing competence level (N = 45)

Performance Level	Score Range	Number of Students	Percentage
Outstanding	18–20 pts	1	2.2%
Satisfactory	14–17 pts	19	42.2%
Fair	11–13 pts	16	35.6%
Insufficient	5–10 pts	9	20.0%
Total		45	100.0%

Note: The initial test results reveal that most students obtained scores within satisfactory and fair performance levels.

The results of the initial test reveal that 55.6% of students obtained scores within the fair and insufficient performance levels. These findings indicated that, although students possessed a basic capacity to express ideas, they experienced considerable difficulties related to grammatical precision, text organization, and lexical variety, which confirms the survey results. At the end of the 12-week intervention, a final test was administered under identical conditions, in which students were asked to write about a family memory. Table 6 presents the results of the final test.

Table 6

Final test results: Writing competence level (N = 45)

Performance Level	Score Range	Number of Students	Percentage
Outstanding	18–20 pts	17	37.8%
Satisfactory	14–17 pts	23	51.1%
Fair	11–13 pts	5	11.1%
Insufficient	5–10 pts	0	0.0%
Total		45	100.0%

Note: The final test results show that no student obtained an Insufficient grade, while the number of students classified as Fair decreased considerably.

The post-test results demonstrate a substantial improvement. It is worth noting that no student obtained an Insufficient grade, and the Outstanding category increased exponentially from 1 to 17 students (37.8%). The overall means increased from 12.96 (SD = 3.01) in the pre-test to 16.51 (SD = 2.41) in the post-test.

Statistical analysis of writing competence

Prior to the comparative analysis, a Shapiro-Wilk normality test was applied to the paired differences (post-test minus pre-test). The analysis yielded a W statistic of 0.928 with a p-value of 0.008. Since $p < 0.05$, the normality hypothesis was rejected; therefore, a non-parametric test was required. The Wilcoxon signed-rank test was applied to compare the pre-test and post-test scores. The test yielded a Z statistic of 5.84 and a p-value < 0.001. The obtained p-value is considerably lower than the established significance level ($p < .05$), which confirms that the observed improvement is not due to chance. Furthermore, the calculated effect size ($r = .87$) is classified as large, indicating that the magnitude of the improvement was substantial and

pedagogically relevant. Consequently, the null hypothesis (H_0) is rejected, and the alternative hypothesis (H_1) is accepted, confirming a significant difference in writing competence after the ChatGPT intervention.

Analysis of results

The initial survey results revealed a gap between the availability of AI tools and their systematic pedagogical integration, since most students reported infrequent use of AI for structural or grammatical correction. This coincides with Dou et al. (2025) and Farrokhnia et al. (2024), who point out that, although students have positive attitudes towards AI, its pedagogical effectiveness largely depends on structured implementation and teacher support. The high preference for ChatGPT (Q15) and the recognition of its usefulness (Q4) support Adel et al. (2024), who highlight that generative AI boosts student engagement when integrated into appropriate frameworks.

The pre-test results confirmed the root causes of the problem: significant difficulties in sentence structure, lexical diversity, and overall coherence. However, the post-test results and the Wilcoxon statistical analysis ($p < .001$, $r = .87$) demonstrated that the 12-week mediated intervention radically transformed these results. The elimination of the "Insufficient" category and the increase in "Outstanding" results provide empirical evidence that ChatGPT, when used as a mediation tool rather than a substitute for cognitive effort, significantly improves academic writing.

These findings coincide with those of Carrillo et al. (2023) and Aliakbar (2024), who found that AI integration improves textual clarity, linguistic precision, and lexical variety in students' compositions. Furthermore, the substantial effect size supports Zhai and Nezakatgoo (2025), who determined that AI facilitates personalized progress tracking, fostering metacognitive processes that strengthen self-assessment. By requiring students to maintain three preliminary drafts and justify their acceptance or rejection of AI suggestions, the intervention successfully mitigated the risks of excessive dependence and erosion of critical thinking warned by Yan et al. (2025) and Baidoo-Anu and Owusu Ansah (2023). Ultimately, the study confirms that automated writing evaluation through artificial intelligence, when combined with explicit teacher guidance, acts as a powerful catalyst for the development of comprehensive writing competence in EFL contexts.

CONCLUSION

Diagnosis of initial difficulties and the usage gap: before the intervention, students at the "International Learning" Institute presented marked deficiencies in written production, particularly in sentence structuring, textual coherence, and lexical diversity. Although there were a favorable perception of artificial intelligence and a clear preference for ChatGPT, its use was

limited, non-systematic, and lacked pedagogical guidance to leverage its potential for formative feedback.

The application of ChatGPT (GPT-4o model) for 12 weeks, structured in the phases of planning, drafting, assisted revision, and editing, proved to be a highly effective pedagogical strategy. The teacher's role was decisive in preventing cognitive dependence, since by requiring students to justify the acceptance or rejection of AI suggestions, critical thinking, autonomy, and real appropriation of the writing process were fostered, rather than the mere copying of generated texts.

The comparison between the pre-test and post-test reveals a statistically significant ($p < 0.001$) and large-magnitude ($r = .87$) improvement in the integrated written production skill. The total eradication of the "Insufficient" level and the substantial increase of students in the "Satisfactory" and "Outstanding" levels confirm that the immediate, personalized, and multidimensional feedback (grammar, cohesion, coherence, and lexicon) provided by the AI, under supervision, accelerates the development of writing competence in English as a second language learners.

The research concludes that the integration of generative AI tools in the English classroom should not be seen as a threat to learning, but as an opportunity to transform traditional methods focused on memorization. For this integration to be successful and ethical, it is essential that educational institutions train teachers in the design of effective prompts, establish responsible use protocols, and promote a "dialogue of knowledge" where AI complements, but does not replace, the student's cognitive effort and human guidance.

REFERENCES

Adel, A., Ahsan, A., & Davison, C. (2024). ChatGPT Promises and Challenges in Education: Computational and Ethical Perspectives. In *Education Sciences* (Vol. 14, Number 8). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/educsci14080814

Aliakbar, T. (2024). *Exploring the Potential of ChatGPT in EFL Language Learning: Learners' Reflections and Practices*. https://doi.org/10.20944/PREPRINTS202412.2218.V1

Asamblea Nacional del Ecuador. (2008). *Constitución de la República del Ecuador 2008* (Decreto Legislativo 0). Decreto Legislativo 0. https://www.gob.ec/sites/default/files/regulations/2020-06/CONSTITUCION%202008.pdf

Baidoo-Anu, D., & Owusu Ansah, L. (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

Baldrich, K., & Domínguez, J. C. (2024). El uso de ChatGPT en la escritura académica: Un estudio de caso en educación. *Revista de Medios y Educación*. https://recyt.fecyt.es/index.php/pixel/article/view/103527/79832

Carrillo, C. C. E., Herrera, B. V. A., & Cortes, S. J. N. (2023). Inteligencia Artificial para la escritura académica en investigación. *Ciencia Latina Revista Científica Multidisciplinar*, *7*(4), 4604–4621. https://doi.org/10.37811/CL_RCM.V7I4.7304

Código Orgánico Integral Penal, COIP (2021). https://www.defensa.gob.ec/wp-content/uploads/downloads/2021/03/COIP_act_feb-2021.pdf

Cohen, L., Manion, L., & Morrison, K. (2018). *Research Methods in Education* (Routledge, Ed.; 8th ed.). Taylor & Francis Group. www.routledge.com/cw/cohen

Cordero, M. L. J., Pachar, L. E. R., & Granados, R. J. F. (2025, March). Ethics in the application of artificial intelligence for the creation of academic work by students in the third year of high school. *593 Digital Publisher CEIT*. https://www.593dp.com/index.php/593_Digital_Publisher/article/view/3150/2554

Dávila, A. (2022). El investigador y el paradigma de investigación. *Revista Investigación, Transcomplejidad y Ciencia*, *3*(2), 1–16. https://revistasuba.com/index.php/INVESTIGACIONTRANSCOMPLEJIDADYCI/article/view/210/139

Dou, A., Xu, W., Li, X., Zhang, S., & Zhang, J. (2025). Artificial Intelligence in Language Learning: Bridging Gaps, Revealing Patterns, and Charting the Future. *International Journal of Distance Education Technologies*, *23*(1). https://doi.org/10.4018/IJDET.385045

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, *61*(3), 460–474. https://doi.org/10.1080/14703297.2023.2195846

George, D., & Mallery, P. (2016). *IBM SPSS Statistics 23 Step by Step A Simple Guide and Reference fourteenth edition* (Taylor & Francis, Ed.). www.routledge.com/9780134320250

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

Lee, J., & Ko, Y. (2026). AI versus humans: Impact of IPA, AI apps, and peer interactions on the English skills of EFL adolescents. *System*, *140*, 104040. https://doi.org/10.1016/J.SYSTEM.2026.104040

Liu, W. (2024). A Systematic Review of Automated Writing Evaluation Feedback: Validity, Effects and Students' Engagement. *Language Teaching Research Quarterly*, *45*, 86–105. https://doi.org/10.32038/ltrq.2024.45.05

Medina, R. M., Rojas, L. R., Bustamante, H. W., Loaiza, C. R., Martel, C. C., & Castillo, A. R. (2023). *METODOLOGÍA DE LA INVESTIGACIÓN TÉCNICAS E INSTRUMENTOS DE INVESTIGACIÓN* (Wilson Sucari, Patty Aza, & Antonio Flores, Eds.; Instituto Universitario de Innovación Ciencia y Tecnología Inudi Perú S.A.C. https://doi.org/10.35622/inudi.b.080

Mills, G. E., & Jordan, A. W. (2022). *Educational Research: Competencies for Analysis and Applications* (13th ed.). https://www.pearson.com/en-us/subject-catalog/p/educational-research-competencies-for-analysis-and-applications/P200000001512/9780137534944?srsltid=AfmBOopLZZuTcNTfhvJnuVdkLKw-WZK9cJur8LTssmCgdchWk742MTBP

Ossa, C., & Willatt, C. (2023). Providing academic writing feedback assisted by Generative Artificial Intelligence in initial teacher education contexts. *European Journal of Education and Psychology*, *16*(2). https://doi.org/10.32457/EJEP.V16I2.2412

Párraga Solórzano, R. J., Acosta Isidor, Z., Ricardo Suárez, J., & Limones Borbor, J. (2024). Impacto de las estrategias metodológicas en el aprendizaje del idioma inglés en estudiantes de nivel elemental en comunidades de la provincia de Santa Elena. *LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades*, *5*(5). https://doi.org/10.56712/LATAM.V5I5.2766

Román, M. G. M., Ibarra, B. S. N., Jordán, E. R. P., Chiluisa, E. A. P., Ibarra, R. E. N., & Pérez, J. C. M. (2026). La inteligencia artificial generativa y la formación del juicio crítico en la adolescencia. *Revista Veritas de Difusão Científica*, *7*(2), 1555-1579.

Secretaría Nacional de Planificación. (2025, August). *PLAN NACIONAL DE DESARROLLO 2025-2029*.

https://intranet.msp.gob.ec/images/Documentos/Descargas/013%20PlanNacionalDeDesarrollo25-29_EcuadorNoSeDetiene_compressed-1.pdf

UNESCO. (2024). Guía para el uso de IA generativa en educación e investigación. *Unesco*, 1–49. https://unesdoc.unesco.org/ark:/48223/pf0000389227.locale=en

Venegas, M. V. L., Esquivel, G. J., Gonzales, B. M. T., & Altamirano, J. Y. (2025). Validez y confiabilidad en investigaciones en ciencias sociales y políticas. Análisis comparativo en los enfoques cuantitativo, cualitativo y mixto. *Cuestiones Políticas*, *43*(83), 97–121. https://doi.org/10.5281/ZENODO.17757052

Vizcaíno, Z. P. I., Cedeño, C. R. J., & Maldonado, P. I. A. (2023). Metodología de la investigación científica: guía práctica. *Ciencia Latina Internacional*, *7*(4), 9723–9760. https://ciencialatina.org/index.php/cienciala/article/view/7658/11619

Yan, Y., Liu, H., & Chau, T. (2025). A Systematic Review of AI Ethics in Education: Challenges, Policy Gaps, and Future Directions. *Journal of Global Information Management*, *33*(1). https://doi.org/10.4018/JGIM.386381

Zhai, Y., & Nezakatgoo, B. (2025). Evaluating AI-Powered Applications for Enhancing Undergraduate Students' Metacognitive Strategies, Self-Determined Motivation, and Social Learning in English Language Education. *Scientific Reports*, *15*(1), 35199. https://doi.org/10.1038/S41598-025-19118-Z