

Use of ChatGPT Assisted Feedback to Improve English Writing among students of First year of Bachillerato

Uso de la retroalimentación asistida por ChatGPT para mejorar la escritura en inglés en estudiantes de primer año de Bachillerato

Raúl Fernando Coyaguillo Anchapaxi ^[0009-0005-3510-1834] José Luis Pomavilla Patiño ^[0009-0005-3510-1834]

Mauro David Villacrés Benalcázar ^[0009-0008-2111-1347]

Investigadores Independientes, Ecuador.

rafer001@hotmail.com jpomavilla@yachaytech.edu.ec mwillacres@yachaytech.edu.ec

CITA EN APA:

Coyaguillo Anchapaxi, R. F., Pomavilla Patiño, J. L., & Villacrés Benalcázar, M. D. (2026). Use of ChatGPT Assisted Feedback to Improve English Writing among students of First year of Bachillerato. *Tesla Revista Científica*, 6(2). <https://tesla.puertomaderoeditorial.com.ar/index.php/tesla/article/view/758>

Recibido: 2026-08-15

Aceptado: 2026-10-02

Publicado: 2026-10-27

TESLA

Revista Científica

ISSN: 2796-9320



Los contenidos de este artículo están bajo una licencia de Creative Commons Attribution 4.0 International (CC BY 4.0)

Los autores conservan los derechos morales y patrimoniales de sus obras.

The contents of this article are under a Creative Commons Attribution 4.0 International (CC BY 4.0) license. The authors retain the moral and patrimonial rights of their works.

Abstract: This educational case study examined changes in English writing performance among first-year Bachillerato students in the context of ChatGPT-assisted feedback. The study involved 171 students from Nelson Isauro Torres Educational Unit in Cayambe, Ecuador, during the 2026–2027 academic year. A quantitative pre-test–post-test approach was used to compare writing performance across four dimensions: grammar, vocabulary, organization and coherence, and task completion. Data were analyzed using descriptive statistics and individual score comparisons. The findings revealed an increase in the overall mean score from 6.27 to 7.06 points. Furthermore, 72.5% of students improved their performance, while 8.8% maintained their scores and 18.7% experienced a decrease. Organization and coherence showed the greatest average improvement, whereas grammar remained the lowest-scoring dimension. These findings indicate positive changes in writing performance while highlighting differences in individual progress. Although the absence of a control group prevents causal conclusions regarding ChatGPT, the study provides contextualized evidence supporting further investigation into technology-assisted feedback, teacher guidance, and individualized writing instruction in secondary education settings where English is learned as a foreign language.

Keywords: ChatGPT; Artificial intelligence; Formative feedback; Secondary education.

Resumen: El presente estudio de caso educativo examinó los cambios en el desempeño de la escritura en inglés de estudiantes de primer año de Bachillerato, en el contexto de la retroalimentación asistida por ChatGPT. Participaron 171 estudiantes de la Unidad Educativa Nelson Isauro Torres, ubicada en Cayambe, Ecuador, durante el año lectivo 2026–2027. Se empleó un enfoque cuantitativo con evaluaciones pretest y posttest para comparar cuatro dimensiones: gramática, vocabulario, organización y coherencia, y cumplimiento de la tarea. Los datos se analizaron mediante estadística descriptiva y comparaciones individuales. Los resultados evidenciaron un incremento del promedio general de 6,27 a 7,06 puntos. Asimismo, el 72,5 % de los estudiantes mejoró su desempeño, el 8,8 % mantuvo sus calificaciones y el 18,7 % presentó una disminución. La organización y coherencia registraron la mayor mejora promedio, mientras que la gramática continuó siendo la dimensión con menor puntuación. Estos hallazgos reflejan avances en la escritura y diferencias en el progreso individual. Aunque la ausencia de un grupo de control impide establecer relaciones causales con ChatGPT, el estudio aporta evidencia contextualizada para futuras investigaciones sobre retroalimentación tecnológica, orientación docente y enseñanza personalizada de la escritura en inglés.

Palabras clave: ChatGPT; Inteligencia artificial; Retroalimentación formativa; Educación secundaria.

1. Introduction

Developing effective writing skills in English remains a significant challenge for students learning English as a Foreign Language (EFL), particularly during secondary education, when learners are expected to communicate ideas with increasing accuracy, coherence, and independence. Writing involves more than applying grammatical rules or selecting appropriate vocabulary; it requires students to organize their thoughts, establish meaningful connections between ideas, and communicate information clearly. These demands can become particularly challenging for learners with limited language proficiency, insufficient writing practice, or difficulties identifying and correcting their own mistakes. Ekizoglu and Demir (2025) emphasize that developing writing competence among secondary school students requires opportunities for continuous practice and individualized feedback, both of which are essential for improving linguistic accuracy and overall written communication.

Within this educational context, feedback plays a fundamental role in helping students recognize their strengths, understand their difficulties, and progressively improve their written work. However, providing timely and personalized feedback to every student can be demanding for teachers, especially in classrooms with diverse learning needs and different levels of English proficiency. Traditional correction practices may identify grammatical or vocabulary errors without always providing sufficient opportunities for students to understand those errors and apply corrections independently. According to Guo and Wang (2024), integrating ChatGPT into the feedback process offers opportunities to support teachers in responding to learners' writing needs. Nevertheless, technological assistance should complement rather than replace the teacher's pedagogical guidance, since meaningful learning depends on how students interpret, evaluate, and apply the feedback they receive.

The emergence of generative artificial intelligence has introduced new possibilities for teaching and learning English, particularly through tools capable of producing immediate responses and suggesting improvements to written texts. Among these technologies, ChatGPT has attracted increasing attention because of its ability to provide explanations, identify grammatical inaccuracies, recommend vocabulary alternatives, and support text revision. Mahapatra (2024), through a mixed-methods intervention study involving university-level English learners, identified improvements in academic writing performance associated with the use of ChatGPT as a formative feedback tool. Similarly, Polakova and Ivenz (2024) reported improvements in several aspects of EFL students' written production, including grammar, conciseness, and the inclusion of relevant information, following exposure to ChatGPT-generated feedback. These findings suggest that artificial intelligence can contribute to writing instruction when students are encouraged to engage actively with the revision process rather than merely accept automated corrections.

Recent research has also highlighted the importance of examining the different dimensions of writing performance rather than focusing exclusively on overall scores. Parra Núñez and Castrillo de Larreta-Azelain (2025), in a systematic review of 30 empirical studies, identified improvements associated with ChatGPT feedback in linguistic accuracy, vocabulary use, text organization, and overall writing quality. However, the authors also identified challenges related to students' excessive reliance on artificial intelligence and emphasized the importance of teacher supervision and critical engagement with automated suggestions. These considerations are particularly relevant in secondary education, where students are still developing the linguistic knowledge, confidence, and autonomy required to evaluate the quality of their own writing.

Although evidence supporting artificial intelligence-assisted writing instruction continues to grow, the application of these technologies in secondary school classrooms requires further contextualized investigation. Ekizoglu and Demir (2025), in a study involving 60 secondary-level EFL students, found that learners receiving AI-assisted writing feedback demonstrated greater improvements in overall writing performance than students receiving traditional teacher feedback alone. Nevertheless, findings from specific educational settings cannot automatically be generalized to other contexts, particularly when students differ in language proficiency, access to technological resources, and previous experience with artificial intelligence. Consequently, examining classroom-based experiences can provide valuable insights into the opportunities and limitations of integrating ChatGPT into English writing instruction.

In Ecuador, English writing instruction represents an important area for strengthening students' communicative competence during Bachillerato. The present study focuses on first-year Bachillerato students at Nelson Isauro Torres Educational Unit, located in Cayambe, Ecuador, during the 2026–2027 academic year. The educational case involves 171 students whose writing performance was assessed through pre-test and post-test activities. The assessment considered four dimensions: grammar, vocabulary, organization and coherence, and task completion. These criteria provide a structured basis for examining changes in students' written production and identifying the areas in which progress or continuing difficulties can be observed.

Accordingly, the objective of this study is to examine changes in English writing performance among first-year Bachillerato students before and after a classroom intervention involving ChatGPT-assisted feedback, considering grammar, vocabulary, organization and coherence, and task completion. Through an educational case study supported by quantitative comparisons of pre-test and post-test results, the research seeks to contribute context-specific evidence to discussions about the pedagogical use of generative artificial intelligence in secondary-level EFL writing instruction. Rather than assuming that technological integration alone guarantees learning improvement, the study recognizes the importance of

teacher guidance, meaningful writing practice, and critical engagement with feedback in supporting students' writing development.

2. Theoretical Background

The integration of artificial intelligence into English language teaching has created new opportunities to support students in developing their writing skills. However, effective writing instruction requires more than access to digital tools. Students need opportunities to express their ideas, receive meaningful feedback, reflect on their mistakes, and gradually become more independent writers. In this context, ChatGPT-assisted feedback can be understood as a pedagogical resource that supports the writing process while preserving the essential role of teachers in guiding student learning.

2.1. English Writing Skills in Secondary Education

Writing is an essential component of English language learning because it allows students to communicate ideas, organize information, and demonstrate their understanding of linguistic structures. Unlike spontaneous oral communication, writing requires learners to consider several elements simultaneously, including grammatical accuracy, vocabulary selection, logical organization, and the clarity of the intended message. For students learning English as a foreign language, coordinating these elements can be particularly demanding, especially when their exposure to English outside the classroom is limited.

At the secondary education level, writing activities should encourage students to move beyond isolated sentences toward the production of meaningful and coherent texts. Ekizoglu and Demir (2025) highlight the importance of supporting secondary school students through feedback practices that help them recognize linguistic difficulties and improve their written performance. Their research provides a relevant perspective on the value of structured feedback for students who are still consolidating foundational English writing skills.

Similarly, Rahimi et al. (2025) examined automated written corrective feedback in EFL learning and emphasized the importance of understanding how students interact with feedback during writing development. Their findings support the view that improving writing performance involves more than correcting individual errors; learners must also develop the ability to recognize weaknesses, revise their work, and make informed linguistic decisions.

In this context, grammar, vocabulary, organization and coherence, and task completion represent complementary dimensions of writing competence. Grammar contributes to the accuracy of sentences, while vocabulary enables students to express ideas with greater precision. Organization and coherence help readers follow the development of a text, whereas task completion reflects the student's ability to respond

appropriately to the writing assignment. Considering these dimensions together provides a more comprehensive understanding of students' progress and the difficulties that may persist throughout the learning process.

2.2. Feedback in English Language Learning

Feedback is a fundamental element of writing instruction because it creates opportunities for learners to understand what they have achieved and identify areas requiring improvement. When students receive clear and constructive comments, they can reconsider their linguistic choices, correct misunderstandings, and gradually strengthen their confidence as writers. Nevertheless, the educational value of feedback depends not only on its accuracy but also on whether students understand the suggestions and use them meaningfully.

Guo and Wang (2024) examined the potential of ChatGPT to support teacher feedback in EFL writing and demonstrated the relevance of considering artificial intelligence as part of a broader instructional process. Their research highlights the importance of maintaining teacher involvement when technological tools are introduced into writing activities. This perspective is especially relevant in school settings, where students may require additional explanations to understand grammatical corrections or determine whether a suggested expression accurately communicates their intended meaning.

Li and Han (2024) further explored ChatGPT-supported teacher feedback in the EFL context, contributing evidence about how technological assistance can be incorporated into writing instruction. Their findings suggest that combining automated suggestions with pedagogical guidance rather than treating feedback as a process limited to identifying mistakes. Students benefit when they are encouraged to examine the reasons behind corrections and consider alternative ways of communicating their ideas.

Furthermore, Yang et al. (2024) investigated learners' interactions with automated writing evaluation feedback, drawing attention to the role of student engagement in the revision process. Their research suggests that the effectiveness of automated feedback cannot be evaluated solely by the number of corrections generated. It is also necessary to consider how learners interpret, accept, reject, or modify the suggestions they receive.

For secondary school students, this distinction is particularly important. Feedback should be understandable, encouraging, and appropriate to their English proficiency level. Excessively complex explanations or extensive corrections may create confusion rather than promote improvement. Therefore, feedback practices should prioritize clarity, gradual development, and active student participation, allowing learners to view writing mistakes as opportunities for improvement rather than simply as indicators of poor performance.

2.3. ChatGPT as a Tool for Writing Improvement

The introduction of ChatGPT into English language classrooms has expanded the possibilities for personalized writing instruction. Unlike conventional correction methods, ChatGPT can generate immediate responses, explain grammatical structures, suggest alternative vocabulary, and provide recommendations for improving the organization of written texts. These features may be especially valuable for students who need additional guidance during the writing process or who feel uncertain about expressing their ideas in English. However, the educational benefits of this technology depend on how it is integrated into classroom activities and whether students remain actively involved in producing and revising their own work.

Mahapatra (2024) examined the impact of ChatGPT on English writing through a mixed-methods intervention study and reported improvements in students' academic writing performance. These findings support the potential of generative artificial intelligence as a source of formative feedback. Similarly, Polakova and Ivenz (2024) identified improvements in EFL students' writing following the use of ChatGPT-generated feedback. Although these studies were conducted in educational contexts that differ from secondary school settings in Ecuador, they offer relevant evidence for understanding how automated feedback may contribute to the development of writing competence.

Beyond correcting grammatical errors, ChatGPT can encourage students to reconsider the way they structure and communicate their ideas. Ding et al. (2025) found that students receiving ChatGPT feedback demonstrated improvements in several aspects of academic writing, particularly grammar, mechanics, and overall writing quality. Their findings also indicated that combining artificial intelligence feedback with instructor guidance could support broader improvements across writing dimensions. This reinforces the importance of considering ChatGPT as a complementary learning resource rather than an independent substitute for classroom instruction.

From the students' perspective, the usefulness of ChatGPT is also related to their confidence, motivation, and willingness to revise their written work. Teng (2024), in a study involving 45 EFL learners, explored students' experiences with ChatGPT-generated writing feedback and identified positive perceptions regarding motivation, self-efficacy, and engagement. Nevertheless, the study also documented concerns about using the technology, highlighting the need to establish clear expectations and maintain students' responsibility for their learning.

For first-year Bachillerato students, these considerations suggest that ChatGPT-assisted feedback should be implemented through activities appropriate to their age and linguistic proficiency. Short paragraph-writing exercises, explanations of common grammatical errors, vocabulary recommendations, and opportunities for rewriting can create a learning environment in which students gradually strengthen

their writing abilities. The teacher's role remains essential in selecting suitable activities, clarifying feedback, and encouraging students to apply corrections through their own efforts.

3. Case Presentation and Methodology

This educational case study examines the integration of ChatGPT-assisted feedback into English writing instruction among first-year Bachillerato students in Ecuador. The study focuses on changes in students' writing performance between two assessment moments, considering four fundamental dimensions: grammar, vocabulary, organization and coherence, and task completion.

The methodological approach prioritizes the analysis of authentic classroom assessment records, allowing the educational experience to be examined within its natural teaching and learning environment. Rather than considering artificial intelligence as an isolated solution, the study approaches writing development as a process that involves linguistic knowledge, continuous practice, formative assessment, and pedagogical guidance.

3.1. Educational Context and Case Description

The study was situated at Nelson Isauro Torres Educational Unit, located in Cayambe, Ecuador, during the 2026–2027 academic year. The institution provides secondary education to students with diverse academic backgrounds and learning needs. The educational experience was conducted within the English language subject, focusing specifically on the development of writing skills among students enrolled in the first year of Bachillerato.

The case involved students attending the afternoon school session under the academic responsibility of the English teacher, Raúl Fernando Coyaguillo Anchapaxi. According to the institutional records, the teacher was responsible for 171 students distributed across five classroom groups. The students' reported age was 15 years, corresponding to a stage of secondary education in which learners continue developing their grammatical knowledge, vocabulary, and ability to communicate ideas through written English.

The selection of this educational setting was relevant because writing instruction at this level requires students to move from basic sentence construction toward more organized and meaningful written communication. In this context, examining the use of ChatGPT-assisted feedback allows an examination of how technology-supported learning may contribute to the development of essential writing abilities in an authentic classroom environment.

Table 1. General Characteristics of the Educational Case

Characteristic	Description
Educational institution	Nelson Isauro Torres Educational Unit
Location	Cayambe, Ecuador
Academic year	2026–2027
Educational level	First Year of Bachillerato
School session	Afternoon
Number of students	171
Reported student age	15 years
Number of groups	Five (A–E)
Subject	English as a Foreign Language
Main skill assessed	English writing
Assessment	Pre-test and post-test

Note. Information obtained from institutional records provided by the English teacher.

The institutional characteristics summarized in Table 1 establish the contextual boundaries of the case study. The inclusion of five classroom groups offers a broader view of writing performance within this specific educational setting. Nevertheless, the participants represent a defined group of students rather than a randomly selected sample of the wider secondary school population. Consequently, the findings should be interpreted in relation to their educational context.

3.2. Participants and Study Design

The study included 171 first-year Bachillerato students from five existing classroom groups. The participant distribution was obtained from the institutional student registration records, which identified 34 students in Group A, 35 in Group B, 32 in Group C, 35 in Group D, and 35 in Group E.

Table 2. Distribution of Participants by Classroom Group

Classroom group	Students	Percentage
First Year A	34	19.9%
First Year B	35	20.5%
First Year C	32	18.7%
First Year D	35	20.5%
First Year E	35	20.5%
Total	171	100%

Note. Percentages are rounded to one decimal place.

The research was organized as an educational case study incorporating a quantitative, within-group pre-test–post-test comparison. This design enabled the examination of differences in students' writing scores across two assessment moments using the same four evaluation dimensions.

The study did not include a documented control group or random assignment of participants. For this reason, the analysis focuses on describing changes in writing performance rather than establishing a definitive causal effect of ChatGPT-assisted feedback. This distinction is essential for interpreting the findings responsibly, since students' progress may also be influenced by regular instruction, writing practice, previous knowledge, and other classroom experiences.

The use of existing academic records also provides a practical methodological advantage: the analysis reflects writing performance documented in the educational setting rather than responses generated specifically to support the researcher's expectations.

3.3. Assessment Instruments

Students' writing performance was assessed using an English writing pre-test and an analytic assessment rubric developed for first-year Bachillerato. The original pre-test, entitled My School Day, required students to write a short paragraph of approximately 50–70 words describing their daily school routine. The activity encouraged students to include information about their morning activities, school attendance, favorite subjects, experiences at school, and after-school routines. The instructions emphasized the use of the Simple Present tense, the verb to be, and basic connectors such as and, but, because, then, and after that.

The assessment lasted 30 minutes and was completed individually. During the pre-test, students were explicitly instructed not to use ChatGPT, translators, dictionaries, or assistance from other individuals. This condition was intended to capture their independent writing performance before the instructional experience. To evaluate the written paragraphs, an analytic rubric was applied using four criteria, each worth a maximum of 2.5 points, for a total possible score of 10 points.

Table 3. Analytic Rubric for English Writing Assessment

Assessment criterion	Description	Maximum score
Grammar	Accuracy in Simple Present and use of to be	2.5
Vocabulary	Appropriateness and variety of vocabulary	2.5
Organization and Coherence	Logical order of ideas and use of connectors	2.5
Task Completion	Relevance, required length, and completeness	2.5
Total	Overall writing performance	10.0

Note. Adapted from the original analytic writing rubric supplied for the educational case.

Each criterion was assessed through performance descriptors corresponding to 2.5, 2.0, 1.0, or 0.5 points. The instrument also allowed a score of zero when there was no relevant or assessable evidence for a particular criterion.

This analytical approach was particularly appropriate because it enabled the identification of specific strengths and weaknesses in students' writing rather than relying exclusively on a general score. For example, a student could demonstrate adequate vocabulary knowledge while continuing to experience difficulties organizing ideas or completing the writing task.

From a pedagogical perspective, the rubric provides a meaningful connection between classroom assessment and formative feedback. By separating writing performance into four dimensions, teachers can identify the areas in which learners require additional guidance and develop activities that respond to their specific needs.

3.4. ChatGPT-Assisted Feedback Intervention

The educational experience was designed around the use of ChatGPT-assisted feedback as a potential strategy for supporting students' English writing development. The pedagogical rationale was to connect writing practice with opportunities for identifying errors, understanding linguistic difficulties, and improving written communication.

Within this framework, ChatGPT may support students by providing grammatical explanations, suggesting vocabulary alternatives, and offering recommendations for organizing ideas more clearly. Nevertheless, the educational purpose of such feedback is not simply to produce a corrected paragraph, but to help learners understand why particular changes are necessary and how those improvements can be transferred to future writing activities.

For this reason, the role of the English teacher is particularly relevant. Artificial intelligence-generated suggestions require pedagogical interpretation, especially when working with students who are still developing foundational language skills. Teacher guidance can help learners distinguish useful corrections from inappropriate suggestions and encourage them to maintain ownership of their writing.

The intervention is understood conceptually through three interconnected components: student writing, ChatGPT-assisted feedback, and teacher-supported reflection and revision. However, the available records do not establish the exact duration, sequence of activities, frequency of ChatGPT use, or specific feedback prompts employed during classroom instruction.

Figure 1. Proposed Framework for ChatGPT-Assisted Writing Assessment and Feedback

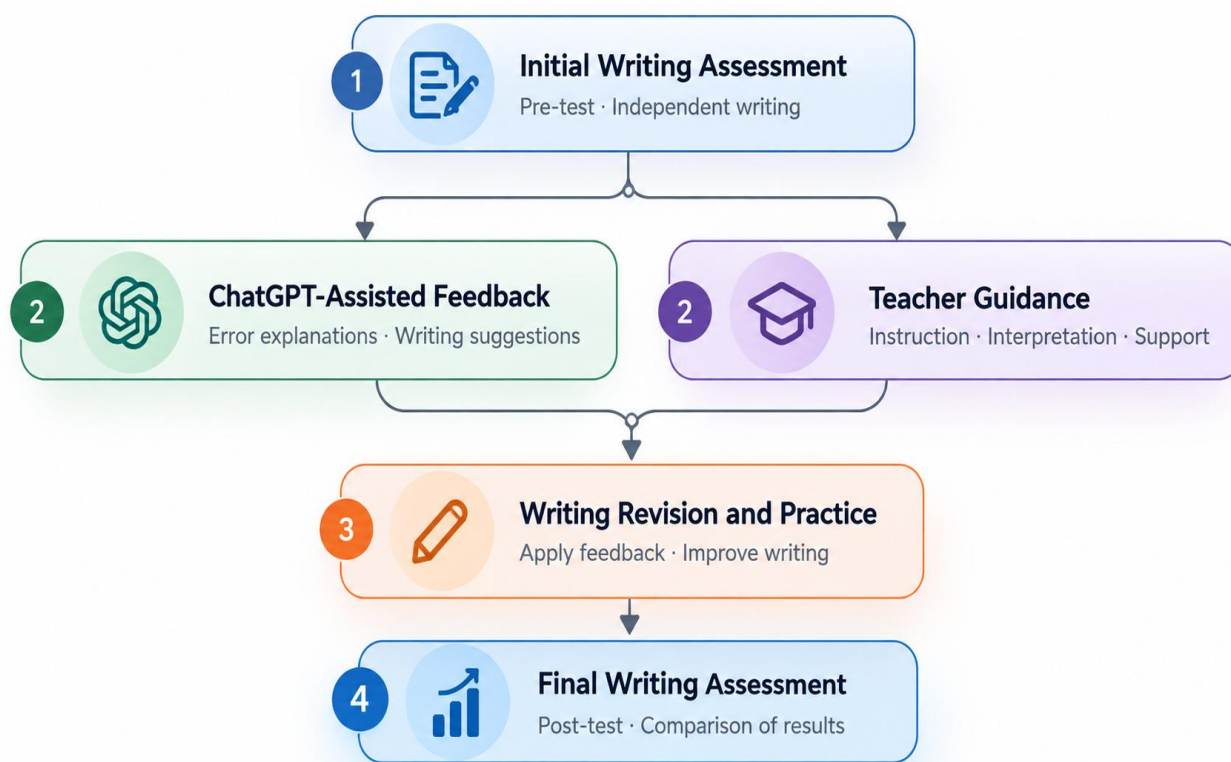


Figure 1 presents the proposed relationship between writing assessment, technological feedback, teacher guidance, and student revision. The framework emphasizes that meaningful writing development requires students to participate actively in identifying and correcting their mistakes. It also highlights the complementary role of artificial intelligence and pedagogical support in creating opportunities for reflection and improvement.

3.5. Data Collection and Analysis

Quantitative data were obtained from the pre-test and post-test writing assessment records provided for the 171 participating students. Each student's performance was documented using an anonymous numerical identifier, allowing the comparison of results across the two assessment moments without including personally identifiable information in the analytical dataset.

The database contained individual scores for grammar, vocabulary, organization and coherence, and task completion. Each dimension contributed a maximum of 2.5 points to the overall writing score, resulting in a total possible score of 10 points.

The analytical procedure focused on three complementary levels of assessment. First, overall writing performance was examined using the mean, median, and standard deviation of total pre-test and post-test scores. Second, the four assessment dimensions were compared to identify areas of improvement or continuing difficulty. Third, individual score differences were examined to classify student progress according to whether scores increased, remained unchanged, or decreased.

Changes in writing performance were calculated by subtracting each pre-test score from its corresponding post-test score:

$$\Delta W = W_{\text{post}} - W_{\text{pre}}$$

where (ΔW) represents the change in writing performance, (W_{post}) is the post-test score, and (W_{pre}) is the pre-test score.

Positive values indicated improvement, zero values indicated no change, and negative values indicated a decrease in performance.

Table 4. Summary of the Data Analysis Procedure

Analytical component	Procedure	Purpose
Overall performance	Mean, median, standard deviation	Describe general writing achievement
Writing dimensions	Comparison of pre-test and post-test means	Identify changes by criterion
Individual progress	Difference between paired scores	Identify improvement, stability or decline
Descriptive distribution	Frequencies and percentages	Describe patterns of student progress

Note. All comparisons were based on the original assessment records. No missing individual scores were identified in the supplied dataset.

The use of descriptive statistics was appropriate for the study's objective because it allowed the identification of changes in writing performance without assuming that all observed differences resulted exclusively from ChatGPT-assisted feedback. Although the dataset supports further inferential analyses, such tests should be selected only after confirming the study design and assumptions.

Overall, this approach provides a transparent and interpretable description of student progress. In addition to examining average performance, it recognizes that learning trajectories may differ among students, an important consideration when evaluating educational interventions in heterogeneous classrooms.

3.6. Ethical Considerations and Methodological Limitations

The analysis was conducted using anonymized academic assessment records. Student names and identification numbers were excluded from the analytical presentation, and individual results were represented through numerical codes. Given that the participants were minors, the publication of this educational experience should be supported by the appropriate institutional authorization and applicable consent or assent procedures.

Several methodological considerations should also be acknowledged. The study examined a single educational institution and did not include a documented comparison group. Furthermore, detailed evidence

of the duration and implementation of the ChatGPT-assisted intervention remains necessary. These characteristics limit the extent to which observed changes can be attributed to artificial intelligence or generalized to other educational populations.

Despite these limitations, the availability of paired assessment scores and an explicit analytic rubric provides a useful foundation for examining writing development in a real secondary school context. The methodological contribution of the case lies in documenting students' performance across identifiable writing dimensions while interpreting the findings within the practical conditions of classroom instruction.

4. Results

The analysis of the pre-test and post-test scores provided evidence of changes in English writing performance among the 171 first-year Bachillerato students at Nelson Isauro Torres Educational Unit. The results were examined using the overall writing scores, the four dimensions of the assessment rubric, and individual patterns of student progress.

Four main findings emerged from the analysis: an increase in overall writing performance, improvements in grammar and vocabulary, progress in organization and task completion, and differences in individual learning outcomes. These findings provide a quantitative description of the changes observed between the two assessment moments, while recognizing that the available data do not establish the exclusive influence of ChatGPT-assisted feedback.

4.1. Overall Improvement in English Writing Performance

The comparison of the pre-test and post-test results revealed an increase in students' overall English writing performance. The mean score increased from 6.27 points in the pre-test to 7.06 points in the post-test, representing an average improvement of 0.79 points on a ten-point scale.

This positive difference suggests that, as a group, students demonstrated stronger writing performance during the second assessment. The change was also reflected in the median score, which increased from 6.10 to 7.05 points. This finding is particularly relevant because it indicates that the improvement was not limited to a small number of students with exceptionally high scores; the central tendency of the overall score distribution also shifted upward.

Table 5. Comparison of Overall English Writing Performance Between Pre and Post test (n = 171)

Statistical indicator	Pre-test	Post-test	Change
Mean score	6.27	7.06	+0.79
Median	6.10	7.05	+0.95

Standard deviation	1.65	1.65	—
Minimum score	1.05	1.10	—
Maximum score	10.00	10.00	—

Note. Scores range from 0 to 10. Values have been rounded to two decimal places.

The results presented in Table 5 indicate a favorable change in the group's average writing achievement. However, the standard deviations remained almost identical across the two assessments, suggesting that differences in performance among students persisted despite the increase in mean scores.

From an educational perspective, this pattern is important because an improvement in the group average does not necessarily imply that all learners progressed at the same rate. While the results reflect overall advancement, they also highlight the need to examine specific writing dimensions and individual trajectories to better understand the nature of student progress.

The observed improvement is consistent with the pedagogical possibility that structured writing practice and feedback may support writing development. Nevertheless, because the study did not include a control group, these results should be interpreted as changes associated with the assessment period rather than as definitive evidence of a causal effect produced by ChatGPT.

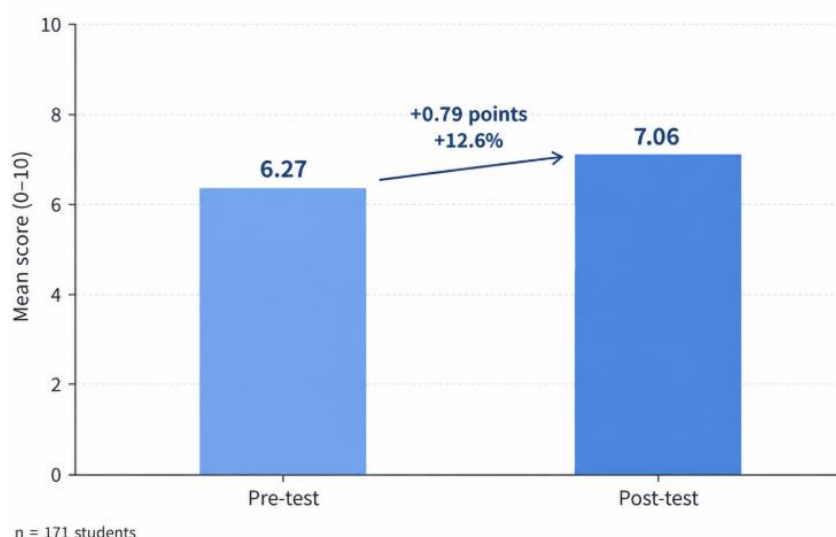


Figure 2. Comparison of mean English writing scores between the pre-test and post-test.
Source: Original assessment database, academic year 2026–2027.

As illustrated in Figure 2, the mean post-test score exceeded the pre-test mean by approximately 0.79 points. Although this difference reflects a positive overall tendency, its educational significance becomes clearer when examined alongside the changes observed in the specific dimensions of writing competence.

4.2. Changes in Grammar and Vocabulary Performance

Grammar and vocabulary are essential components of written communication because they influence students' ability to express ideas accurately and understandably. The analysis of these two dimensions showed improvement between the pre-test and post-test assessments, although the magnitude of the changes varied.

The mean grammar score increased from 1.45 to 1.62 points, representing a gain of 0.17 points. This result suggests modest progress in students' use of grammatical structures, particularly those assessed through the writing task, such as the Simple Present tense and the verb to be. However, grammar remained the lowest-scoring dimension in the post-test, indicating that grammatical accuracy continued to represent an area requiring pedagogical attention.

Vocabulary performance showed a slightly greater increase, rising from 1.55 to 1.74 points, equivalent to an average improvement of 0.19 points. This finding suggests that students were generally more successful in selecting vocabulary appropriate to the writing task during the second assessment.

The comparison of these two dimensions is presented in Table 6.

Table 6. Changes in Grammar and Vocabulary Scores

Writing dimension	Pre-test mean	Post-test mean	Mean change
Grammar	1.45	1.62	+0.17
Vocabulary	1.55	1.74	+0.19

Note. Maximum score per dimension: 2.5 points.

The findings in Table 6 suggest that vocabulary developed slightly more than grammar during the assessment period. One possible interpretation is that selecting or replacing words may be more immediately manageable for beginning EFL writers than consistently applying grammatical rules across an entire paragraph. However, the available numerical records do not allow this explanation to be verified directly.

In pedagogical terms, these findings reinforce the importance of differentiated feedback. Vocabulary development may benefit from opportunities to explore alternative expressions and contextualized word choices, whereas grammatical improvement often requires repeated explanation, application, and independent practice.

Therefore, although both dimensions exhibited positive changes, the comparatively lower performance in grammar indicates the need for continued instructional support that encourages students to understand linguistic structures rather than rely exclusively on suggested corrections.

4.3. Development of Organization, Coherence, and Task Completion

Beyond grammatical accuracy and vocabulary knowledge, effective writing requires students to communicate ideas in a logical sequence and respond adequately to the assigned task. These abilities are particularly relevant in the early stages of English writing development, when students may know individual words and grammatical structures but experience difficulties connecting them into meaningful paragraphs.

The results showed that organization and coherence exhibited the largest mean improvement among the four assessed dimensions. The average score increased from 1.59 points in the pre-test to 1.81 points in the post-test, representing a gain of 0.22 points. This finding suggests that students demonstrated greater ability to structure their written responses and establish connections between ideas during the second assessment.

Task completion also improved, increasing from 1.69 to 1.89 points, with an average gain of 0.20 points. This dimension obtained the highest mean score in both assessments, suggesting that students were generally more successful in responding to the writing assignment than in demonstrating grammatical accuracy or a wide range of vocabulary.

Table 7. Comparison of Organization, Coherence, and Task Completion

Writing dimension	Pre-test mean	Post-test mean	Mean change
Organization and Coherence	1.59	1.81	+0.22
Task Completion	1.69	1.89	+0.20

Note. Scores are expressed on a 2.5-point scale.

The information presented in Table 7 highlights an important aspect of writing development: improvement cannot be evaluated solely through the reduction of grammatical mistakes. The ability to organize information, connect sentences, and address a writing task appropriately is equally essential for effective communication.

From a pedagogical perspective, the improvement in organization and coherence is encouraging because it reflects progress in a dimension that integrates linguistic knowledge with the ability to structure ideas. Nevertheless, the assessment records do not contain the original student paragraphs, so it is not

possible to identify which specific connectors, organizational strategies, or revisions contributed to the observed changes.

The comparatively high task completion scores also suggest that students were generally able to engage with the assigned writing activity. However, fulfilling the requirements of a task does not necessarily guarantee grammatical precision or sophisticated language use. For this reason, teachers should continue encouraging learners to combine adequate task responses with progressively more accurate and coherent written communication.

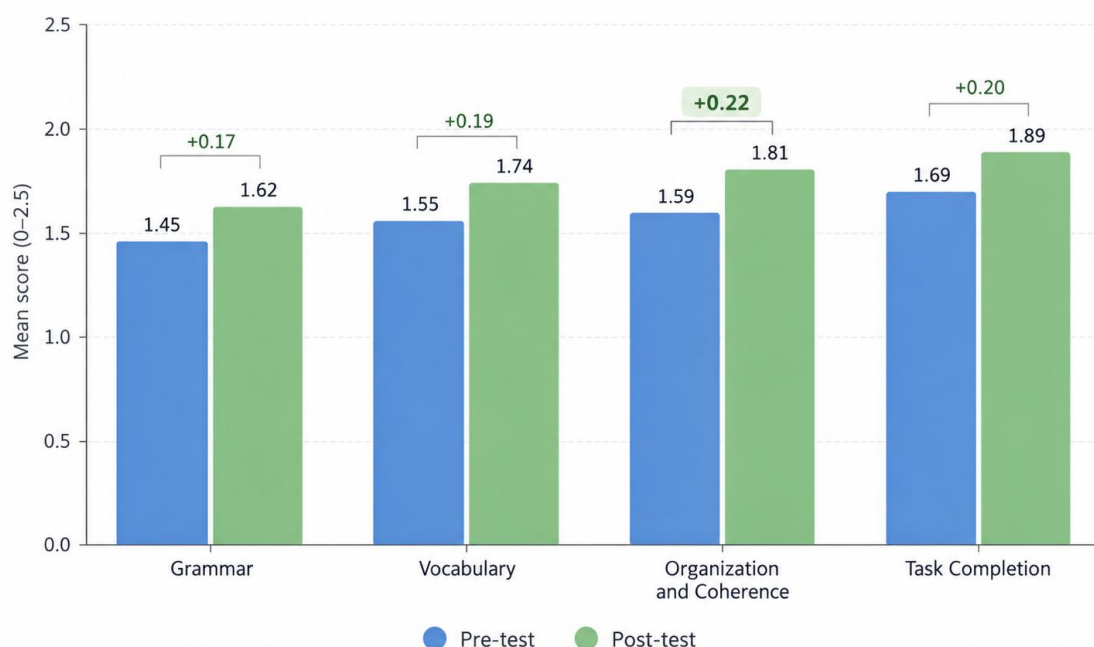


Figure 3. Comparison of mean scores across the four writing assessment dimensions.
Source: Original pre-test and post-test records.

Figure 3 illustrates the positive changes observed in every writing dimension. Although all four criteria increased, their progression was not identical. Organization and coherence presented the largest absolute gain, while grammar remained the dimension with the lowest mean performance. This distinction highlights the importance of examining writing development as a multidimensional process rather than relying exclusively on total assessment scores.

4.4. Individual Differences in Student Writing Progress

Although the overall results indicated improvement in English writing performance, individual score comparisons revealed that students did not experience the same level of progress. This finding is particularly important in educational research because group averages may conceal differences in learning trajectories, previous knowledge, and responses to classroom instruction.

The analysis of the 171 paired records showed that 124 students (72.5%) obtained higher scores in the post-test, while 15 students (8.8%) maintained their initial scores. In contrast, 32 students (18.7%) recorded lower scores in the second assessment.

These differences demonstrate that, despite the favorable overall tendency, improvement was not universal. The results therefore provide a more balanced understanding of the educational experience, identifying both progress and areas requiring further attention.

Table 8. Individual Changes in English Writing Performance

Student progress	Frequency (n)	Percentage
Improved performance	124	72.5%
Unchanged performance	15	8.8%
Decreased performance	32	18.7%
Total	171	100%

Note. Progress was determined by comparing each student's total post-test score with their corresponding pre-test score.

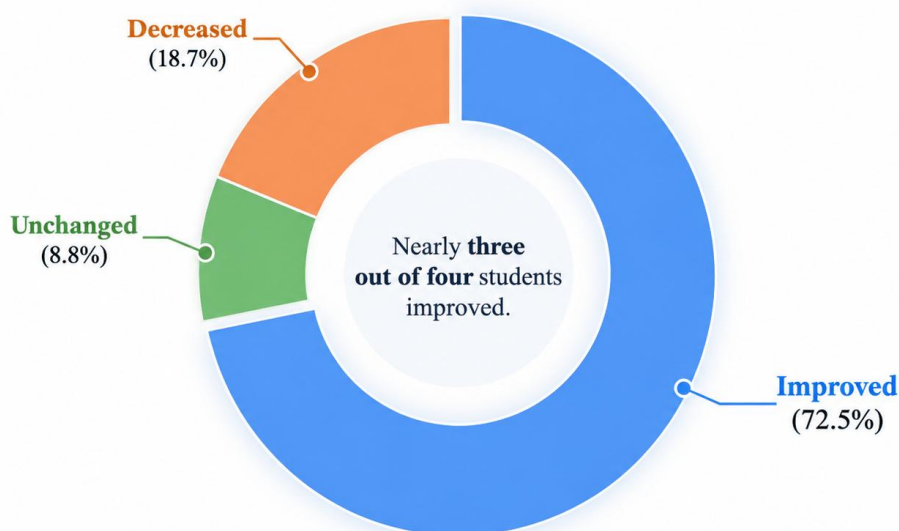


Figure 4. Distribution of Individual Writing Performance Changes

As shown in Table 8 and Figure 4, nearly three out of four students demonstrated an increase in their total writing scores. This represents a relevant finding because it indicates that the positive change in group performance was accompanied by improvement among a substantial proportion of the participating students.

Nevertheless, the presence of 32 students whose scores decreased deserves particular attention. Their results suggest that the educational experience did not produce uniformly favorable outcomes. Differences in previous language proficiency, engagement with writing activities, understanding of

feedback, or classroom learning conditions may have influenced students' performance, although these explanations cannot be confirmed using the available quantitative records alone.

The 15 students whose scores remained unchanged also represent an important group. Stable scores do not necessarily indicate an absence of learning, as students may develop particular abilities without producing changes in their overall assessment results. However, further evidence, such as analysis of their written paragraphs or classroom observations, would be needed to examine this possibility.

From an educational perspective, these findings highlight the importance of individualized instructional support. While technology-assisted feedback may create additional opportunities for writing practice, students' learning needs remain diverse. Therefore, teachers should consider different forms of guidance, additional practice opportunities, and formative assessment strategies to support learners who demonstrate limited progress or declining performance.

Overall, the individual analysis strengthens the interpretation of the findings by moving beyond general averages and recognizing the diversity of students' writing development.

5. Discussion

The findings of this educational case study provide valuable insights into the development of English writing skills among first-year Bachillerato students in a secondary school context in Ecuador. The comparison of pre-test and post-test scores revealed a positive overall change in writing performance, accompanied by improvements in grammar, vocabulary, organization and coherence, and task completion. However, the results also demonstrated differences in individual progress, suggesting that writing development is neither uniform nor exclusively determined by access to technological resources. These findings can be interpreted in relation to recent research on ChatGPT-assisted feedback, while recognizing the methodological limitations of the present study.

One of the most relevant findings was the increase in the overall writing mean score from 6.27 to 7.06 points, representing an improvement of 0.79 points. This positive tendency is consistent with Mahapatra (2024), who reported improvements in academic writing performance among English as a Second Language (ESL) learners following an intervention involving ChatGPT-generated formative feedback. Similarly, Polakova and Ivenz (2024) identified improvements in several aspects of EFL students' writing, including grammatical accuracy, conciseness, and the inclusion of relevant information. Although these investigations were conducted with university students, their findings provide a useful basis for understanding the potential contribution of artificial intelligence to writing instruction.

Nevertheless, the present study offers a different educational perspective because it focuses on students at the beginning of Bachillerato, who are still consolidating essential linguistic structures. In this context, an increase in writing scores should not be understood merely as a numerical achievement. It may also reflect progress in students' ability to communicate familiar experiences, organize basic information, and respond to writing assignments with greater clarity. From a pedagogical standpoint, these abilities are particularly important because they establish the foundations for more complex written communication at later educational stages.

Regarding specific writing dimensions, the results showed improvements in both grammar and vocabulary, although grammar remained the lowest-scoring criterion in the post-test. This finding deserves attention because grammatical development frequently requires students to understand rules and apply them consistently across different communicative situations. Ekizoglu and Demir (2025), in their research involving secondary-level EFL students, reported improvements in writing performance associated with AI-assisted feedback, highlighting its potential to support linguistic development. However, the persistence of grammatical difficulties in the present case suggests that feedback alone may not be sufficient to ensure the consolidation of grammatical knowledge.

In our interpretation, this result highlights the difference between recognizing a grammatical correction and independently applying the underlying rule. A student may understand a suggested change when reviewing a paragraph but continue making similar mistakes when producing a new text without assistance. Consequently, writing instruction should encourage learners to explain corrections, practice grammatical structures in different contexts, and progressively assume responsibility for their written production. ChatGPT-generated explanations may be useful within this process, but their educational value depends on how students engage with them and how teachers reinforce the learning experience.

Another important finding was the improvement in organization and coherence, which increased by 0.22 points and represented the largest mean gain among the four writing dimensions. This result is particularly meaningful because coherent writing requires students to establish relationships between ideas rather than simply construct grammatically acceptable sentences. Guo and Wang (2024) demonstrated that ChatGPT can generate feedback addressing content, organization, and language, suggesting that its potential extends beyond identifying surface-level linguistic errors. Similarly, Li and Han (2024) found that ChatGPT-supported teacher feedback could encourage students to engage more deeply with writing revisions when automated suggestions were incorporated into pedagogical guidance.

These findings offer an important point of comparison with the present case. The increase in organization and coherence may indicate that students developed a better understanding of how to structure their written ideas during the assessment period. However, without access to the original student paragraphs and detailed intervention records, it is not possible to determine whether this improvement resulted from

the use of connecting words, clearer sentence sequencing, teacher explanations, ChatGPT suggestions, or a combination of these experiences. This distinction is essential because meaningful educational research should differentiate observed outcomes from possible explanations.

Task completion also demonstrated positive development, reaching the highest mean score among the four assessed dimensions. This finding suggests that students were generally more successful in responding to the requirements of the writing assignment than in demonstrating complete grammatical accuracy. In practical terms, students may be able to communicate the intended message and include relevant information even when their written texts contain linguistic errors. Such a result supports the importance of evaluating writing through multiple complementary criteria rather than reducing written communication to grammatical correctness alone.

Parra Núñez and Castrillo de Larreta-Azelain (2025), in their systematic review of 30 empirical studies, identified improvements associated with ChatGPT feedback in linguistic accuracy, text organization, and overall writing quality. Their findings support the relevance of examining different aspects of writing performance, as these dimensions may respond differently to feedback and instructional practice. The present case contributes to this discussion by showing that, within the assessed group, improvements were observed across all four dimensions, although their magnitude varied.

An additional finding requiring careful consideration was the variation in individual student progress. Although 72.5% of participants improved their writing scores, 8.8% maintained their initial performance and 18.7% obtained lower post-test scores. These results suggest that positive changes in group averages do not necessarily represent equivalent learning outcomes for every student. From an educational perspective, this diversity is important because learners may differ in their previous knowledge, confidence, engagement, and ability to interpret feedback. Nevertheless, these factors were not directly measured in the present study and should therefore be considered possible explanations rather than confirmed causes.

The observed differences also highlight the importance of maintaining teacher involvement when implementing artificial intelligence in the classroom. Li and Han (2024) emphasized the value of combining automated feedback with teacher guidance, particularly when students require support in understanding and applying suggestions. Similarly, Guo and Wang (2024) identified complementary characteristics in teacher-generated and ChatGPT-generated feedback, indicating that both sources may contribute differently to the writing process. These perspectives reinforce the need to consider technological feedback as part of a broader pedagogical strategy rather than as an independent mechanism for improving academic performance.

In the specific context of first-year Bachillerato, this consideration becomes particularly relevant. Students at this educational level may require feedback that uses accessible language, explains mistakes without overwhelming them, and encourages independent writing. Excessive dependence on automated corrections could potentially limit opportunities for developing self-monitoring and autonomous revision skills. Therefore, the pedagogical objective should not be to obtain error-free paragraphs through artificial intelligence, but to help students progressively produce clearer, more coherent, and more accurate texts through their own understanding.

Despite the encouraging results, several methodological limitations must be acknowledged. The study was conducted in a single educational institution using a one-group pre-test–post-test design, without random assignment or a comparison group. In addition, the available documentation does not fully describe the duration and implementation of the ChatGPT-assisted intervention. Consequently, the observed improvements cannot be attributed exclusively to ChatGPT, since regular classroom instruction, repeated writing practice, familiarity with assessment procedures, and other learning experiences may also have influenced performance. The absence of student interviews, classroom observations, and original written texts further limits the possibility of explaining why some students improved while others did not.

Nevertheless, the study provides useful descriptive evidence from an authentic secondary education setting. Its principal contribution lies in examining writing development among 171 Bachillerato students through an analytic rubric that considers four essential components of written communication. Unlike investigations focused exclusively on university populations, this case draws attention to the importance of studying artificial intelligence-supported writing instruction among younger EFL learners and within specific institutional contexts.

The results indicate that the assessment period was associated with favorable changes in English writing performance, particularly in organization and coherence, while also revealing persistent grammatical difficulties and heterogeneous individual progress. From a pedagogical perspective, the results encourage a balanced approach in which ChatGPT-assisted feedback, when appropriately implemented, supports rather than replaces teacher guidance, independent writing practice, and students' critical engagement with their own learning. Future investigations incorporating control groups, detailed intervention protocols, qualitative evidence, and longer follow-up periods would help clarify the specific contribution of ChatGPT-assisted feedback to sustainable writing development.

6. Conclusions

The present educational case study examined changes in English writing performance among 171 first-year Bachillerato students at Nelson Isauro Torres Educational Unit in Cayambe, Ecuador, within the context of ChatGPT-assisted feedback. The findings revealed an overall improvement in students' writing

scores, with the mean increasing from 6.27 in the pre-test to 7.06 in the post-test. Furthermore, 72.5% of participants demonstrated positive changes in their writing performance. These results provide encouraging evidence of progress during the assessment period and highlight the relevance of exploring innovative strategies to support English language learning in secondary education.

The analysis of the four assessment dimensions demonstrated improvements in grammar, vocabulary, organization and coherence, and task completion. Organization and coherence showed the greatest average increase, suggesting that students became more successful in structuring and connecting their written ideas. Nevertheless, grammar remained the lowest-scoring dimension, indicating that the accurate application of linguistic structures continues to require sustained instructional attention. These findings emphasize that writing development involves multiple interconnected abilities and that improvements in one dimension do not necessarily guarantee equivalent progress in others.

An important contribution of this study is the recognition that students do not progress at the same pace or respond identically to educational experiences. Although most participants improved their writing scores, a proportion maintained or decreased their performance. This variation reinforces the importance of individualized feedback, continuous assessment, and teacher guidance. From a pedagogical perspective, ChatGPT may offer valuable opportunities for explaining linguistic difficulties and supporting text revision; however, its potential should be directed toward strengthening students' understanding and independence rather than encouraging the automatic acceptance of generated corrections.

Finally, the findings should be interpreted considering the limitations of the study, particularly the absence of a control group and the incomplete documentation of the ChatGPT-assisted intervention. Therefore, the observed improvements cannot be attributed exclusively to artificial intelligence. Future research should incorporate clearly documented instructional procedures, comparison groups, student perceptions, and follow-up assessments to determine whether writing improvements are sustained over time. Ultimately, the educational value of ChatGPT lies not simply in helping students produce more accurate texts, but in its potential to support a learning process in which students develop the confidence, critical judgment, and linguistic independence necessary to communicate effectively in English.

FINANCIACIÓN

Los autores no recibieron financiación para el desarrollo de la presente investigación.

CONFLICTO DE INTERESES

Los Autores declaran que no existe conflicto de intereses con su investigación

CONTRIBUCIÓN DE AUTORÍA

En concordancia con la taxonomía establecida internacionalmente para la asignación de créditos a autores de artículos científicos (<https://credit.niso.org/>). Los autores declaran sus contribuciones en la siguiente matriz:

	Autor 1.	Autor 2.	Autor 3.
Participar activamente en:			
Conceptualización	X	X	X
Análisis formal	X	X	X
Adquisición de fondos	X	X	X
Investigación	X	X	X
Metodología	X	X	X
Administración del proyecto	X	X	X
Recursos	X	X	X
Redacción –borrador original	X	X	X
Redacción –revisión y edición	X	X	X
La discusión de los resultados	X	X	X
Revisión y aprobación de la versión final del trabajo.	X	X	X

8. REFERENCIAS BIBLIOGRÁFICAS

Aksakallı, C., & Daşer, Z. (2025). Unlocking EFL learners’ insights into ChatGPT use for L2 writing: The impacts of usage frequency and gender variations. *Current Psychology*, 44, 7957–7977. [10.1007/s12144-025-07437-3](https://doi.org/10.1007/s12144-025-07437-3)

Ding, L., Zou, D., & Kohnke, L. (2025). ChatGPT as an automated writing evaluation tool: How students perceive it and how it affects their writing. *Education and Information Technologies*, 30, 26777–26799. [10.1007/s10639-025-13775-3](https://doi.org/10.1007/s10639-025-13775-3)

Ekizoglu, M., & Demir, A. N. (2025). The role of AI assisted writing feedback in developing secondary students writing skills. *Discover Education*, 4, Article 454. [10.1007/s44217-025-00919-3](https://doi.org/10.1007/s44217-025-00919-3)

Guo, K., & Wang, D. (2024). To resist it or to embrace it? Examining ChatGPT’s potential to support teacher feedback in EFL writing. *Education and Information Technologies*, 29, 8435–8463. [10.1007/s10639-023-12146-0](https://doi.org/10.1007/s10639-023-12146-0)

Jmaiel, H. A., Abukhait, R. O., Mohamed, A. M., Shaaban, T. S., Al-khresheh, M. H., & AL-Qadri, A. H. (2025). The role of ChatGPT in enhancing EFL students’ ESP writing skills: An experimental study of gender and major differences. *Discover Education*, 4, Article 240. [10.1007/s44217-025-00700-6](https://doi.org/10.1007/s44217-025-00700-6)

Li, M., & Han, J. (2024). Exploring ChatGPT-supported teacher feedback in the EFL context. *System*, 126, Article 103502. [10.1016/j.system.2024.103502](https://doi.org/10.1016/j.system.2024.103502)

Lu, D., & Zeng, Y. (2025). Exploring the use of ChatGPT-generated model texts as a feedback instrument: EFL students’ text quality and perceptions. *Innovation in Language Learning and Teaching*. [10.1080/17501229.2025.2525341](https://doi.org/10.1080/17501229.2025.2525341)

Mahapatra, S. (2024). Impact of ChatGPT on ESL students’ academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, 11, Article 9. [10.1186/s40561-024-00295-9](https://doi.org/10.1186/s40561-024-00295-9)

Mekheimer, M. (2025). Generative AI-assisted feedback and EFL writing: A study on proficiency, revision frequency and writing quality. *Discover Education*, Article 170. [10.1007/s44217-025-00602-7](https://doi.org/10.1007/s44217-025-00602-7)

Parra Núñez, R., & Castrillo de Larreta-Azelain, M. D. (2025). The impact of using ChatGPT for feedback in EFL writing: A systematic review. *Revista Internacional de Lenguas Extranjeras*, (23), 95–119. [10.17345/rile23.4221](https://doi.org/10.17345/rile23.4221)

Polakova, P., & Ivenz, P. (2024). The impact of ChatGPT feedback on the development of EFL students’ writing skills. *Cogent Education*, 11(1), Article 2410101. [10.1080/2331186X.2024.2410101](https://doi.org/10.1080/2331186X.2024.2410101)

- Rahimi, M., Fathi, J., & Zou, D. (2025). Exploring the impact of automated written corrective feedback on the academic writing skills of EFL learners: An activity theory perspective. *Education and Information Technologies*, 30, 2691–2735. [10.1007/s10639-024-12896-5](https://doi.org/10.1007/s10639-024-12896-5)
- Teng, M. F. (2024). “ChatGPT is the companion, not enemies”: EFL learners’ perceptions and experiences in using ChatGPT for feedback in writing. *Computers and Education: Artificial Intelligence*, 7, Article 100270. [10.1016/j.caeai.2024.100270](https://doi.org/10.1016/j.caeai.2024.100270)
- Yang, Y., Hussin, S., & Hashim, H. (2025). Revising with intelligence: ChatGPT feedback and its impact on EFL students' revision and self-efficacy. *Forum for Linguistic Studies*, 7(7). [10.30564/fls.v7i7.9845](https://doi.org/10.30564/fls.v7i7.9845)
- Yang, H., Gao, C., & Shen, H.-Z. (2024). Learner interaction with, and response to, AI-programmed automated writing evaluation feedback in EFL writing: An exploratory study. *Education and Information Technologies*, 29, 3837–3858. [10.1007/s10639-023-11991-3](https://doi.org/10.1007/s10639-023-11991-3)