

Article

What's Next for Feedback in Writing Instruction? Pre-Service Teachers' Perceptions of Assessment Practices and the Role of Generative AI

Maria Navío-Inglés ^{1,*}, Jesús Guzmán Mora ¹, Paula O'Connor-Jiménez ² and Almudena García González ³

¹ Department of Hispanic and Classical Philology, Faculty of Education of Albacete, University of Castilla-La Mancha, 02071 Albacete, Spain

² LabinTic, Laboratory of Integration of Technology in Classrooms, Faculty of Education of Albacete, University of Castilla-La Mancha, 02071 Albacete, Spain

³ Department of Hispanic and Classical Philology, Faculty of Letters of Ciudad Real, University of Castilla-La Mancha, 13071 Ciudad Real, Spain

* Correspondence: maria.navio@uclm.es

Abstract

Providing effective feedback on writing remains a challenge for teachers. Among the tools and strategies being explored to address this issue is generative artificial intelligence (AI), due to its potential to deliver immediate and tailored feedback that complements teacher input. However, successful implementation requires understanding the views of those likely to integrate such tools into their future teaching practice. This quasi-experimental study explores pre-service teachers' perceptions of the inclusion of generative AI-generated feedback. A control group received teacher feedback on their writing, while an experimental group received a combination of teacher and AI-generated feedback. After the intervention, participants' views on the feedback received and their preparedness for assessing writing were analyzed. Results reveal more positive perceptions among the experimental group, along with greater confidence in their ability to teach and assess writing. Although both groups acknowledged the need to improve their own linguistic knowledge and assessment strategies, and emphasized the importance of teacher-led feedback, participants in the experimental group also advocated for including AI tools. This highlights the need to strengthen linguistic knowledge and assessment training in teacher education, as well as the positive attitudes and openness towards generative AI among those who have experienced its potential in feedback provision.

Keywords: generative AI; writing; pre-service teachers; perceptions; language



Academic Editor: Lawrence Jun Zhang

Received: 13 October 2025

Revised: 2 November 2025

Accepted: 11 November 2025

Published: 14 November 2025

Citation: Navío-Inglés, M., Guzmán Mora, J., O'Connor-Jiménez, P., & García González, A. (2025). What's Next for Feedback in Writing Instruction? Pre-Service Teachers' Perceptions of Assessment Practices and the Role of Generative AI. *Education Sciences*, 15(11), 1534. <https://doi.org/10.3390/educsci15111534>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Feedback for Writing Skills Development

Writing is a key competence, but its development represents one of the main challenges in the teaching–learning process of language across different educational stages, as reflected in the educational curricula of many countries. This skill is not limited to the orthographic and grammatical encoding of ideas; it also involves the planning, drafting, and revision of a text that must meet demands of coherence, cohesion, appropriateness, and discursive structure (Godwin-Jones, 2022). Consequently, knowing how to write requires simultaneously activating both linguistic knowledge and knowledge related to the process of text creation, which makes writing a highly demanding task for learners.

In this context, feedback becomes an essential component of the improvement process (Graham et al., 2015). Systematic writing practice cannot be separated from continuous guidance that allows students to identify errors, understand the characteristics of well-constructed texts, and revise their work with greater autonomy. The specialized literature has repeatedly emphasized that feedback is one of the most impactful factors in the development of writing skills, not only from a cognitive perspective but also from an affective dimension (Graham et al., 2015; Nunes et al., 2022). Students' perceptions of their own performance, as well as their motivation and self-efficacy to write, are deeply influenced by the feedback they receive (Graham et al., 2015). Moreover, the effectiveness of such feedback can also be affected by how students engage with it (Winstone et al., 2017).

Nevertheless, although the formative value of feedback is widely recognized, its effective implementation in real educational contexts presents several challenges for teachers. Text correction entails a high workload for instructors (Xu & Zhang, 2022). In many cases, the dynamic that unfolds is based on the periodic exchange of written assignments, where the student submits a text and the teacher subsequently corrects it and returns it with annotations or a grade (Ruiz Flores, 2009). This practice, common across various educational levels, relegates feedback to a phase that comes after the writing process, preventing students from incorporating comments while drafting. As a result, feedback becomes a tool for product evaluation rather than a guide to support the process.

Moreover, teachers do not always feel adequately prepared to handle writing evaluation, particularly those with less training in this area, such as pre-service teachers (Latifi & Noroozi, 2021; Ropohl & Rönnebeck, 2019). As noted by Zhu et al. (2025), addressing the correction of students' texts requires mastery of several areas in which pre-service teachers are still being instructed: the content being taught, the pedagogical principles underlying it, and an understanding of students' needs. Specifically, regarding writing, it is essential to address both lower-level aspects, such as spelling, and higher-level ones, such as text structure (Link et al., 2022). However, less experienced teachers tend to focus more on the former and give less attention to more complex writing issues (Junqueira & Payant, 2015), perhaps due to a lack of experience and confidence (Chen, 2023).

Given these challenges, new forms of support that can complement teachers' efforts in text correction and feedback provision regarding students' writing should be explored. Nonetheless, in this search for possible solutions, it is equally important to consider the training and perspectives of the professionals who may be responsible for implementing such forms of support in the future. After all, their beliefs about these tools can influence their attitudes toward writing instruction (Hall, 2016), and therefore, the way they approach it.

1.2. Generative AI in Writing and Language Learning

The popularization of generative artificial intelligence (AI) is making a turning point in the use of technology in education (Robert et al., 2025). Models such as the generative pre-trained transformer (GPT), integrated in tools like ChatGPT and Microsoft Copilot, are capable of producing text using natural language, with an expression increasingly similar to that of humans (Lim et al., 2023). As a result, it is generating significant interest in the field of education (Zhang et al., 2025).

In recent years, its educational potential has begun to be explored, for example, including its use as support for writing instruction (Barrot, 2024). Authors such as Su et al. (2023) suggest that generative AI can intervene at different stages of the writing process, from idea generation to stylistic improvement, as well as linguistic and structural revision. In this regard, some studies highlight its usefulness for facilitating supporting writing process, offering immediate improvement suggestions, and fostering a more reflective approach to writing (Li et al., 2024; Steiss et al., 2024).

Therefore, its potential as a source of immediate feedback during the writing process becomes especially relevant. In contrast to the traditionally delayed feedback that teachers typically provide to a group of students, generative AI enables learners to obtain instant responses as they write, opening new possibilities for integrating real-time revision (Zhu et al., 2025). In line with this, authors such as Escalante et al. (2023) highlight the positive impact that combining both types of feedback can have on students' writing. Furthermore, in the case of pre-service teachers, the role of such tools can extend even further, assisting teachers themselves in the process of generating feedback and learning how to do it. One example of this is found in the already mentioned work of Zhu et al. (2025), in which the introduction of ChatGPT helped pre-service teachers improve both their feedback strategies and the depth of the feedback they provide.

Alongside the potential formative effects of using generative AI to provide feedback, it is particularly important to understand the perceptions of those who will interact with the tool, as these also influence their willingness to use it (Feng et al., 2025). For this reason, studies that explore users' perceptions of generative AI-based tools have begun to emerge. In the study conducted by Kim et al. (2025), a group of university students used a generative AI system to complete academic writing tasks. The results showed that students valued several aspects positively, such as the immediacy of the responses provided by the tool, the guidance offered at different stages of the writing process, its ability to assist with information search and retrieval, and the way it reduced anxiety about making mistakes compared to writing in front of human instructors. However, they also expressed concerns regarding the system's limited contextual understanding, its potential to generate inaccurate or misleading content, and its lack of pedagogical skills. Similarly, Zhan and Yan (2025) found that students experienced mostly positive emotional responses when interacting with generative AI. Yet, their results also revealed differences in the level of critical engagement: while some students approached AI-generated feedback with skepticism, others accepted it without questioning the accuracy or appropriateness of the responses. Finally, the cross-sectional study carried out by Henderson et al. (2025) showed that participants viewed generative AI as a valuable tool for feedback. In spite of it, they continued to perceive teacher-generated feedback as more helpful, trustworthy, personal, and expert.

Thus, the challenges described in relation to the teaching and learning process of writing, as well as the provision of truly effective and formative feedback, have prompted the search for new solutions (Kohnke et al., 2023; Zhang et al., 2025). Recent advances in generative AI position this tool as a resource with strong potential to offer immediate feedback during the writing process, especially when used as a complement to teacher revision (Escalante et al., 2023; Guo & Wang, 2024). However, its successful integration into real educational settings also requires considering the perspectives of those who may implement it in their future teaching practice, but there is still a lack of studies that address this issue (Celik et al., 2022; Moorehouse & Kohnke, 2024). Taking all this into account, the present study aims to explore the perspectives of pre-service primary school teachers regarding text correction practices and their self-perceived preparedness to carry out this task, with particular attention to the possibilities and limitations they identify in the use of generative AI-based tools. To this end, the perceptions of two groups of participants are compared following an educational intervention in which they received different types of feedback. One group received only delayed teacher feedback, while the other received a combination of immediate feedback generated by generative AI and delayed teacher feedback. The objectives of the study are as follows:

- To analyze and compare pre-service teachers' perceptions of delayed teacher feedback and its combination with immediate AI-generated feedback.

- To explore how they assess their preparation to teach and evaluate writing, as well as the strategies they consider most suitable for their future teaching practice, with special attention to the potential integration of generative AI tools.

2. Materials and Methods

A quasi-experimental study with control group (CG) and experimental group (EG) was carried out. The study compares the perceptions of two groups of pre-service teachers who were exposed to different feedback conditions on argumentative texts written in their primary language of instruction and official language of the region. The study received approval from the Social Research Ethics Committee from the University of Castilla-La Mancha (CEIS-2025-103060).

2.1. Participants

The sample consisted of 89 undergraduate students enrolled in a Primary Education Teaching Degree program at the University of Castilla-La Mancha (Spain). Specifically, the sample included 61 women and 21 men, aged between 20 and 25. The unequal distribution of male and female students can be attributed to the historically higher enrollment of women in education-related university programs (Navarro-Guzmán & Casero-Martínez, 2012). All participants were being trained to teach pupils aged approximately 6 to 12 in primary school settings. At this stage of their studies, they had already received training in the teaching of Spanish as a first language, as part of the compulsory courses included in the degree program.

2.2. Instruments

The data were collected through four open-ended questions directly aligned with the research objectives. Prior to their administration, the questions were reviewed by a panel of three experts: one university lecturer with extensive experience in the field of Language and Literature Didactics, and two university professors specializing in educational technology.

Each expert independently received a document providing a brief overview of the study and the initial version of the proposed questions. They were invited to evaluate the questions and suggest improvements where necessary. In doing so, they were asked to consider the relevance and clarity of each question in relation to the study's context and objectives, its appropriateness for the participant profile, and the overall clarity of wording. Based on their feedback, the final versions of the questions were as follows:

1. To what extent do you feel the feedback you received during this study helped you improve your writing? Justify your answer.
2. After participating in the activity, what advantages and disadvantages do you identify in the type of feedback you received?
3. How well prepared do you feel to teach and assess writing in your future teaching practice? What aspects do you think you would need to improve or strengthen?
4. How would you approach the teaching and assessment of writing in your future classes? Would you include any aspects of the methodology used in this study? Justify your answer.

2.3. Procedure

The 89 participants were assigned to the control group (CG) or the experimental group (EG) based on the existing class distribution. This means the CG had 41 students and the EG had 48. Across three weekly sessions, each lasting one hour, all participants were asked to write argumentative essays, regardless of group. To carry out these tasks, the writing instructions were taken from the university entrance exams used in the region where the study was conducted. As they are developed by experts, this ensured an appropriate

and consistent level of difficulty throughout the intervention. Examples of the proposed topics include the effectiveness of anti-smoking measures in the country or the possible replacement of human workers by AI. For each topic, participants were provided with an initial reading on the issue so that they had some background knowledge about the content they would address in their essay.

Students in the CG wrote their texts and submitted them via the university's Moodle platform to the instructors, who provided corrections and returned feedback a few days before the following session, so that students could review it. The EG followed the same process as the CG, but with the additional support of the generative AI tool Microsoft Copilot, which provided immediate feedback to help resolve content and language questions, as well as revise and improve drafts before submission. Therefore, the CG received delayed teacher feedback only, while the EG received a combination of immediate AI-generated feedback and delayed teacher feedback. To ensure proper use of the tool, students in the EG received an introductory session on generative AI in addition to the three writing sessions.

After completing the writing sessions, all 89 participants were given one hour to respond to the four open-ended questions described earlier. The analysis focused exclusively on participants' written perceptions collected through the open-ended questions, rather than on the linguistic quality or performance of the essays produced.

Feedback Details

After each session, two teacher educators applied the same strategy to provide feedback to both groups. First, they used a rubric specifically designed to evaluate each relevant aspect of essay writing. The rubric included three levels for each item, with detailed explanations allowing students to identify weaknesses in areas such as each typical part of an essay, coherence, cohesion, appropriateness, lexical precision, grammatical accuracy and spelling, and content. Therefore, although they are expected to be familiar with this type of text from their prior education, they were made aware of what was expected for the writing task. Additionally, the rubric featured a comments section where instructors could include any supplementary observations they deemed relevant. Besides that, instructors also provided direct annotations on the students' drafts. These annotations primarily focused on identifying spelling and grammar errors, as well as any issues that were more effectively addressed directly in the text rather than through the rubric.

This combined approach aimed to standardize the correction process without sacrificing personalized feedback. The use of a widely adopted instructional tool such as a rubric, along with direct annotations, helped ensure that the workload remained manageable for instructors while still offering feedback tailored to each student's writing needs.

Regarding the generative AI feedback, a specific set of interaction guidelines was provided to students. First, they were required to draft an initial version of their essay without any assistance. Then, they interacted with Copilot to receive feedback and used it to produce a final version of the text. To guide the interaction, students were given an initial prompt designed to define the tool's role and generate feedback aligned with the aspects evaluated in their essays. After this initial input, students were allowed to ask additional questions freely, provided they did not prompt the AI to generate a full text. Once they had revised their drafts based on the AI's suggestions, they submitted their final version for teacher evaluation. This procedure ensured that students engaged in the core cognitive effort required to produce their own texts and that AI was introduced only as a complementary support, not as a substitute for student work, nor as a replacement for teacher assessment.

2.4. Data Analysis

The responses to the four open-ended questions were analyzed using qualitative content analysis with a hybrid approach combining deductive and inductive codification. Deductive coding categories based on core dimensions of writing competence and pedagogical preparedness with inductive categories that emerged from participants' responses were adopted. Coding was carried out using the software Atlas.ti (v.9). A first coder conducted an initial inductive coding process by identifying meaning units and grouping them into emergent codes based on the data. Subsequently, a second coder reviewed the initial coding, suggested adjustments, and, in agreement with the first coder, a final coding scheme was established. This scheme was then consistently applied across the dataset to ensure the reliability and validity of the analysis. Also, the qualitative analysis was complemented by frequency counts and percentage descriptions to support the interpretation of trends across groups. Thus, their inclusion is aimed at contextualizing the contrast trends between the two groups.

The four open-ended questions were analyzed separately, as each addressed different dimensions of the study and responded to specific research objectives. This meant that the set of codes for each question varied according to its focus, while codes referring to shared aspects were kept consistent across questions to ensure methodological coherence and facilitate interpretation.

For the first question, coding was organized around two main axes: whether the feedback was perceived as useful or not, and which linguistic aspects participants considered it to be most helpful for. Thus, codes captured both the perceived usefulness or lack of usefulness of the feedback received and the linguistic areas in which improvement was perceived, grouped into broad categories related to writing (e.g., grammar, structure and organization, cohesion, vocabulary).

In the second question, codes were created to capture the perceived advantages and disadvantages of the type of feedback received (teacher, AI, or both). Within each category, subcodes were established to reflect the nature of these advantages and disadvantages. The aim here was to extract participants' evaluations and descriptions of the feedback received rather than perceived effects on their writing or teaching skills. Accordingly, codes referred to dimensions such as the format of feedback, its precision and relevance, and the presence of improvement suggestions.

For the third question, coding focused on participants' self-perceived preparedness to teach and assess writing, as well as the areas they felt required improvement. Responses were first categorized according to whether pre-service teachers considered themselves sufficiently or insufficiently prepared. Then, the areas for improvement were classified under broader dimensions related to linguistic knowledge, assessment process, teaching resources and tools, and affective aspects (e.g., self-confidence). Given the linguistic emphasis of this study, the language-related aspects mentioned as areas for improvement were further categorized into the general dimensions of written language already reported in the first question.

Finally, the coding of the fourth question focused on whether participants would include any aspect of the methodology implemented during the study and what teaching approaches they proposed. The resulting codes first distinguished between inclusion and non-inclusion of elements from the intervention, and then grouped the methodological trends and resources considered most relevant by participants, such as teacher feedback, active or motivating methodologies, topics of interest or close to the students' reality, peer feedback or inclusion of AI. For this last category, and given the specific focus of this study, the reasons provided for including AI were also analyzed and categorized.

Finally, each participant's responses were also coded according to their group (CG or EG) to enable cross-group comparison.

3. Results

The results are organized according to the research objectives established at the beginning of the article. The first objective, concerning participants' perceptions of delayed teacher feedback and its combination with immediate feedback generated by generative AI, corresponds to the first two open-ended questions they answered. The second objective, related to pre-service teachers' perceived preparation and their views on evaluation strategies, is addressed through the third and fourth questions. Additionally, all codes identified and mentioned in this section are underlined for greater clarity. The reported codes are illustrated with excerpts from participants' responses, identified as S.XX.

3.1. Objective 1. Perceptions of Delayed Teacher Feedback and Hybrid Feedback

The first question posed after the intervention referred to the perceived usefulness of the feedback received by the pre-service teachers. In the group that only received teacher feedback, 36 participants (87.8%) considered it useful, while 5 of them (12.2%) felt it had not been sufficiently helpful. In contrast, in the group that received hybrid feedback, 46 pre-service teachers (95.8%) found the feedback experienced during the intervention sessions to be useful, while only 2 (4.2%) stated otherwise.

When analyzing the specific aspects for which the feedback had been helpful, students in the CG emphasized that the feedback provided by the instructors mainly supported them in terms of appropriateness in their writing ("Even though we only completed a few writing tasks, I must say I got used to writing [...] and to identifying which expressions can or cannot be used in an essay" S.57), text structure and organization ("In my case, I started writing the first essay without any organized structure, just including ideas as they came to mind, without an introduction or a conclusion. When I was able to review the feedback, I realized those mistakes and was able to correct them and gradually improve" S.50) and grammar ("It helped me reduce grammatical errors, which is essential for my future role as a teacher" S.63).

In the case of the GE, participants highlighted the usefulness of educators' feedback for structure and organization and grammar ("I believe the feedback provided by the teacher was key to improving my argumentative essay writing. The feedback helped me identify both structural and grammatical errors" S.8). Regarding the AI-generated feedback, students in the EG particularly emphasized its usefulness for textual structure and organization ("I think it did help me improve my writing because thanks to it, I learned how to formulate a thesis and understand the parts of an essay, something essential to know where to begin" S.3), vocabulary and cohesion ("The AI provides immediate feedback on the text you submit, and that helped me identify issues I had, such as repetition or connectors, and it offered me alternatives to make the text more varied" S.26). Furthermore, 34 of the 48 students in the EG mentioned the benefits of the combination of both types of feedback, often assigning different roles to each ("Both types of feedback, from the teacher and from the AI, are useful, but in different ways" S.29). Nevertheless, 12 students explicitly expressed a preference for teacher feedback ("I still prefer the teacher's feedback because I feel more confident when it comes to corrections" S.21).

The second question asked students to identify the advantages and disadvantages of the type of feedback they had received. In the CG, participants explicitly identified advantages of teacher feedback in 33 cases, while they pointed out disadvantages in 35. In contrast, in the EG, 28 students reported advantages in 28 instances and reported disadvantages of the same type of feedback in 21. Regarding the AI-generated feedback received during the writing process, students identified advantages on 36 occasions, while they referred to

disadvantages on 27. These frequencies reflect individual mentions, not participants, as some students pointed out both advantages and disadvantages in their responses.

In the CG, the most frequently mentioned advantages of human corrections alluded to the clarity and organization in relation to the correction format (“First of all, the comments section highlights your main mistakes, and the corrections on the file allow you to see all the errors in detail. Also, the assessment is not general, it evaluates multiple aspects, which lets you see in which areas you struggle the most” S.55), the clear identification of errors and strengths in students’ texts (“It helps bring visibility to the errors, or, on the contrary, to the parts that are done well” S.66), and the precision and relevance of the feedback (“Since the feedback was so thorough, it offered a global perspective on what we had written, because it addressed many aspects like structure, grammar, spelling, formality, use of impersonal voice, or supporting arguments, something that helped us apply what was indicated” S.86). Among the perceived disadvantages, students in the CG also referred to the format of the feedback, particularly, they criticized the way it was delivered or its structure in some cases (“Some of the feedback could have been given in a face-to-face meeting, in order to understand the corrections better and ask questions” S.51). Other commonly cited drawbacks included the demotivation caused by negative assessment (“It is very frustrating to get a zero in a section that you always thought you were good at” S.67), and the lack of improvement suggestions (“One disadvantage I see in this model is the lack of guidelines for improvement” S.69).

In the case of the group that received hybrid feedback, the most frequently mentioned positive aspects echoed those reported by the CG when evaluating teacher corrections. However, the main drawbacks highlighted were the delayed nature of teacher feedback (“The disadvantages of this kind of feedback are that sometimes it is slow, and it takes too long to receive it, which means you have already forgotten about the text and it is harder to understand the feedback. Also, when you get it, you are in a different setting and under different conditions, because no one has two days that are exactly alike” S.27), and, again, the format of the feedback. Regarding this last point, students noted issues such as limitations in some subsections of the feedback or its structure (“One of the disadvantages I see in teacher feedback is the use of numerical scores in rubrics. It is true that is an effective way to assess and it has some advantages, but when it comes to improving or getting more involved in the next essay, I do not think it helps much. Many times, we just look at the grades we received and do not pay much attention to which areas we need to improve or what our strengths are in writing” S.46), and the way the feedback was delivered (“The only disadvantage is that it was given virtually, and although it is hard to go one by one in person commenting on the mistakes, I prefer face-to-face interaction because it is more personal and lets you ask questions on the spot if needed” S.28).

On the other hand, when expressing their views on the AI-generated feedback, the most frequently mentioned positive aspects were its immediacy and availability during the writing process (“It allows you to improve your essay as you are writing it, which I think makes you much more aware of your mistakes and gives you a wider range of alternatives that could be useful in future essays” S.22), as well as the presence of improvement suggestions and examples (“Copilot offered me multiple examples of vocabulary and connectors that helped improve the essay and try not to be so repetitive” S.31). As for the negative aspects, the most commonly reported issue was the lack of precision in the feedback at times, either because it did not match the information that students expected (“I noticed some drawbacks, such as the lack of alignment between the responses and the prompts you give it. Although I have to admit that this issue could stem from a lack of precision when formulating the questions or from trying to address too many things at once” S.22), or due to corrections that were perceived as inappropriate or

incomplete (“It had limitations and would get stuck in a loop, repeating the same text and citing it as an error even though the text had already been corrected before” S.39). Some participants also mentioned technical issues with the tool (“For example, in some sessions the AI had problems, and we lost time in class” S.45).

3.2. Objective 2. Perceptions on Preparation and Strategies for Writing Assessment

In the third question, pre-service teachers were asked about their perceived preparedness to teach and assess writing, as well as the aspects they felt they needed to improve. In the CG, 22 students (53.7%) reported feeling inadequately or only partially prepared, 13 students (31.7%) felt sufficiently prepared, and 6 students (14.6%) did not clearly state a position. In the EG, 18 students (37.5%) reported insufficient preparation, 23 (47.9%) believed they were sufficiently prepared, and 7 (14.6%) did not take a clear stance.

The areas for improvement identified were the same across both groups: linguistic knowledge (“This activity made me realize I still have a lot to improve in this area. Previously, I thought I was quite skilled when it came to writing and expressing myself” S.33), knowledge of the assessment process for writing and language (“I believe I especially need to practice how to assess texts so I can correct them with a solid foundation and clear purpose” S.84), teaching resources and tools (“I think I especially need to work on how to implement it in Primary classrooms in a sustained way and ensure students develop proper use of punctuation, accents, and spelling; also on how to teach them to express themselves more formally in exams, for example, how to teach them to speak or to describe an object or a person in detail [...]” S.47), and affective aspects (“I need to improve my confidence so I can pass that on to my future students” S.77). The Sankey diagram shows the frequency of each code by group (Figure 1).

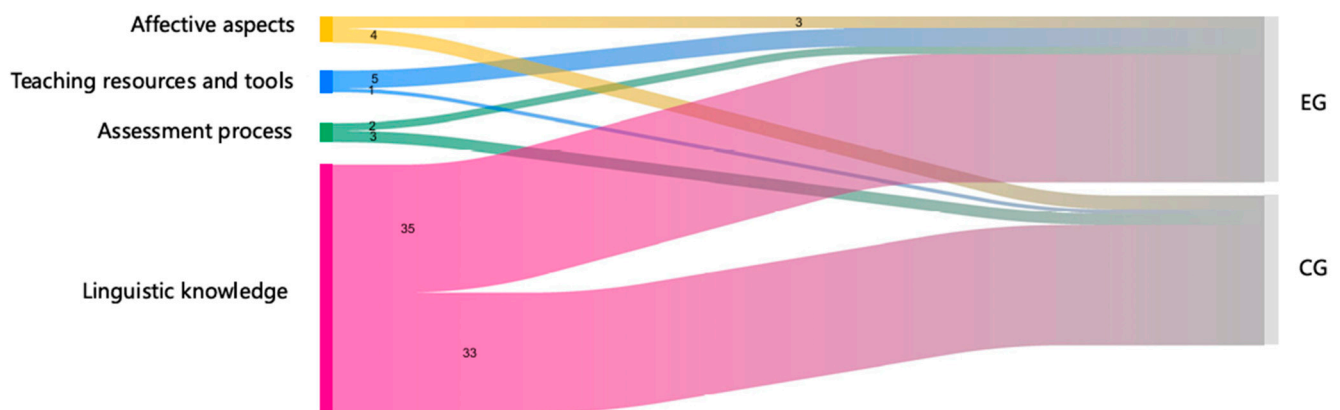


Figure 1. Frequency of each code by group.

In both groups, the most frequently mentioned area for improvement was linguistic knowledge. Even so, within that category, the specific aspects students aimed to improve differed by group. In the CG, pre-service teachers most frequently mentioned the need to improve spelling (“I need to improve my spelling and punctuation for future teaching” S.54), and vocabulary (“I should expand my vocabulary to have more resources” S.64). In the GE, the most frequently cited aspects were vocabulary (“Avoid repeating the same word, in other words, expand my vocabulary so I do not repeat” S.3) and the structure and organization of their texts (“I believe I need to strengthen how I organize the information I want to express in my texts” S.41).

Finally, through the fourth question, pre-service teachers were asked to explain how they would approach the teaching and assessment of writing skills, and whether they would incorporate any aspect of the methodology implemented during the sessions. In

the CG, 25 students (61%) stated they would include some elements of the methodology, 6 students (14.6%) expressed opposition to its inclusion, and 10 students (24.4%) did not take a clear position. In the EG, 41 students (85.4%) indicated they would incorporate some aspects of the methodology used in their group, and 7 students (15.6%) did not take a position, while none of the participants expressed opposition.

In the group that received only teacher feedback, the most frequently repeated proposals were the inclusion of teacher assessment (“I would give students feedback so that they become aware of both their weaknesses and their strengths when writing” S.70), and the implementation of active or motivating methodologies (“In my classes, I would give importance to writing through creative activities and dynamics that encourage students to enjoy writing, so that it does not feel like an imposed or uncreative task” S.53). In the group that received hybrid feedback, pre-service teachers primarily advocated for the inclusion of AI. In this regard, in addition to highlighting the potential of this technology to support writing instruction, they also referred to its usefulness for supporting teachers’ work (“It would even help us as teachers to produce feedback and to have more time for other activities” S.45). They also referred to its motivational potential and linked it to the importance of digital competence (“The methodology used during this practice is perfectly applicable to the classroom with some adaptations. I believe it can motivate students to use artificial intelligence to improve a text they have already written themselves, thereby enhancing their writing skills. Moreover, in today’s world, digital competence is essential” S.1). Besides that, participants also suggested working with topics of interest or close to the students’ reality (“I would ask them to write a text on a topic that is familiar or appealing to them” S.4), as well as including teacher assessment (“Moreover, every writing-related activity would always be accompanied by feedback from the teacher so that students can become aware of their mistakes and know how to correct them” S.6).

4. Discussion

Given the complexity involved in the teaching and learning of writing skills, feedback becomes an essential component of the process, even in higher education (Gan et al., 2021; Zhan & Yan, 2025). Even so, implementing it effectively remains a challenge (Leenknecht & Carless, 2023). In this context, generative AI emerges as a tool worth exploring due to its ability to provide instant support (Zhang et al., 2025). The present study aimed to analyze the perceptions of a group of pre-service primary school teachers regarding teacher feedback and hybrid feedback, as well as their perceived preparedness for writing assessment and future teaching practices involving generative AI.

Regarding the first research objective, most participants in both groups considered the type of feedback they received to be useful. This aligns with the value attributed to feedback in the existing literature, which highlights its relevance to the student learning process (Graham et al., 2015; Nunes et al., 2022). However, it is worth noting that the positive perception of its usefulness was higher and nearly unanimous in the EG. Furthermore, a large portion of participants in this group pointed out specific benefits of hybrid feedback, which supports proposals such as that of Escalante et al. (2023), who suggest that immediate feedback mediated by emerging technologies like generative AI can be particularly useful and engaging if introduced as a complement to the teacher’s role, rather than as a substitute for it. When examining the aspects in which pre-service teachers believe the feedback helped them, this idea appears to be reinforced. Both the CG and EG reported that teacher feedback supported their improvement in areas such as appropriateness, textual structure and organization, and grammar. Also, the EG participants stated that Copilot helped them improve especially in textual structure and organization, vocabulary, and cohesion. Thus, each type of feedback

appears to serve a particular function from the perspective of the participants, reinforcing this idea of complementary types of feedback (Escalante et al., 2023).

Although the EG generally perceived generative AI as a valuable support for their writing, some pre-service teachers explicitly expressed a clear preference for teacher feedback. Similarly, participants in the studies by Nazaretsky et al. (2024) and Henderson et al. (2025) also showed an inclination toward teacher-provided feedback, which they considered to be more credible. This finding is also consistent with earlier studies conducted prior to the rise in generative AI. For instance, Lipnevich and Smith (2008) found that students tended to value feedback from teachers more highly than that generated by a computer, even if it is the same.

When identifying the advantages and disadvantages of the feedback received, the CG appears to report more negative aspects than positive ones, whereas the opposite happens in the EG, especially regarding the feedback provided by the AI. This difference may be due to the way students engage with feedback, how they process it, and how they feel about it, depending on whether the source is a human or AI (Bozkurt & Sharma, 2024). This may also help explain the presence of demotivation related to feedback reported in the CG, which was not observed in the EG. Some studies have found reduced anxiety levels when interacting with chatbots (Tai & Chen, 2024), and report that students believe feedback delivered by generative AI tends to have a more positive tone than that provided by teachers, whereas teacher corrections may be perceived as having a negative tone and causing negative feelings (Henderson et al., 2025). Consequently, if AI feedback was received more positively and helped EG students improve their texts, it is possible that the subsequent corrections provided by teachers were fewer or less severe, thus reducing negative perceptions of the teacher feedback received.

On the other hand, the EG specifically highlighted immediacy and availability as key advantages of AI, in line with its potential to provide students with instant support, as noted by other authors (Li et al., 2024). At the same time, this underscores the previously mentioned complementarity between AI and teacher feedback, as pre-service teachers pointed out benefits that are directly linked to the introduction of this technological tool to provide a different kind of support from that of the teacher, yet aimed at the same goal. However, one of the disadvantages noticed by participants was the lack of precision in the feedback generated by the tool. As several pre-service teachers explained, this may stem from poorly formulated prompts, since the response obtained largely depends on how the prompt is written (Bearman et al., 2024; Tlili et al., 2023). Nonetheless, it is also important to consider that this technology has been reported to hallucinate and generate inaccurate information (Alkaissi & McFarlane, 2023; Kim et al., 2025).

Regarding the second objective of this study, pre-service teachers' perceived preparedness to teach and assess writing varies depending on the group. While more students in the CG reported feeling insufficiently prepared, the opposite can be seen in the EG, where a greater number felt adequately prepared. This may be related to the previously mentioned perception of how hybrid feedback helped them progress. It could also be linked to the effect observed by Zhu et al. (2025) on teacher training. The authors describe a refinement in the way participants in their study provided feedback on written texts after interacting with ChatGPT. In line with these authors' findings, but focusing on self-perceptions rather than actual improvement of skills, the experiences reported by our participants suggest that engaging with AI during the feedback process may have fostered a perceived development of the competences required to teach and assess writing, making them feel more confident in this task, even though our students admitted that they still need to improve their knowledge about it, as reflected in the fourth open-ended question.

Nevertheless, this is not the only area that pre-service teachers believe they need to strengthen. In both the CG and EG, students primarily identified linguistic knowledge as the main area for improvement, reflecting the writing shortages often observed from early educational stages through to higher education (Fernández Merino, 2015). Although students in both groups agreed on the need to improve their language skills, they differed somewhat in the specific linguistic aspects they felt required more work. CG participants mainly emphasized spelling and vocabulary, while those in the EG focused on vocabulary and the structure and organization of their texts. This focus on structural aspects among EG students is not surprising, considering that, in previous responses, they indicated that both teacher and AI-generated feedback has helped them progress in this part of their writing. This may have led them to pay greater attention to structure and regard it as a key area for further improvement.

Finally, when it comes to teaching and assessing writing in their future teacher practice, more students from the group that received hybrid feedback indicated a willingness to replicate some parts of the methodology used with them. These results are consistent with the more positive perception they have of the feedback they received, which may influence their predisposition to adopt it, as students' perception of feedback shapes how they engage with it (Jonsson & Panadero, 2018), and a positive view increases the likelihood that it will be accepted and applied (Panadero et al., 2023). At the same time, teachers' beliefs about resources implemented to support writing, like Copilot in this case, may affect their willingness to use it and their approach to writing instruction (Hall, 2016). This would also explain their openness to integrating generative AI into their future teaching practices, as well as the motivational potential they attribute to it. In this regard, it is worth noting that, in addition to viewing AI as a valuable tool to support the teaching and assessment of writing, pre-service teachers also highlighted its potential to assist educators, which aligns with the idea that AI could help reduce instructors' workload to allow them to focus on more meaningful issues (Chan & Tsi, 2024). Moreover, EG participants also connected their proposal to include AI in their future classrooms with the importance of developing students' digital competence. This is especially relevant considering the prominent role that AI is expected to play in the future of education (Robert et al., 2025) and the growing emphasis on developing AI literacy (e.g., Moorehouse & Kohnke, 2024).

5. Conclusions

This study highlights relevant differences in how pre-service teachers perceive teacher feedback compared to hybrid feedback on their writing. While the feedback received in each group was generally seen as useful by all participants, this perception was notably stronger in the experimental group, who often attributed complementary functions to AI and teacher feedback. These findings support the potential of hybrid feedback models that combine AI's immediacy with teacher's expertise while maintaining personalization.

Participants in the EG reported perceived improvements even in higher-level writing aspects, such as structure and organization of their texts, challenging prior assumptions that tools for automated feedback may support surface-level corrections. Despite this, students still expressed a clear preference for teacher feedback.

The experimental group also showed greater confidence in their preparation to teach and assess writing and were willing to include generative AI tools in their future practice. Overall, they reported a more positive view of both the experience and their own abilities, reinforcing the value of integrating generative AI as a complement to teacher's work.

Limitations and Future Research

This study presents some limitations that raise opportunities for future research. Expanding the sample size would be of interest, as it would allow the inclusion of new variables, such as gender, to analyze in greater depth how these may influence participants' perceptions. Likewise, the qualitative design, although appropriate for exploring pre-service teachers' views in detail, does not allow for generalizable conclusions. Future studies could incorporate mixed-methods designs to examine how hybrid feedback affects writing development and to analyze the evolution of students' perceptions.

Author Contributions: Conceptualization, M.N.-I.; methodology, M.N.-I., J.G.M., A.G.G.; software, M.N.-I.; formal analysis, M.N.-I., P.O.-J.; investigation, M.N.-I., J.G.M.; writing—original draft preparation, M.N.-I.; writing—review and editing, M.N.-I., J.G.M., P.O.-J., A.G.G.; visualization M.N.-I., J.G.M., P.O.-J., A.G.G.; supervision, J.G.M., A.G.G.; project administration, J.G.M., A.G.G.; funding acquisition, A.G.G., M.N.-I., P.O.-J. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Regional Government of Castilla-La Mancha and the European Social Fund Plus (EFS+) Grant 2023-PREJCCM-000054; the University of Castilla-La Mancha and the European Regional Development Fund (ERDF) Grants 2022-GRIN-34470 and 2022-GRIN-34039, 2025-GRIN-38341; and the Regional Government of Castilla-La Mancha and the European Regional Development Fund (ERDF) Grant SBPLY/24/180225/000132.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by the Research in Social Investigation Ethics Committee of the University of Castilla-La Mancha (protocol code CEIS-2025-103060 and date of approval: 10 July 2025) for studies involving humans.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Dataset available on request due to restrictions.

Acknowledgments: We would like to thank the participants for their collaboration in this investigation.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

- Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in ChatGPT: Implications in scientific writing. *Cureus*, 15(2), e35179. [\[CrossRef\]](#)
- Barrot, J. S. (2024). ChatGPT as a Language learning tool: An emerging technology report. *Technology, Knowledge and Learning*, 29, 1151–1156. [\[CrossRef\]](#)
- Bearman, M., Tai, J., Dawson, P., Boud, D., & Ajjawi, R. (2024). Developing evaluative judgement for a time of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, 49(6), 893–905. [\[CrossRef\]](#)
- Bozkurt, A., & Sharma, R. C. (2024). Are we facing an algorithmic renaissance or apocalypse? Generative AI, ChatBots, and emerging human-machine interaction in the educational landscape. *Asian Journal of Distance Education*, 19(1), 1–12.
- Celik, I., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research. *TechTrends*, 66(4), 616–630. [\[CrossRef\]](#)
- Chan, C. K. Y., & Tsi, L. H. (2024). Will generative AI replace teachers in higher education? A study of teacher and student perceptions. *Studies in Educational Evaluation*, 83, 101395. [\[CrossRef\]](#)
- Chen, W. (2023). Investigating novice EFL writing teachers' beliefs and practices concerning written corrective feedback across contexts: A case study from a complexity theory perspective. *Language Awareness*, 32(3), 465–486. [\[CrossRef\]](#)
- Escalante, J., Pack, A., & Barret, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. *International Journal of Educational Technology in Higher Education*, 20, 57. [\[CrossRef\]](#)
- Feng, J., Yu, B., Tan, W. H., Dai, Z., & Li, Z. (2025). Key factors influencing educational technology adoption in higher education: A systematic review. *PLoS Digit Health*, 4(4), e0000764. [\[CrossRef\]](#)

- Fernández Merino, M. (2015). La expresión escrita. In E. Martínez Carro, M. Simón Parra, M. García Sánchez, M. Fernández Merino, & M. Ortiz Canseco (Eds.), *Didáctica de la Lengua Española en Educación Primaria* (pp. 139–166). UNIR Editorial.
- Gan, Z., An, Z., & Liu, F. (2021). Teacher feedback practices, student feedback motivation, and feedback behavior: How are they associated with learning outcomes? *Frontiers in Psychology*, 12, 697045. [\[CrossRef\]](#)
- Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology*, 26(2), 5–24. [\[CrossRef\]](#)
- Graham, S., Hebert, M., & Harris, K. R. (2015). Formative assessment and writing: A meta-analysis. *The Elementary School Journal*, 115(4), 523–547. [\[CrossRef\]](#)
- 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(7), 8435–8463. [\[CrossRef\]](#)
- Hall, A. H. (2016). Examining shifts in preservice teachers' beliefs and attitudes toward writing instruction. *Journal of Early Childhood Teacher Education*, 37(2), 142–156. [\[CrossRef\]](#)
- Henderson, M., Bearman, M., Chung, J., Fawns, T., Buckingham Shum, S., Matthews, K. E., & de Mello Heredia, J. (2025). Comparing Generative AI and teacher feedback: Student perceptions of usefulness and trustworthiness. *Assessment & Evaluation in Higher Education*, 1–16. [\[CrossRef\]](#)
- Jonsson, A., & Panadero, E. (2018). Facilitating students' active engagement with feedback. In A. Lipnevich, & J. Smith (Eds.), *The Cambridge handbook of instructional feedback (Cambridge handbooks in psychology)* (pp. 531–553). Cambridge University Press. [\[CrossRef\]](#)
- Junqueira, L., & Payant, C. (2015). "I just want to do it right, but it's so hard": A novice teacher's written feedback beliefs and practices. *Journal of Second Language Writing*, 27, 19–36. [\[CrossRef\]](#)
- Kim, J., Yu, S., Detrick, R., & Li, N. (2025). Exploring students' perspectives on Generative AI-assisted academic writing. *Education and Information Technologies*, 30, 1265–1300. [\[CrossRef\]](#)
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537–550. [\[CrossRef\]](#)
- Latifi, S., & Noroozi, O. (2021). Supporting argumentative essay writing through an online supported peer-review script. *Innovations in Education and Teaching International*, 58(5), 501–511. [\[CrossRef\]](#)
- Leenknecht, M. J., & Carless, D. (2023). Students' feedback seeking behaviour in undergraduate education: A scoping review. *Educational research review*, 40, 100549. [\[CrossRef\]](#)
- Li, B., Lowell, V. L., Wang, C., & Li, X. (2024). A systematic review of the first year of publications on ChatGPT and language education: Examining research on ChatGPT's use in language learning and teaching. *Computers and Education: Artificial Intelligence*, 7, 100266. [\[CrossRef\]](#)
- Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, 21(2), 100790. [\[CrossRef\]](#)
- Link, S., Mehrzad, M., & Rahimi, M. (2022). Impact of automated writing evaluation on teacher feedback, student revision, and writing improvement. *Computer Assisted Language Learning*, 35(4), 605–634. [\[CrossRef\]](#)
- Lipnevich, A. A., & Smith, J. K. (2008). Response to assessment feedback: The effects of grades, praise, and source of information. *ETS Research Report Series*, 2008(1), i-57. [\[CrossRef\]](#)
- Moorehouse, B. J., & Kohnke, L. (2024). The effects of generative AI on initial language teacher education: The perceptions of teacher educators. *System*, 122, 103290. [\[CrossRef\]](#)
- Navarro-Guzmán, C., & Casero-Martínez, A. (2012). Análisis de las diferencias de género en la elección de estudios universitarios. *Estudios Sobre Educación*, 22, 115–132. [\[CrossRef\]](#)
- Nazaretsky, T., Mejia-Domenzain, P., Swamy, V., Frej, J., & Käser, T. (2024). AI or human? Evaluating student feedback perceptions in higher education. In R. Ferreira Mello, N. Rummel, I. Jivet, G. Pishtari, & J. A. Ruipérez Valiente (Eds.), *European conference on technology enhanced learning* (pp. 284–298). Springer Nature.
- Nunes, A., Cordeiro, C., Limpo, T., & Castro, S. L. (2022). Effectiveness of automated writing evaluation systems in school settings: A systematic review of studies from 2000 to 2020. *Journal of Computer Assisted Learning*, 38(2), 599–620. [\[CrossRef\]](#)
- Panadero, E., Alqassab, M., Fernández Ruiz, J., & Ocampo, J. C. (2023). A systematic review on peer assessment: Intrapersonal and interpersonal factors. *Assessment & Evaluation in Higher Education*, 48(8), 1053–1075. [\[CrossRef\]](#)
- Robert, J., Muscanell, N., McCormack, M., Pelletier, K., Arnold, K., Arbino, N., Young, K., & Reeves, J. (2025). *2025 EDUCAUSE horizon report: Teaching and learning edition*. EDUCAUSE Publications.
- Ropohl, M., & Rönnebeck, S. (2019). Making learning effective—Quantity and quality of pre-service teachers' feedback. *International Journal of Science Education*, 41(15), 2156–2176. [\[CrossRef\]](#)
- Ruiz Flores, M. (2009). *Evaluación de lengua escrita y dependencia de lo literal*. Graó.
- Steiss, J., Tate, T., Graham, S., Cruz, J., Hebert, M., Wang, J., Moon, Y., Tseng, W., Warschauer, M., & Olson, C. B. (2024). Comparing the quality of human and ChatGPT feedback of students' writing. *Learning and Instruction*, 91, 101894. [\[CrossRef\]](#)

- Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, 57, 100752. [\[CrossRef\]](#)
- Tai, T. Y., & Chen, H. H. J. (2024). The impact of intelligent personal assistants on adolescent EFL learners' listening comprehension. *Computer Assisted Language Learning*, 37(3), 433–460. [\[CrossRef\]](#)
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(1), 15. [\[CrossRef\]](#)
- Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: A systematic review and a taxonomy of recipience processes. *Educational Psychologist*, 52(1), 17–37. [\[CrossRef\]](#)
- Xu, L., & Zhang, T. (2022). Engaging with multiple sources of feedback in academic writing: Postgraduate students' perspectives. *Assessment Evaluation in Higher Education*, 48(7), 995–1008. [\[CrossRef\]](#)
- Zhan, Y., & Yan, Z. (2025). Students' engagement with ChatGPT feedback: Implications for student feedback literacy in the context of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, 1–14. [\[CrossRef\]](#)
- Zhang, Z., Aubrey, S., Huang, X., & Chiu, T. K. F. (2025). The role of generative AI and hybrid feedback in improving L2 writing skills: A comparative study. *Innovation in Language Learning and Teaching*, 1–19. [\[CrossRef\]](#)
- Zhu, S., Li, J., Yao, Y., & Zhu, X. (2025). What they provide and how: An intervention study on pre-service teachers' GenAI-assisted writing feedback. *The Internet and Higher Education*, 67, 101040. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.