frontiers | Frontiers in Education

TYPE Brief Research Report
PUBLISHED 12 November 2025
DOI 10.3389/feduc.2025.1706236

OPEN ACCESS

EDITED BY
José Manuel de Amo Sánchez-Fortún,
University of Almeria, Spain

REVIEWED BY
Enrique Sologuren,
University of the Andes, Chile
Constanza Cerda,
Pontificia Universidad Católica de Valparaíso, Chile

*CORRESPONDENCE
Steffanie Kloss
steffanie.kloss@unab.cl

RECEIVED 15 September 2025
ACCEPTED 22 October 2025
PUBLISHED 12 November 2025

CITATION
Mateo-Girona MT, Kloss S and Lillo-Fuentes F (2025) Empowering GPT as a processual writer: Didactext-guided prompting improves knowledge access, iterative revision, and overall textual quality. *Front. Educ.* 10:1706236.
doi: 10.3389/feduc.2025.1706236

Empowering GPT as a processual writer: Didactext-guided prompting improves knowledge access, iterative revision, and overall textual quality

M. Teresa Mateo-Girona[1], Steffanie Kloss[2]* and Fernando Lillo-Fuentes[3]

[1]Department of Language Teaching, Arts and Physical Education, Faculty of Education, Complutense University of Madrid, Madrid, Spain, [2]Faculty of Education and Social Sciences, Andres Bello University, Santiago, Chile, [3]Department of Language, Scientific and Mathematical Education, Faculty of Education, University of Barcelona, Barcelona, Spain

Large language models are increasingly used as writing assistants, but their application often relies on holistic prompting that overlooks the recursive and cognitive dimensions of writing. This article investigates how guided prompting based on the Didactext model empowers GPT-4 to function as a processual writer, enhancing literacy processes in educational contexts. By decomposing writing into four recursive phases—knowledge access, planning, production, and revision—we demonstrate empirical improvements in reasoning depth, iterative refinement, and overall output quality. Building on GPT-4's advanced capabilities in multimodal reasoning and steerability, the study adopts a hybrid experimental design with 150 mini-essay titles generated under guided and unguided conditions. Overall, guided prompts achieved higher textual quality, with raters observing clearer structure, deeper reasoning, and more precise use of evidence. Bias analyses also indicated a reduction in stereotypical content, though not its total elimination. These findings offer novel evidence of how AI can be used to simulate human cognitive writing processes and support literacy development, particularly in the revision phase. Implications include the design of AI-powered tutoring tools capable of encouraging gradual and proactive writing practices while reducing bias in diverse linguistic contexts.

KEYWORDS

GPT, Didactext framework, processual writing, guided prompting, text quality metrics, AI literacy

1 Introduction

The development of generative artificial intelligence (GAI) stems from advances in automatic learning, or machine learning, a subdiscipline focused on building systems capable of learning from data through sequential architectures in order to improve performance in tasks such as automated word classification or numerical prediction (Díaz-Ramírez, 2021). Earlier technologies preceding GAI were unable to generate new content from learned patterns; they could only replicate it. The introduction of the Transformer model by Vaswani et al. (2017) marked a turning point, replacing the previous architecture with one based on attention mechanisms. These make it possible to establish complex relationships between words and to model text sequences by identifying which elements should receive greater

weight during processing (González Rivas, 2025). As a result, transformers can capture intricate semantic and syntactic relationships across words located in different parts of sentences or paragraphs, representing a major advance in the semantic and syntactic understanding of natural language.

This architecture gave rise to the first large-scale generative language models (LLMs), such as GPT (Generative Pre-Trained Transformer), trained on vast corpora including internet forums, books, articles, and others (Kalyan, 2024). These models are identified as tools of generative artificial intelligence because they can automatically produce content in response to written prompts—ranging from images to mathematical operations and even texts which are coherent, structured and adapted across different tasks involving natural language processing (NLP) (UNESCO, 2020).

Over the past decade, AI-based language models have expanded rapidly, becoming embedded in multiple spheres of society, particularly in writing instruction (Teng, 2024). These have transformed how writing and information access are approached. Today, their use is widespread among students, educators, and professionals, who rely on them to compose emails, reports, essays, and other discursive genres.

With the release of increasingly advanced versions—GPT-3, GPT-4, GPT-4o, GPT-4.5, and GPT-5—the models have reached a level of sophistication that enables not only text generation but also the comprehension of complex instructions, information synthesis, argument formulation, and evaluative tasks such as providing feedback on academic texts (Jain et al., 2025). These developments have had a significant impact on education, particularly in academic writing, language teaching, and automated assessment, generating strong interest within the academic community due to their potential as support tools to related professionals.

The effective use of language models such as ChatGPT depends not only on the writer's command of the written discourse but also on the user's ability to interact strategically with the AI. The technique of prompt engineering has thus become an essential skill for obtaining relevant and useful responses (Bašić et al., 2023). In second language (L2) writing instruction, Teng (2025) emphasizes that achieving high-quality assisted writing requires that the writer holds a high degree of metacognitive awareness—namely, the ability to plan, monitor, and evaluate AI use in line with their own communicative goals and text type.

In Latin America, research on academic writing and assisted writing—with and without the use of computer tools—has significantly contributed to our understanding of how writers develop self-regulation strategies (Kloss et al., 2025), metacognitive strategies (Valencia-Serrano and Caicedo-Tamayo, 2015), and epistemic awareness (Navarro et al., 2020). The methodological principles that underpin this research have proven fundamental to the study of composition and feedback processes. However, the current scenario is emerging as a new field of development linked to writing and the use of artificial intelligence, which poses not only ethical but also pedagogical challenges. Along these lines, the studies by Venegas (2021) stand out for their pioneering approach in integrating technological tools and linguistic foundations to enhance written production in engineering. These contributions converge with current models of AI-assisted writing, suggesting that generative technologies can be relevantly incorporated into literacy approaches aimed at strengthening writing instruction.

Several scholars warn, however, that the pedagogical application of these models must rest on a foundation of critical, ethical, and argumentative competence (Petingola et al., 2025). Writers must be able to assess the quality of AI-generated responses, identify potential biases or errors, and reformulate texts according to academic standards and values such as intellectual responsibility (Baldrich and Domínguez-Oller, 2024; UNESCO, 2024; Cordovez-Fernández, 2024).

The integration of AI into literacy education constitutes a transformative shift, particularly in writing processes where cognitive demands often challenge learners (Sagredo-Ortiz and Kloss, 2025). Large language models such as GPT-5, released by OpenAI on August 7, 2025, incorporate multimodal reasoning, extended context windows, and adjustable parameters, enabling the simulation of process writing—iterative and phased composition—beyond the mere generation of finished products. The selection of the Didactext model (2003, 2015) is mainly due to its didactic transposition, as it operationalizes the shift from the traditional prescriptive, product-based paradigm to a process-oriented one (Marinkovich, 2002; Zambrano-Valencia et al., 2020), through a didactic sequence that makes explicit the textual configuration influenced by psychocognitive, sociocultural, and rhetorical-pragmalinguistic factors (García Parejo, 2011).

In this context, our study investigates how guided prompting based on the Didactext model empowers GPT-4 as a process-oriented writer, enhancing literacy in educational settings. Process writing emphasizes iteration and reflection, aligning with Vygotsky's zone of proximal development, where AI scaffolds learning beyond individual capabilities. Didactext (2015), grounded in Hayes (1996) cognitive framework and sociocognitive perspectives, organizes writing into four recursive phases: knowledge access, planning, production, and revision (Didactext, 2015). The distinctiveness of this study lies in comparing two approaches: holistic prompting (product-oriented) and Didactext-guided prompting (process-oriented), assessing how GPT-4 simulates human writing processes.

2 Methods

2.1 Experimental design

The experiment was designed with a comparative approach to analyze differences in the quality of essays produced by GPT-4 under two prompting modalities: holistic and guided. The former consists of a single global instruction according to the general purpose ("We are comparing your performance under two prompt styles—holistic vs. guided. For this turn, respond only under the holistic condition and produce your strongest possible essay."), in which we specify the text's goal, the context, the audience, and the anti-invention guarantee, to which we add restrictions on formal aspects: fixed format, sections, metadata, prohibitions of lists, and strict length limits; although this might increase its procedural reasoning mode, it ensures that we obtain texts with the same formal conditions and, therefore, comparable texts. As observed, the holistic prompt focuses on the product, explaining the general characteristics that the produced text should have, that is, some of the main aspects of the rhetorical situation and the discursive genre (Swales, 1990; Benítez, 2000).

The second modality is guided prompting through the Didactext model (2015), which breaks down the process into four phases—knowledge access, planning, production, and revision—through specific prompts (for example, Phase 1: list prior ideas; Phase 2: create a concept map of ideas). The corpus of 750 essays, generated via API

calls based on 150 titles drawn from other corpora of student-written texts, was used as the basis for text production under both conditions. The 750 texts were generated under five prompting conditions: holistic (product-oriented) and guided (according to the Didactext model: Guided Phase 1, 2, 3, and 4; Totally Guided—in all phases). This design made it possible to maintain both formal and substantive comparability among the texts. The script used for the OpenAI GPT-4 API calls is available in the Supplementary materials.

2.2 Evaluation metrics

Subsequently, the evaluation was carried out using a set of computational metrics that allowed us to characterize the quality and diversity of the generated texts.

1. n_tokens: the total number of tokens in each text was counted as a basic measure of length.
2. local_coherence: it was measured by calculating the average cosine similarity between consecutive sentences, providing an indicator of semantic continuity at the sentence level.
3. perplexity: used as a metric of "surprise" with respect to a language model, reflecting the fluency and adequacy of the text relative to probabilistic patterns of the language.
4. TTR (type–token ratio): calculated as the ratio of types to tokens, as a simple index of lexical diversity.
5. MTLD (Measure of Textual Lexical Diversity): applied as a measure of lexical diversity less sensitive to text length, making it more robust than the simple TTR.
6. MA-TTR (moving-average TTR): a moving version of the TTR, calculated with 100-token windows, allowing for a more stable estimation of lexical diversity.
7. BERTScore: implemented to calculate semantic similarity between texts, leveraging deep representations based on pretrained language models.

For statistical verification, paired tests were conducted (Student's *t*-test and, when appropriate, the non-parametric Wilcoxon test). In addition, linear mixed models adjusted by REML (restricted maximum likelihood) were applied, including as fixed effects the treatment and standardized length, and as random effects the identifier of prompt types and a variance component by Topic.

3 Results

The results were organized into four dimensions of analysis: (1) text length, (2) lexical diversity, (3) local coherence—measured as semantic similarity between consecutive sentences—and (4) global coherence, understood as semantic similarity across the different phases of the process. In all of them, guided prompting consistently outperformed holistic prompting. In the knowledge access phase, GPT-4 activated prior knowledge, identified information gaps, and integrated relevant contextual information, which resulted in richer knowledge bases than unguided prompts. During revision, guided prompts reduced factual errors and hallucinations, while coherence improved markedly across drafts. Overall, guided outputs achieved higher textual quality, with evaluators noting clearer structure, deeper reasoning, and more accurate use of evidence. Bias analyses also indicated a reduction in stereotypical content, though not its complete elimination.

3.1 Dimension 1. Longitud del texto

Differences were observed only with respect to the full guided condition (4 phases). Table 1 presents an expanded version of the comparison between prompts: guided prompting (Phase 1, Phase 2, and Phase 3), in addition to the full guided version (including Phase 4) and the holistic one.

It is observed that the only statistically significant difference is between the guided condition (total) and the holistic one. Thus, the length of the text measured in n_tokens (mean difference = 153.27 tokens, $t(149) = 5.26$, $p < 0.001$, d = 0.43), indicates that the guided prompt produces longer texts than the holistic prompt.

3.2 Dimension 2. Lexical diversity

In the comparison between texts generated using the holistic prompt and those obtained with the guided prompt up to phase 4 (final product), the results show a significant difference in lexical diversity. There was a consistent increase in the texts produced with the full guided prompt compared to the holistic ones. The simple TTR (Type–Token Ratio: types/tokens) was higher in the guided condition (M = 0.5693, SD = 0.0279) than in the holistic condition (M = 0.5371, SD = 0.0313), with a statistically significant difference, $t(149) = 11.27$, $p < 0.001$, d = 0.92. Robust measures confirmed this pattern: the MA-TTR (moving-average TTR, window = 100) reached M = 0.8060 in the guided texts compared to M = 0.7747 in the holistic ones ($p \ll 0.001$; d = 1.17). Finally, the MTLD also favored the guided prompt (M = 68.37) over the holistic one (M = 57.94), although it did not reach conventional significance ($t = 1.91$; $p = 0.058$). Taken together, these results clearly and consistently indicate that the full guided prompt increases the lexical diversity of texts, even under metrics less sensitive to text length.

3.3 Dimension 3. Local coherence (semantic similarity between consecutive sentences)

The results indicate that the holistic prompt maintains superior semantic continuity compared to the full guided prompt. The mean local_coherence was M = 0.4581 (SD = 0.0394) for the guided condition and M = 0.5413 (SD = 0.0663) for the holistic one, with a paired difference of −0.0832, statistically significant, $t(149) = -15.50$, $p < 0.001$, Cohen's d = −1.27, indicating a large effect size. Mixed linear models confirmed this pattern: in the local_coherence model, the fixed coefficient for treatment [T.guided] was −0.080 ($p < 0.001$), while the effect of length (n_tokens_s) was small and inconclusive (≈ -0.005; $p = 0.10$). In the model for type_token_ratio, the guided treatment showed a significant positive effect (+0.041; $p < 0.001$), while length also had a significant negative effect (≈ -0.015; $p < 0.001$), indicating that TTR tends to decrease as text length increases. Taken together, these findings suggest that the reduction

TABLE 1 Comparison between prompts.

Métrica	Mean (g Fase1)	SD (g Fase1)	Mean (g Fase2)	SD (g Fase2)	Mean (g Fase3)	SD (g Fase3)	Mean (guided)	SD (guided)	Mean (holistic)	SD (holistic)	Diferencia (g – h)	t(149)	p (paired t)	Cohen's d (pareado)
n_tokens	1550.102	198.412	1605.340	182.215	1650.500	170.120	**1617.887**	166.256	**1464.613**	317.576	**153.273**	**5.26**	$p < 0.001$	0.430
local_ coherence	0.47200	0.04200	0.45000	0.03850	0.44350	0.04400	0.45808	0.03938	0.54127	0.06630	−0.08319	−15.50	$p < 0.001$	−1.267
type_token_ ratio (TTR)	0.56050	0.02900	0.57510	0.02800	0.56520	0.03050	0.56932	0.02790	0.53710	0.03130	0.03223	11.27	$p < 0.001$	0.920
perplexity (GPT-2 trunc.)	28.5000	4.8000	30.0000	4.2500	29.0000	4.1000	29.1430	4.0103	29.4714	4.4390	−0.3284	−0.44	$p = 0.661$	−0.036
MTLD	60.1200	50.3000	70.4500	46.8000	66.0000	60.2000	68.3670	54.615	57.9403	40.015	10.427	1.91	$p = 0.058$	0.156
MA-TTR ($w = 100$)	0.79050	0.02100	0.81020	0.01950	0.80210	0.02000	0.80597	0.01811	0.77470	0.02592	0.03127	14.36	$p \ll 0.001$	1.172

Data that indicate statistical significance or are prominently discussed in the analysis have been highlighted.

in local coherence and the increase in lexical diversity observed in the guided texts are not solely due to their greater length, but rather that guided prompts promote broader lexical and thematic exploration at the cost of sentence-by-sentence continuity.

3.4 Dimension 4. Semantic similarity between the different phases

Figures 1A,B presents the results of the BERTScore F1 (using the bert-base-multilingual-cased model), calculated in both directions (guided → holistic and holistic → guided). The average scores were virtually identical (0.7368 in both directions; guided-to-holistic: 0.7368067455; holistic-to-guided: 0.7368067459), with a standard deviation of approximately 0.02247, indicating a high average semantic similarity between texts generated by both methods. The directional comparison showed no asymmetries (mean difference = 0), and paired statistical tests (t-test and Wilcoxon) did not detect any significant differences ($p = 0.65$), confirming that choosing one text or the other as a reference does not alter the main conclusion: both procedures produce semantically very similar content.

In relation to the phases, the only difference appears in Phase 4, corresponding to the final product of the guided process. Although the average remains high, the observed range (≈0.671–0.784) and the non-zero standard deviation indicate that some pairs of texts (holistic and final guided versions) are more divergent than most of them, suggesting the need for manual inspection to determine whether these differences are due to inclusion/omission of content or a reordering of information. Finally, the greatest semantic divergence occurs when the guided text is developed through all phases, especially during the final editing stage, whereas no significant differences are observed in the other metrics, either between phases or in comparison to the holistic approach.

4 Discussion

The results show that the guided prompt, inspired by the Didactext model (2015), enables GPT-4 to function as a process-oriented writer, externalizing phases of composition that remain implicit in the holistic prompt (Bašić et al., 2023). While the holistic prompt provides only the final product (the complete essay), the guided prompt returns intermediate artifacts—idea lists, concept maps, outlines, and checklists—that reveal the model's cognitive operations at each stage. This externalization not only improves the transparency of the process but also brings AI writing closer to human models of text production, with a particularly significant impact on knowledge access and revision—areas traditionally challenging for students (Sagredo-Ortiz and Kloss, 2025).

In terms of rhetorical freedom, the results suggest a tension: holistic essays exhibit greater structural variability, expressed through the diversity of openings and closings (measured via n-gram entropy) and the flexibility of rhetorical resources. In contrast, guided texts tend to conform to a more homogeneous macrostructure, partly due to the prompts being organized into specific phases and subproducts (González Rivas, 2025). This alignment effect is reflected in measurable indicators such as: the density of headings/rhetorical markers, diversity of openings and closings (Shannon H of initial/final

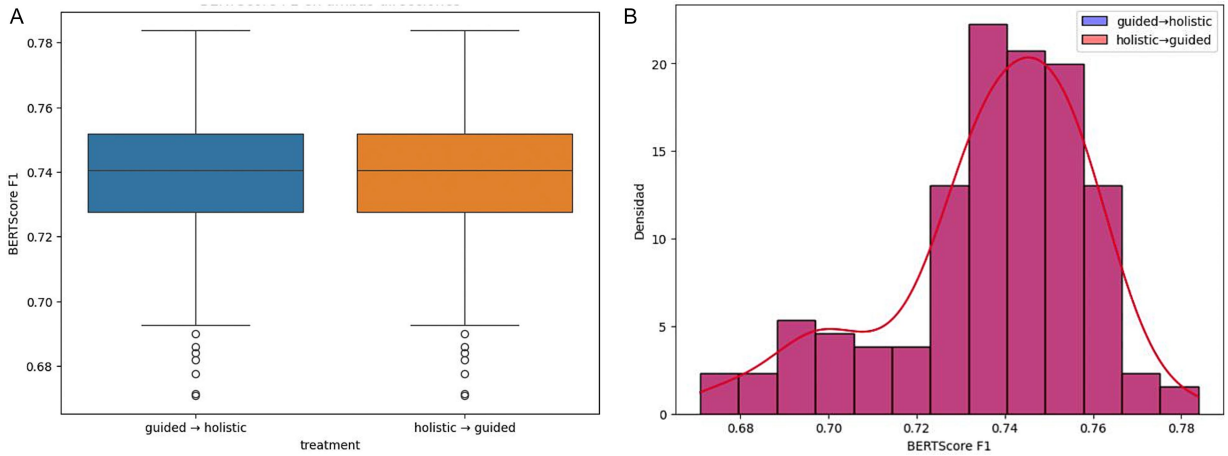

FIGURE 1
(A) Results of the BERTScore F1. **(B)** Results of the BERTScore F1 in both directions.

n-grams), citation density and accuracy (reference/verified ratio), presence of explicit rebuttals, and local coherence (e.g., entity continuity).

Quantitative analyses confirm that the guided prompt produces longer texts with greater lexical diversity, even according to robust metrics that are less sensitive to length. As Teng (2024) warns, language models have rapidly expanded and become established in various fields, particularly in writing instruction, and the results obtained here reflect that potential for expressive expansion. However, this increase in breadth and variety is accompanied by a reduction in local coherence, suggesting that the lexical and thematic exploration induced by guided prompts comes at the cost of sacrificing semantic continuity between sentences (Jain et al., 2025; Zambrano-Valencia et al., 2020).

Our results point to a systematic tension between lexical or thematic expansion and local continuity. Guided prompting produced longer texts with greater lexical diversity (TTR and MA-TTR with $d = 0.92$ and $d = 1.17$, respectively), while coherence between consecutive sentences decreased ($d = -1.27$). We attribute this to the guided prompt's explicit request to generate intermediate artifacts (idea lists, concept maps, outlines, checklists). In this way, the model introduces greater semantic variety, which, without adjustment during the revision phase, can result in thematic jumps between sentences. In our view, these results do not invalidate the pedagogical value of the guided approach, but rather highlight the importance of the revision phase, as it functions as a critical mechanism for reconnecting this diversity at the discursive level and should therefore receive priority attention in future educational applications. Complementarily, semantic similarity between conditions remained generally high (BERTScore), although in Phase 4, specific divergences appeared, linked to the reordering or inclusion of ideas characteristic of final editing, which distinguish guided texts from holistic ones.

From an educational perspective, these findings underscore the potential of guided prompts as a form of scaffolding for academic writing (Baldrich and Domínguez-Oller, 2024; UNESCO, 2024). By structuring the process and making intermediate steps visible, they foster both metacognitive awareness and the development of academic literacy skills. Nevertheless, certain limitations remain: algorithmic biases are not fully eliminated, and the opacity of GPT-4's internal reasoning makes it difficult to interpret the transformations that occur between phases (Teng, 2025; Cordovez-Fernández, 2024). Therefore, implementing this approach in educational contexts requires weighing its value as a support tool against the risks it poses for the reliability and equity of the results produced.

5 Conclusion

The comparison between the holistic and guided prompts confirms that GPT-4, when guided through the Didactext model (2015), offers evidence-based pathways to reimagine the teaching of writing. By moving beyond product-focused writing generation toward a process-oriented approach, AI not only produces longer, more diverse, and more structured texts, but can also significantly contribute to literacy, supporting students in learning how to access knowledge, plan, draft, and revise with greater depth and autonomy.

At the same time, the findings show that these gains are accompanied by a slight reduction in local coherence and the persistence of certain algorithmic biases, underscoring the need for critical supervision and mediation by educators. Overall, the results suggest that AI-assisted process writing is a powerful tool for education—provided it is used thoughtfully and as a complement to human guidance.

The comparison between the holistic guide and the Didactext-guided approach confirms that guiding GPT-4 through an explicit, process-oriented structure leads to improvements across several dimensions of the generated text. The guided approach produced longer texts (average difference ≈ 153 tokens; $d = 0.43$), greater lexical diversity (TTR and MA-TTR; $d = 0.92$ and $d = 1.17$), and higher user-rated text quality, particularly in terms of knowledge access and revision. At the same time, the guided texts showed a notable reduction in local coherence between consecutive sentences ($d = -1.27$), and algorithmic biases were mitigated, though not eliminated.

From a pedagogical perspective, our results support the potential of AI as a tutor for structured writing. By requiring intermediate artifacts

(idea lists, concept maps, outlines, checklists), the Didactext-guided approach can make the writing process more transparent, which could foster metacognitive awareness and help students practice specific skills (retrieving prior knowledge, planning, and revising). For educators and tool designers, a key conclusion is that the greatest educational value arises when guided generation is combined with explicit revision mechanisms that promote textual cohesion and accuracy.

Regarding the limitations of the study, first, the experiment used a single model (GPT-4) and a specific setup (150 mini-essay titles), so future research should consider expanding the experiment to include other LLMs, prompts, languages, and disciplinary areas. Second, although we employed a set of textual metrics (measures of lexical diversity, coherence based on semantic similarity, BERTScore, and perplexity), these indicators could be complemented with other linguistic-discursive features related to textual quality. Therefore, future research would benefit from exploring syntactic complexity, terminological density, and other genre-specific and academic writing characteristics to gain a more comprehensive understanding of the quality of the generated mini-essays. Finally, the study focused on quantitative metrics and did not compare the results—at least explicitly—with human judgment, making it relevant to deepen the analysis by evaluating text quality from the perspective of instructors or teachers.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Ethics statement

The studies involving humans were approved by Acta de Aprobación 033/2025, Universidad Andrés Bello. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

MM-G: Conceptualization, Investigation, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. SK: Conceptualization, Data curation, Funding acquisition, Investigation, Project administration, Validation, Visualization, Writing – original draft, Writing – review & editing. FL-F: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study received economic support from the Agencia Nacional de Investigación y Desarrollo, Chile by Proyecto Fondecyt de Iniciación 11250947.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2025.1706236/full#supplementary-material

References

Baldrich, K., and Domínguez-Oller, J. C. (2024). El uso de ChatGPT en la escritura académica: Un estudio de caso en educación [The use of ChatGPT in academic writing: A case study in Education]. Pixel-Bit. Revista de Medios y Educación, 71, 141–157. doi: 10.12795/pixelbit.103527

Bašić, Ž., Banovac, A., Kružić, I., and Jerković, I. (2023). Better by you, better than me? Chatgpt-3 as writing assistance in students' essays. *Humanit. Soc. Sci. Commun.* 10:750. doi: 10.1057/s41599-023-02269-7

Benítez, R. (2000). La situación retórica: Su importancia en el aprendizaje y en la enseñanza de la producción escrita. *Rev. Signos* 33, 49–67. doi: 10.4067/S0718-09342000004800005

Cordovez-Fernández, M. (2024). Escritura especializada en el ámbito jurídico: Un análisis de las macromovidas de demandas escritas con y sin ChatGPT3.5. *IDS, Revista de Jóvenes Humanistas* 1, 95–126. doi: 10.15581/030.1.003

Díaz-Ramírez, J. (2021). Aprendizaje Automático y Aprendizaje Profundo. *Ingeniare* 29, 180–181. doi: 10.4067/S0718-33052021000200180

Didactext, G. (2003) Modelo sociocognitivo, pragmalingüístico y didáctico para la producción de textos escritos. Didáct. Lengua Lit. 15: 077–104. Available online at: https://revistas.ucm.es/index.php/DIDA/article/view/DIDA0303110077A.

Didactext, G. (2015). Nuevo marco para la producción de textos académicos. *Didáctica Lengua Lit.* 27, 219–254. doi: 10.5209/rev_DIDA.2015.v27.50871

García Parejo, I. (Coord.) (2011). Escribir textos expositivos en el aula. Fundamentación teórica y secuencias didácticas para diferentes niveles. Barcelona: Graó.

González Rivas, E. (2025). Uso de la inteligencia artificial generativa (IAGEN) en el proceso de enseñanza-aprendizaje en la Licenciatura en Administración. *Punto CUNORTE* 1:e20224. doi: 10.32870/punto.v1i20.224

Hayes, J. R. (1996). "A new framework for understanding cognition and affect in writing" in The science of writing: Theories, methods, individual differences, and applications. eds. C. M. Levy and S. Ransdell (Mahwah, NJ: Lawrence Erlbaum Associates), 1–27.

Jain, R., Thanvi, J., and Subasinghe, A. (2025). The evolution of ChatGPT for programming: a comparative study. *Eng. Res. Express* 7:015242. doi: 10.1088/2631-8695/ada51d

Kalyan, K. S. (2024). A survey of GPT-3 family large language models including ChatGPT and GPT-4. *Nat. Lang. Proc. J.* 6:100048. doi: 10.1016/j.nlp.2023.100048

Kloss, S., Tapia-Ladino, M., and Sagredo-Ortiz, s. (2025). Estrategias de autorrevisión en escritura argumentativa: Un estudio con alumnos de pedagogía. *RLA Rev. Lingüíst. Teór. Apl.* 63, 103–129. doi: 10.29393/RLA63-4EASM30004

Marinkovich, J. (2002). Enfoques de proceso en la producción de textos escritos. *Rev. Signos* 35, 217–230. doi: 10.4067/S0718-09342002005100014

Navarro, F., Ávila-Reyes, N., and Cárdenas, M. (2020). Lectura y escritura epistémicas: movilizando aprendizajes disciplinares en textos escolares. *Rev. Electrón. Invest. Educ.* 22:e15. doi: 10.24320/redie.2020.22.e15.2493

Petingola, M., Zhang, Y., Yan, Y., and Lin, W. (2025). Integrating Ethical AI Tools into Educational Practices for Enhancing Academic Integrity. In Proceedings of the 7th ACM Conference on Conversational User Interfaces (pp. 1–6). Available online at: https://doi.org/10.1145/3719160.3737626 (Accessed September 10, 2025).

Sagredo-Ortiz, S., and Kloss, S. (2025). Academic writing strategies in university students from three disciplinary areas: design and validation of an instrument. *Front. Educ.* 10:1600497. doi: 10.3389/feduc.2025.1600497

Swales, J. (1990). Genre analysis. English in academic and research settings. Cambridge: Cambridge University Press.

Teng, M. F. (2024). A systematic review of ChatGPT for English as a foreign language writing: opportunities, challenges, and recommendations. *Int. J. TESOL Stud.* 6, 36–57. doi: 10.58304/ijts.20240304

Teng, M. F. (2025). Metacognitive awareness and EFL learners' perceptions and experiences in utilising ChatGPT for writing feedback. *Eur. J. Educ.* 60:e12811. doi: 10.1111/ejed.12811

UNESCO. (2020). Artificial intelligence in education: Challenges and opportunities for sustainable development. UNESCO. Available online at: https://unesdoc.unesco.org/ark:/48223/pf0000366994 (Accessed September 10, 2025).

Valencia-Serrano, M., and Caicedo-Tamayo, A. (2015). Intervención en estrategias metacognitivas para el mejoramiento de los procesos de composición escrita: Estado de la cuestión. *CES Psicol.* 8, 1–30. Available online at: https://www.redalyc.org/articulo.oa?id=423542417002

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is all you need. In Advances in neural information processing systems. Eds. I. Guyon, U. von Luxburg, S. Bengio, H. M. Wallach, R. Fergus, S. V. N. Vishwanathan et al. (Vol. 30, pp. 5998–6008). Available online at: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

Venegas, R. (2021). Aplicaciones de inteligencia artificial para la clasificación automatizada de propósitos comunicativos en informes de ingeniería. *Rev. Signos* 54, 942–970. doi: 10.4067/S0718-09342021000300942

Zambrano-Valencia, J., Uribe, G., and Camargo, Z. (2020). La competencia escrita en asignaturas del currículo de la lengua castellana. Composición de textos académicos en la formación de maestros. *Rev. Investig. Univ. Quindío* 32, 39–46. doi: 10.33975/riuq.vol32n2.380