AN INTERNATIONAL HYBRID CONFERENCE

12th COTEFL
Conference on Teaching English as A Foreign Language

Paths to Maximize the AI's Roles in Shaping Future Language Education and Literacy Studies

5th CELTI
Conference on English Language Teaching

Training Students to Identify and Correct Fabricated References in ChatGPT-Generated Literature Reviews

Author's:
Hendi Pratama

Affiliation:
Universitas Negeri Semarang, Semarang, Indonesia

Corresponding email:
hendipratama@mail.unnes.ac.id

DOI:
https://doi.org/10.24090/celti.2025.1335

Pages:
25-36

Abstract: The widespread use of generative AI tools such as ChatGPT among university students has introduced new challenges to academic integrity, particularly in the form of hallucinated citations: fabricated references that appear plausible but cannot be verified. This study investigates the effectiveness of a targeted instructional intervention in improving students' ability to identify and correct such false citations. A total of 139 students applied to participate through online recruitment; 50 were selected, 34 attended the training, and 27 completed both pre- and post-tests. The intervention consisted of a three-hour online session conducted via Zoom, where students were trained in responsible AI use, APA 7th citation formatting, and reference verification techniques. Pre- and post-tests required students to write short literature reviews using AI tools, and their submissions were assessed based on four indicators: the number of hallucinated sources, non-journal citations, reputable journal references, and APA formatting errors. Results revealed a significant reduction in fabricated and low-quality references, along with an increase in the use of reputable sources and improved citation formatting. These findings align with previous research emphasizing the risks of uncritical AI use and the importance of digital literacy education. The study concludes that targeted training is an effective method for mitigating citation hallucinations and enhancing students' academic writing skills. The research highlights the necessity of integrating AI literacy into academic curricula to ensure ethical and effective engagement with emerging technologies in higher education.

Keywords: *chatGPT; generatif AI; targeted training; AI Hallucinations*

Introduction

The integration of generative artificial intelligence (AI) tools such as ChatGPT into academic writing has seen rapid growth among university students. Although students report improvements in grammar, structure, and fluency, concerns have arisen about the ethical implications and growing dependency on AI tools for completing academic work (Aljuaid, 2024; Launonen et al., 2024). This surge in reliance has brought academic integrity into question, as students may be bypassing the development of foundational writing and critical thinking skills (Medina et al., 2024; Janković & Kulić, 2025).

One of the most pressing concerns is the phenomenon of AI hallucinations, i.e. ChatGPT's tendency to fabricate plausible-sounding but non-existent references. These hallucinated citations can mislead both students and educators, leading to a loss of academic credibility and weakening students' research

literacy (Buchanan et al., 2023; Lingard, 2023). Students often lack the skills to detect such false citations, particularly in the early stages of their academic training, making this a pedagogical issue as much as a technological one.

To address this, some educational initiatives have introduced training programs focused on equipping students to critically evaluate AI-generated content. Research suggests that interventions like paraphrasing exercises and structured guidance can significantly enhance students' ability to identify misinformation and improve writing quality (Emran et al., 2024; Khalifa & Albadawy, 2024). These findings highlight the need for proactive curriculum design that fosters both digital literacy and academic integrity.

The present study investigates the effectiveness of a brief, targeted training session designed to help students recognize hallucinated references in AI-generated academic texts. By analyzing pre- and post-intervention outcomes, this study evaluates how instructional support impacts students' abilities to verify references and uphold scholarly standards. Previous studies have shown that guided use of ChatGPT, alongside educational scaffolding, can lead to improved writing efficacy and reduced dependency on AI tools (Cheng et al., 2025; Duong et al., 2024). This underscores the importance of integrating AI literacy into higher education to ensure students engage with these technologies critically and ethically.

Literature Review

The emergence of ChatGPT and other large language models (LLMs) has radically reshaped the landscape of academic writing. While these tools offer powerful affordances in generating coherent and well-structured texts, recent scholarship has raised serious concerns about their tendency to produce hallucinated references, citations that appear plausible but are entirely fabricated or factually inaccurate. These hallucinated citations pose significant risks to academic integrity, particularly when used by students with limited experience in academic research and source evaluation. Research consistently shows that citation hallucinations are prevalent and problematic. In one of the earliest systematic evaluations, Walters and Wilder (2023) found that 55% of citations generated by ChatGPT-3.5 were fabricated, and 43% of even the "real" citations contained substantive errors. While the newer GPT-4 model performed better, with hallucination rates dropping to 18%, the persistence of inaccurate referencing remained a significant limitation. Similarly, Buchanan et al. (2023) found that over 30% of citations produced by GPT-3.5 in economics contexts did not exist, and even GPT-4 failed to eliminate this problem entirely.

This issue has implications not only for factual accuracy but also for the development of students' academic skills. Franzoni Velázquez et al. (2024) reported that while students could learn to recognize some fake citations through targeted training, many still failed to evaluate the completeness and relevance of references generated by ChatGPT without explicit instruction. As Day (2023) argues, the ability to identify hallucinated sources is increasingly becoming a core academic skill, as students must now contend not just with evaluating peer-reviewed research but also verifying AI-generated information.

The proliferation of fabricated citations is especially concerning in higher education contexts where critical thinking and evidence-based reasoning are central learning objectives. Scholars like Coetzer and van Aardt (2024) highlight that overreliance on ChatGPT erodes students' capacity to formulate independent arguments, leading to the dilution of the "student voice" and potentially

undermining academic decolonization efforts, particularly in several regional institutions. This sentiment is echoed by Janković and Kulić (2025), who stress that while ChatGPT can enhance writing fluency, students must be explicitly trained to critically evaluate and contextualize AI-generated content to preserve academic integrity.

Another major concern is the inconsistent formatting of citations, even when they refer to real sources. Rashidov (2024) proposed an AI-enhanced system to automatically format references in APA, MLA, or IEEE styles using ChatGPT, but acknowledged that current models still require significant human oversight to ensure accuracy and consistency. Wu and Dang (2023) further demonstrated that only 10% of references generated by ChatGPT in the context of medical writing were fully correct, while DOI hallucinations occurred in the majority of cases.

Nevertheless, some researchers advocate for the responsible integration of AI tools into academic instruction. Duong et al. (2024) reported that ChatGPT enhanced students' grammar, coherence, and vocabulary when used with caution and under guided supervision. Similarly, Xu and Jumaat (2024) found that ChatGPT could effectively scaffold academic writing strategies among EFL students, particularly in planning and revision, as long as ethical and verification practices were taught alongside AI usage.

Ethical frameworks have also been proposed to address this issue. Jarrah et al. (2023) emphasize that ChatGPT can be ethically integrated into academic writing if it is acknowledged appropriately, and if its output is verified rigorously by human authors to prevent inadvertent plagiarism. Scholars such as AlZaabi et al. (2023) echo these calls, recommending the development of educational guidelines and digital literacy frameworks to help students navigate the benefits and limitations of AI tools in research and writing. While ChatGPT holds potential as a supplementary writing aid, its hallucinated citations and formatting errors can seriously undermine academic credibility. These risks are amplified when students are not trained to detect or question AI-generated references. To address these concerns, scholars advocate for instructional interventions, ethical use frameworks, and the reinforcement of human criticality as essential to sustaining integrity in the age of AI-assisted scholarship.

Methodology

This study employed a quasi-experimental pre-test/post-test design to evaluate the effectiveness of a short instructional intervention on students' ability to detect hallucinated and low-quality citations in AI-generated academic writing. This design is frequently used in educational contexts where random assignment is not feasible but where controlled comparisons are desired (Dörnyei, 2007; Mackey & Gass, 2016).

1. Participant Recruitment and Selection

Participants were recruited using a purposive sampling method via researcher's Instagram Stories, targeting university students interested in academic writing and artificial intelligence. A total of 139 individuals expressed interest by responding to the call to participate. After screening for availability and access to necessary technology, 50 students were selected. Of these, 34 joined the scheduled training session, and only 27 students completed both the pre-test and post-test tasks. While this attrition is common in online learning research, final participation met the threshold for small-scale experimental studies in applied linguistics (Nunan, 1992).

2. Training Intervention

Participants attended a synchronous, three-hour online training session delivered via Zoom. This format aligns with best practices in distance learning and allows for both instruction and real-time interaction (Hampel & Stickler, 2012). All participants were required to use laptops and keep their cameras turned on throughout the session to ensure engagement and support monitoring. The training content created and compiled by the researcher and focused on academic writing principles, APA 7th citation formatting, reference verification, and the use of AI tools such as ChatGPT and Grok for literature generation and source validation.

Before the training, participants completed a writing task submitted in Microsoft Word format. They were asked to compose a four-paragraph literature review essay on a given topic, incorporating at least eight journal references using APA 7th edition format. The post-test, administered immediately following the training, required participants to perform a similar writing task on a different topic, also under timed conditions and with AI tool use permitted.

3. Data Collection and Analysis

The pre- and post-test submissions were analyzed based on four indicators:

a. Number of invalid sources i.e. references that were fabricated or could not be verified.

b. Number of non-journal sources including general websites, blogs, and news articles.

c. Number of reputable journal references i.e. articles indexed in Scopus, DOAJ, or SINTA.

d. APA 7th edition citation errors based on adherence to formatting guidelines for both in-text citations and reference lists.

Each submission was coded independently by two raters trained in academic writing assessment, following analytic rubric principles recommended in language assessment (Weigle, 2002). Discrepancies were resolved through discussion to ensure inter-rater reliability. Quantitative comparison using Wilcoxon of pre- and post-test results was administered to measure the effectiveness of the intervention. By focusing on authentic tasks, AI-supported workflows, and verifiable citation practices, this methodology aligns with recent calls in applied linguistics to adapt instruction to new digital literacies (Godwin-Jones, 2018) while preserving academic rigor.

Findings and Discussion

1. Errors Found in Students Academic Writing

Participants' reference lists revealed a striking number of invalid or "hallucinated" sources: citations that simply do not exist in any academic database. For instance, one student cited "Brown, T., & Lee, H. (2025). The psychological benefits of multilingualism in professional settings. Journal of Applied Psychology, 110(3), 321–335. https://doi.org/10.1037/ apl0000987," yet no such article appears in the Journal of Applied Psychology in 2025 . In another instance, a reference to "Dibrova, V., Dukhanina, N., & Hrabar, O. (2025). Advancing foreign language proficiency: The integration of multimedia tools in higher education. Educational Technology Research and Development, 73(2), 694–704. https://doi.org/10.33806/ijaes.v25i2.524" turned out to repurpose an Arabic-English studies DOI and a non-existent volume. These phantom citations undermine credibility and suggest that students must be trained to cross-check every source against reliable indexes before inclusion.

Equally common were citations of non-journal materials such as news articles, blogs, or preprints, presented as if they were peer-reviewed studies. One student listed “Efendioğku, İ. H. (2022). The impact of conspicuous consumption in social media on purchasing intentions. arXiv Preprint. https://arxiv.org/abs/2202.12345,” failing to recognize that arXiv papers are preprints without formal peer review unless subsequently published in a journal . Another simply referenced “BMC Psychology. (2023). Pros & cons: impacts of social media on mental health. BMC Psychology. jurnal.iicet.orgbmcpsychology.biomedcentral.com” with no authors, volume, or pagination, more akin to a news blurb than an article. Such misclassification obscures the difference between scholarly research and popular commentary, and students must learn to distinguish reputable journals from general web content.

Mistakes also arose when students cited journals that, while real, are not indexed in recognized databases such as Scopus, DOAJ, or SINTA, or when they misnamed legitimate outlets. For instance, “Ruslita, G., & Seran, A. (2024). Media dan konsumerisme: Studi kritis pahlawan konsumtif dalam budaya populer. Journal of Mandalika Literature, 6(1), 480–492. https://doi.org/10.36312/jml.v6i1.3976” could not be located in any major index and likely represents a small or predatory publication . Similarly, the citation “Ansar, Ansyar. (2023). Pengaruh Intensitas Iklan Media Sosial Terhadap Perilaku Konsumtif Pada Mahasiswa. Nathiqiyyah, 6(2), 8–14. https://doi.org/10.46781/nathiqiyyah.v6i2.767” refers to a journal absent from Scopus, DOAJ, or SINTA . While local or niche journals may sometimes be acceptable, students should be guided to prefer and justify use of higher-tier, widely indexed publications when possible.

Even when students correctly identified valid, peer-reviewed journals, their APA 7th edition formatting was often riddled with errors. Journal titles were frequently left in plain text rather than italicized, for example, “Journal of Education and Learning, 19(1), 45–56”, and major words in article titles were sometimes not properly capitalized, as in “Advancing foreign language proficiency: The integration of multimedia tools…” . Author names were occasionally inverted or missing initials (“Ansar, Ansyar.” instead of “Ansyar, N. A.”) , and DOI formats vacillated between bare prefixes (“DOI:10.69760/egjlle.250003”) and full URLs (“https://doi.org/10.69760/egjlle.250003”) . Some students omitted volume, issue, or page numbers altogether, as seen in the BMC Psychology citation. Addressing these errors requires focused instruction on APA’s precise requirements for both in-text citations and reference lists.

By examining data taken from participants’ writings including fabricated entries that vanish on verification, misclassified preprints and blogs, obscure or misnamed journals, and pervasive APA formatting mistakes, we gain a clear roadmap for improving students’ research skills. The following training emphasizes rigorous source validation, checking DOIs and journal indexes, careful selection of reputable outlets, and meticulous adherence to APA 7th rules. Only then can their essays reliably reflect genuine scholarship and stand up to academic scrutiny.

2. The Training

The training session was designed to equip students with the skills to produce high-quality, AI-assisted academic writing while avoiding common pitfalls such as plagiarism and hallucinated citations. The session began with a pre-test that tasked students with composing a four-paragraph literature review essay, integrating a minimum of eight journal articles and formatted according to APA 7th edition guidelines. This activity served to benchmark students’ baseline understanding of source credibility and

referencing mechanics. Throughout the session, the facilitator emphasized the importance of writing with academic weight, originality, and responsibility, particularly when employing AI tools like ChatGPT.

Following the initial task, the training transitioned into a structured instructional module covering the principles of scholarly writing and source evaluation. Students were introduced to best practices for crafting effective academic titles, identifying core concepts, and using action-oriented language. A significant focus was placed on teaching APA 7th edition citation rules, including paraphrased and direct quotation techniques, multi-author citation conventions, and reference list formatting. The participants also explored AI-powered research assistants, such as Scispace, Elicit, and Consensus, for identifying relevant sources, and were advised on the manual verification of references using tools like Grok to assess citation validity and journal reputability. The session reinforced that while AI can assist in generating references, final responsibility lies with the writer to ensure accuracy and ethical compliance.

In the final segment, students engaged in a post-test mirroring the pre-test format but on a new topic. This allowed participants to immediately apply the skills learned during training. The post-test required careful integration of AI tools for sourcing and formatting while avoiding non-journal or unverifiable references. Participants were assessed on their ability to reduce the number of hallucinated citations, minimize APA formatting errors, and increase their use of reputable journal sources e.g., those indexed in Scopus, DOAJ, or SINTA. The session concluded with a discussion on the ethical use of AI in academic contexts, including institutional guidelines that caution against overreliance on AI detection software due to high false-positive rates. Through this structured approach, the training effectively bridged AI utility with academic integrity.

3. Quantitative Findings

The Wilcoxon signed-rank test was employed in this study due to the small sample size of 27 participants who completed both the pre-test and post-test tasks. With a limited number of observations, the assumptions of normality required for a traditional paired t-test may not hold, making the non-parametric Wilcoxon test more appropriate. This test effectively compares the paired differences without assuming a normal distribution, providing a robust analysis of the training's impact on students' citation practices despite the constrained sample.

Table 1. Wilcoxon Sign-rank Test of Students' Writings

Paired Samples T-Test

Measure 1		Measure 2	W	z	df	p
Invalid_Pre	-	Invalid_Post	171.000	3.058		0.002
NonJournal_Pre	-	NonJournal_Post	145.000	2.591		0.009
Reputable_Pre	-	Reputable_Post	22.000	-3.657		< .001
APA_Mistakes_Pre	-	APA_Mistakes_Post	331.000	3.949		< .001

Note. Wilcoxon signed-rank test.

Descriptives

	N	Mean	SD	SE	Coefficient of variation
Invalid_Pre	27	2.259	2.212	0.426	0.979
Invalid_Post	27	0.667	1.301	0.250	1.951
NonJournal_Pre	27	1.444	1.783	0.343	1.234
NonJournal_Post	27	0.556	1.121	0.216	2.018
Reputable_Pre	27	4.037	2.710	0.522	0.671

Paired Samples T-Test

Measure 1		Measure 2			W	z	df	p
Reputable_Post	27	6.667	1.519	0.292		0.228		
APA_Mistakes_Pre	27	6.111	4.750	0.914		0.777		
APA_Mistakes_Post	27	2.148	1.748	0.336		0.814		

The paired samples t-test results, using the Wilcoxon signed-rank test, reveal a significant reduction in the number of invalid sources post-training. The pre-test mean for invalid sources was 2.259, which decreased to 0.667 after the intervention, with a Wilcoxon statistic (W) of 171.000, a z-value of 3.058, and a p-value of 0.002. This indicates that the training effectively equipped students to identify and eliminate fabricated or unverifiable references, such as those hallucinated by AI tools like ChatGPT, thereby enhancing the credibility of their academic work.

The analysis also shows a significant decrease in the use of non-journal sources following the training. The pre-test mean of 1.444 dropped to 0.556 post-test, supported by a Wilcoxon statistic (W) of 145.000, a z-value of 2.591, and a p-value of 0.009. This reduction suggests that students became more adept at distinguishing between scholarly sources and less reliable materials like blogs or news articles, reflecting an improved ability to adhere to academic writing standards after the instructional session.

There was a notable increase in the use of reputable journal references post-training, with the pre-test mean of 4.037 rising to 6.667. This improvement is statistically significant, as indicated by a Wilcoxon statistic (W) of 22.000, a z-value of -3.657, and a p-value of <0.001. The increase highlights the training's success in guiding students to prioritize high-quality, indexed journals, such as those in Scopus, DOAJ, or SINTA, thereby strengthening the scholarly foundation of their literature reviews.

The training also led to a significant reduction in APA 7th edition citation errors, with the pre-test mean of 6.111 decreasing to 2.148 post-test, supported by a Wilcoxon statistic (W) of 331.000, a z-value of 3.949, and a p-value of <0.001. This substantial drop indicates that students improved their adherence to APA formatting guidelines, including proper italicization, capitalization, and DOI formatting, after receiving targeted instruction. The lower post-test variability, as shown by a coefficient of variation of 0.814 compared to 0.777 pre-test, further suggests greater consistency in applying these standards.

4. Discussion

The findings of this study demonstrate that targeted training can significantly improve students' ability to identify and correct hallucinated citations in AI-generated academic writing. This aligns with the broader consensus in recent scholarship that instructional intervention is critical for fostering academic integrity in the age of artificial intelligence (Franzoni Velázquez et al., 2024; Day, 2023). The reduction in fabricated and unverifiable sources post-training suggests that students, when adequately guided, can develop the critical literacy skills required to use ChatGPT responsibly. This supports the notion that digital literacy, particularly AI literacy, should be considered a core competency in higher education (Godwin-Jones, 2018).

The statistically significant reduction in invalid citations, from an average of 2.26 to 0.67, provides empirical evidence that focused instruction can meaningfully impact students' referencing practices. These findings parallel those of Walters and Wilder (2023), who observed high levels of citation hallucination in GPT-3.5 outputs, even when users believed the references were accurate. This study reinforces the conclusion that such hallucinations can only be mitigated through user education, not by

relying on AI alone. Furthermore, the training appears to have empowered students to verify sources using external tools and databases, echoing Day's (2023) argument that verification and cross-referencing must become fundamental research practices in the digital age.

A second key result was the decrease in the use of non-journal sources, including blogs and general websites, from 1.44 to 0.56. This finding is consistent with Coetzer and van Aardt (2024), who warn that students' uncritical use of ChatGPT can blur the line between scholarly and informal sources. By explicitly training students to recognize peer-reviewed articles and emphasizing the importance of indexing systems such as Scopus, DOAJ, and SINTA, the intervention helped students make more discerning choices. This supports the work of Rashidov (2024), who highlighted the need for enhanced source classification in AI-generated outputs and advocated for user training as a solution.

The significant improvement in the number of reputable sources, from an average of 4.04 to 6.67, suggests that the training not only reduced error but also elevated the academic quality of students' writing. This aligns with findings by Xu and Jumaat (2024), who showed that when students are guided in using ChatGPT as a writing scaffold rather than a replacement, they are more likely to produce well-researched and structured work. The current study confirms that AI-assisted writing can support deeper engagement with academic content, but only if students are explicitly taught to use the technology critically.

Perhaps most striking was the decrease in APA 7th edition citation errors, from 6.11 to 2.15 per student. This improvement illustrates that referencing mechanics, often treated as peripheral in writing instruction, are a central challenge in the AI era. Unlike human writers who may gradually learn formatting rules, AI tools frequently output inconsistently formatted references (Wu & Dang, 2023). As a result, student overreliance on such outputs without validation leads to widespread technical inaccuracies. The reduction in APA mistakes post-training suggests that focused instruction on citation style, often overlooked in curricula, remains essential and must be integrated into any AI literacy framework.

These findings can also be interpreted through the lens of socio-cognitive writing theory, which views writing as a problem-solving task that requires metacognitive regulation (Flower & Hayes, 1981). From this perspective, the training functioned as a scaffolding tool, helping students monitor, evaluate, and revise their AI-assisted texts. The dramatic reduction in both hallucinated and poorly formatted citations indicates that students became more reflective and analytical about their source use: a key tenet of academic writing development.

In relation to Martin and White's (2005) appraisal theory, which the broader study references in its discourse analysis section, we might understand students' post-training writing as increasingly informed by "judgment", in the form of improved capacity to evaluate the credibility and propriety of AI-generated references. Whereas initial writing samples reflected a passive acceptance of AI suggestions, post-test results demonstrate more critical agency and authorial control.

This study supports the broader theoretical call for integrating digital ethics and AI competencies into the curriculum (Jarrah et al., 2023; AlZaabi et al., 2023). As universities continue to adapt to the presence of generative AI tools in academic life, training programs such as the one administered here may serve as a model for ensuring students use these tools responsibly, ethically, and skillfully. Rather than banning AI or treating it solely as a threat, the findings support a proactive pedagogy that recognizes AI as a writing partner, albeit one that must be handled with discernment and oversight.

Conclusion

This study demonstrates that a short, targeted training session can significantly reduce the number of hallucinated citations and APA formatting errors in AI-assisted academic writing. The intervention effectively helped students distinguish between valid and invalid references, prioritize reputable journal sources, and apply proper citation mechanics. These outcomes underscore the critical importance of integrating AI literacy and academic integrity instruction into higher education curricula. In an era where students increasingly rely on generative tools like ChatGPT, structured training ensures they can use such tools responsibly, enhancing the quality, credibility, and ethics of their academic work.

Despite the promising results, this study is not without limitations. The sample size was relatively small (n=27) and self-selected, which may limit the generalizability of the findings. The short-term nature of the intervention also restricts insights into long-term skill retention. Future research should explore the sustained impact of AI literacy instruction over a semester or academic year and examine how different disciplinary contexts influence students' use of AI-generated references. Expanding the training into blended or modular formats, incorporating peer review, and integrating automated verification tools could further strengthen students' capacity to manage AI-assisted writing with confidence and critical awareness.

Acknowledgments

The author would like to express sincere gratitude to Universitas Negeri Semarang for providing the facilities and access necessary to conduct this research. The institutional support was instrumental in ensuring the smooth implementation of the study and the success of the training program. Appreciation is also extended to all the undergraduate student participants from various universities across Indonesia, whose enthusiasm and commitment made this research possible. Their active engagement and contributions were invaluable in exploring the educational impact of AI-assisted academic writing.

References

Aljuaid, H. (2024). The impact of artificial intelligence tools on academic writing instruction in higher education: A systematic review. Arab World English Journal (AWEJ) Special Issue on ChatGPT. https://ssrn.com/abstract=4814342

AlZaabi, A., ALAmri, A., Albalushi, H., Aljabri, R., & AlAbdulsalam, A. A. (2023). ChatGPT applications in academic research: A review of benefits, concerns, and recommendations. bioRxiv. https://doi.org/10.1101/2023.08.17.553688

Buchanan, J., Hill, S., & Shapoval, O. (2023). ChatGPT hallucinates non-existent citations: Evidence from economics. The American Economist. https://doi.org/10.1177/05694345231218454

Cheng, A., Calhoun, A., & Reedy, G. (2025). Artificial intelligence-assisted academic writing: recommendations for ethical use. Advances in Simulation, 10(1), 22. https://doi.org/10.1186/s41077-025-00350-6

Coetzer, Z., & van Aardt, P. V. (2024). Unsilencing the student voice: Detecting and addressing ChatGPT-generated texts presented as student-authored texts at a university writing centre. International Journal of Critical Diversity Studies, 6(2), 151–174. https://doi.org/10.13169/intecritdivestud.6.2.00151

Day, T. (2023). A preliminary investigation of fake peer-reviewed citations and references generated by ChatGPT. The Professional Geographer, 75(4), 1024–1027. https://doi.org/10.1080/00330124.2023.2190373

Duong, N. H., Tong, T. M. H., & Le, D. (2024). Utilizing ChatGPT in checking academic writing for postgraduate students. Proceedings of the AsiaCALL International Conference. https://doi.org/10.54855/paic.24614

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. College Composition & Communication, 32(4), 365-387. https://doi.org/10.58680/ccc198115885

Franzoni Velázquez, A. L., Huerta, E., & Jensen, S. (2024). Retracting ChatGPT: Completeness and relevance of academic references. Discover Education. https://doi.org/10.1007/s44217-024-00333-1

Godwin-Jones, R. (2018). Using mobile technology to develop language skills and cultural understanding. Language Learning & Technology, 22(3), 104–120. https://www.lltjournal.org/item/2999

Hampel, R., & Stickler, U. (2012). The use of videoconferencing to support multimodal interaction in an online language classroom. ReCALL, 24(2), 116–137. https://doi.org/10.1017/S095834401200002X

Khalifa, M., & Albadawy, M. (2024). Using artificial intelligence in academic writing and research: An essential productivity tool. Computer Methods and Programs in Biomedicine Update, 100145. https://doi.org/10.1016/j.cmpbup.2024.100145

Jarrah, A. M., Wardat, Y., & Fidalgo, P. (2023). Using ChatGPT in academic writing is (not) a form of plagiarism: What does the literature say? Online Journal of Communication and Media Technologies. https://doi.org/10.30935/ojcmt/13572

Janković, A., & Kulić, D. (2025). Use and misuse of ChatGPT in academic writing among the English language students. Information Technologies and Learning Tools, 105(1). https://doi.org/10.33407/itlt.v105i1.5955

Lingard, L. (2023). Writing with ChatGPT: An illustration of its capacity, limitations & implications for academic writers. Perspectives on medical education, 12(1), 261. https://doi.org/10.5334/pme.1072

Mackey, A., & Gass, S. M. (2016). Second language research: Methodology and design (2nd ed.). Routledge.

Martin, J. R., & White, P. R. R. (2005). The language of evaluation: Appraisal in English. Palgrave Macmillan.

Medina, W. A. A., Velásquez, C. H. F., Zegarra, S. D. R. O., Romero, G. P. M., Andía, A. Q., Bellido, I. E. A., ... & Bravo, N. A. (2024). Using ChatGPT in University Academic Writing: A Bibliometric Review Study on the Implications for Writing Reports, Papers, Essays, and Theses. International Journal of Learning, Teaching and Educational Research, 23(12), 360-381. https://doi.org/10.26803/ijlter.23.12.19

Nunan, D. (1992). Research methods in language learning. Cambridge University Press.

Rashidov, A. (2024). Automating citation formatting in scientific publications using ChatGPT. 2024 Asian Conference on Communication and Networks (ASIANComNet), 1–6. https://doi.org/10.1109/asiancomnet63184.2024.10811018

Walters, W. H., & Wilder, E. I. (2023). Fabrication and errors in the bibliographic citations generated by ChatGPT. Scientific Reports, 13, Article 12856. https://doi.org/10.1038/s41598-023-41032-5

Weigle, S. C. (2002). Assessing writing. Cambridge University Press.

Wu, R. T., & Dang, R. (2023). ChatGPT in head and neck scientific writing: A precautionary anecdote. American Journal of Otolaryngology, 44(6), 103980. https://doi.org/10.1016/j.amjoto.2023.103980

Xu, T., & Jumaat, N. F. (2024). ChatGPT-empowered writing strategies in EFL students' academic writing: Calibre, challenges and chances. International Journal of Interactive Mobile Technologies (iJIM), 18(15). https://doi.org/10.3991/ijim.v18i15.49219

This page has been intentionally left blank.