Artificial Intelligence and Academic Writing: A Case Study of STEM Undergraduates

Nino Jojua
Associate Professor,
European University, Georgia
Email: nino.jojua@eu.edu.ge

Abstract: *The emergence of generative artificial intelligence (AI), particularly tools such as ChatGPT, has introduced new complexities into the teaching and learning of academic writing. For students in STEM disciplines, who often regard writing as a secondary skill compared to technical expertise, AI offers unprecedented support in grammar, vocabulary, and organization. Yet, it simultaneously raises concerns related to authenticity, critical engagement, and academic honesty. This article presents findings from a mixed-methods study conducted with undergraduate STEM students enrolled in composition courses at San Diego State University Georgia. By combining textual analysis of student assignments with reflective surveys, the study identifies both opportunities and risks associated with AI-assisted writing. Results reveal improvements in surface-level accuracy, vocabulary range, and time management, counterbalanced by challenges in synthesis, overreliance on automated support, and instances of fabricated references. The study argues for a reimagined pedagogy that integrates AI not as a substitute for student thinking but as a scaffold for deeper engagement, fostering reflective practice, ethical awareness, and sustainable autonomy in writing development.*

Keywords: AI assisted learning, academic integrity, academic writing, generative artificial intelligence

INTRODUCTION

Generative AI has become an influential force in higher education, transforming not only how knowledge is accessed but also how it is produced. Within academic writing, tools such as ChatGPT can generate entire essays, reframe arguments, and correct linguistic errors within seconds. For instructors, this dual capacity to enhance and to undermine student learning requires careful pedagogical reflection.

The context of STEM education adds another dimension to the problem. STEM students often see writing as ancillary to their core technical training. As such, they are particularly prone to viewing AI as an expedient tool for completing writing tasks rather than as a means of cultivating communication skills. While AI may provide shortcuts to fluency and clarity, there is a danger that reliance on it could erode essential cognitive skills such as argumentation, synthesis, and critical inquiry.

Despite the rapidly growing body of scholarship on generative AI in education, several gaps remain in our understanding of how such tools affect student writing, particularly in STEM contexts. Much of the existing research has focused on either general improvements in language fluency (Tran, 2024) or on ethical concerns surrounding plagiarism and authenticity (Okaiyeto et al., 2023; Doshi & Hauser, 2024). Fewer studies have examined the *cognitive trade-offs* involved in AI use, such as the potential weakening of synthesis, critical thinking, and independent problem-solving skills (Stadler et al., 2024). Moreover, most empirical investigations have been conducted with humanities or language-major cohorts, leaving STEM students—who often approach writing as a secondary rather than central academic skill—largely underrepresented in the literature. This study addresses these gaps by focusing specifically on STEM undergraduates and analyzing not only the textual outcomes of their AI-assisted writing but also their reflective perceptions of its affordances and limitations.

Therefore, this research aims to illuminate how STEM undergraduates use ChatGPT in academic writing and what pedagogical consequences emerge from this practice. The study was guided by two questions:

1. How does AI-assisted writing manifest in the quality and features of student assignments?
2. What perceptions do students hold regarding the benefits and limitations of using AI in their writing process?
3. And based on the answers to the previous research questions, how to amplify benefits and reduce hazards of AI in writing.

LITERATURE REVIEW

The growing body of scholarship on generative artificial intelligence (AI) in education paints a picture of both promise and peril, with debates centering on how AI reshapes writing practices, learning habits, and academic integrity.

Productivity, Skill Bridging, and Integrity Risks

Recent work positions large language models (LLMs) as accelerators of the writing process: they help students overcome starting inertia, outline arguments, and tidy prose, thereby *bridging skill gaps* that are acute for multilingual and STEM writers who often treat composition as ancillary to technical learning (Ray, 2023). This productivity gain can democratize access to feedback and

models of academic discourse, reducing time-to-draft and improving document formatting. Yet the same affordances generate *integrity vulnerabilities*: when boundaries around acceptable support are ambiguous, AI-assisted paraphrase or wholesale generation risks sliding into *AI-enabled cheating*, fabricated sources, and opaque authorship (Okaiyeto, Bai, & Xiao, 2023). The literature thus emphasizes the need for explicit norms and transparent disclosure practices in assessment and instruction to prevent conflating fluency with authorship.

Creativity Gains vs. Declines in Collective Novelty

At the individual level, LLMs can *boost ideation* by supplying analogies, alternative framings, and rapid brainstorming prompts, which students perceive as creative momentum (Doshi & Hauser, 2024). However, when many learners draw from similar model distributions, a cohort-level effect appears: outputs converge toward median patterns, reducing the **diversity of novel ideas** across a class or field. For STEM composition—where solution multiplicity, problem framing, and methodological originality matter—this homogenizing tendency is consequential. The literature therefore recommends instructional designs that preserve divergent thinking (e.g., open-ended problem contexts, data- or lab-anchored prompts) while still leveraging AI for micro-level support.

Reduced Cognitive Burden and the Depth-of-Processing Trade-off

LLMs lower *cognitive load* by handling micro-decisions (wording, transitions) and macro-structuring (headings, paragraph flow). While such scaffolding can free working memory for higher-order reasoning, evidence indicates a **depth-of-processing trade-off**: students expend less mental effort and more readily accept plausible but shallow arguments, weakening synthesis, evaluation, and evidential reasoning (Stadler, Bannert, & Sailer, 2024). In writing-studies terms, polished surface features can mask brittle rhetorical architecture. This echoes process theories that treat composing as recursive problem solving; removing productive struggle can short-circuit learning opportunities embedded in planning, drafting, and revising.

Language Gains: Coherence, Cohesion, and Accuracy—With Affective Costs

Empirical classroom studies show measurable improvements in **coherence, cohesion, lexical range, and grammatical accuracy** when AI is used for drafting and micro-revision (Tran, 2024). These gains are especially salient for multilingual STEM students who must communicate technical content with precision. However, parallel findings document *affective costs*: habitual reliance on AI correlates with **insecurity when tools are unavailable** and a diminished sense of intellectual ownership (Nebieridze & Jojua, 2024). Students may reframe "good writing" as "LLM-polished writing," weakening self-efficacy and metacognitive control over choices such as stance, evidence selection, and genre moves.

Synthesis: A Double-Edged Pedagogical Problem

Taken together, the literature depicts AI as a **double-edged** intervention. It can equalize access to linguistic resources and expedite drafting (Ray, 2023; Tran, 2024), yet it risks eroding the very capacities writing instruction is meant to cultivate: originality at the cohort level (Doshi & Hauser, 2024), deep engagement with sources and problems (Stadler et al., 2024), and durable writerly agency (Nebieridze & Jojua, 2024). For STEM composition specifically, where genres demand precise claims, warranted evidence, and discipline-appropriate stance, improvements in surface form do not guarantee epistemic quality. The literature therefore points toward pedagogies that (a) keep the human *process* visible (planning, revision, source synthesis), (b) set **ethical boundaries** and disclosure expectations, and (c) design **context-rich, AI-resistant tasks** that require local data, lab artifacts, or in-class reasoning—so that AI functions as scaffold rather than substitute.

METHODOLOGY

This study employed a **mixed-methods design** to examine both the textual outcomes of AI-assisted writing and the subjective experiences of student users. The approach allowed for a comprehensive understanding of how generative AI tools, such as ChatGPT, influence academic writing practices in STEM disciplines.

Participants
The study involved **28 undergraduate STEM students** majoring in engineering, computer science, and natural sciences, all enrolled in an academic writing course. Participants' English proficiency ranged from **B2 to C1** according to the Common European Framework of Reference for Languages (CEFR), ensuring sufficient language skills to engage meaningfully with AI-assisted writing tasks.

Study Period
Data were collected over the **Fall semester (September–December 2024),** spanning three major assignments.

Data Sources

1. **Textual Corpus:** First-draft response papers submitted for the three assignments were collected to examine writing development and AI-assisted patterns.
2. **Survey:** A reflective questionnaire was administered to gather participants' attitudes toward AI-assisted writing, strategies employed, and perceptions of its benefits and limitations.

Rubrics and Analysis

The textual corpus was evaluated and coded using a structured rubric that focused on **organization, content development, grammar/mechanics, and adherence to APA style**. Quantitative analyses identified general performance patterns, while qualitative analyses highlighted anomalies and nuanced effects of AI assistance. Survey responses were subjected to **thematic analysis** to extract recurring themes, insights, and student reflections regarding AI use.

This mixed-methods framework enabled the study to triangulate findings, linking observable textual outcomes with learners' perceptions and experiences, thereby providing a holistic view of AI's impact on academic writing practices.

FINDINGS

Assignment Analysis

The assignment analysis has revealed the following findings and recurrent patterns:

- ✓ In 17 cases (approximately 60%) the writing of these students, who demonstrated difficulties in productive skills during the course, was characterized by virtually no errors in grammar, vocabulary, or punctuation, yet exhibited overly complex grammatical constructions and a tendency toward slightly elevated vocabulary.
- ✓ In 13 cases (approximately 46%), unrealistic essay organization, often comprising more than five to six short paragraphs, frequently accompanied by outline-like subtitles such as "Introduction," "Point 1," and "Conclusion," was identified reflecting a direct imitation of the prompt's metadiscourse rather than an independently structured argument.
  Additionally, the papers also demonstrated Content peculiarities such as instead of synthesizing several authors' ideas, summarizing each of them separately in different paragraphs.

- ✓ In rare cases (approximately 7 %), in- text citations and referencing such authors that were not pre- assigned by the instructor and, upon diligent checking, simply could not be retrieved and researched;
- ✓ In the majority of the papers (20 out of 28=approx. 71%) the reference page was accurately compiled according to the required APA which used to be a very problematic area for students before.
- ✓ Almost all the papers (approx. 89%) were uploaded on time demonstrating an unusual time management on the given course.

To reflect on the the above-given findings, they suggest a significant impact of AI-generated assistance on students' writing practices

- ✓ **Organization-wise** excessive segmentation of content, paragraphing and the use of prompt-derived subtitles, indicating a lack of organic structural development and coherence.

- ✓ **Content-wise,** a surface-level engagement with academic argumentation and difficulty synthesizing multiple sources, instead opting to summarize them individually; ; At times fabricated references, pointing to a misuse or misunderstanding of academic integrity when using AI tools.

- ✓ **Language wise:** while a majority of papers were free from surface-level errors such as grammar, vocabulary, and punctuation, this fluency was often paired with unusually complex sentence structures and elevated vocabulary;

- ✓ **Formatting - wise:** accurately formatted APA reference pages—an area that had previously posed challenges for most students.

- ✓ **Transferrable skills-wise:** Time management and respect for university deadlines: improved time management, potentially linked to AI assistance.

Survey Insights

The survey demonstrated varying levels of ChatGPT's effects on students across three types of critical thinking tasks: *problem-solving, creative tasks, and text analysis.*

Problem-solving

- **Independent Thinking**
  - Solve independently (n = 22, 38.6%)
- **Involvement of ChatGPT**
  - Have ChatGPT refine solutions (n = 14, 24.6%)
  - Have ChatGPT provide guidance (n = 18, 31.6%)
- **Complete Dependence**
  - Have ChatGPT do most of the problem (n = 1, 1.8%)
  - Have ChatGPT solve the problem (n = 2, 3.5%)

Creative Tasks

- **Independent Thinking**
  - Brainstorm ideas (n = 27, 47.4%)
- **Involvement of ChatGPT**

- Generate ideas after ChatGPT simplifies the task (n = 18, 31.6%)
  - Select from several of ChatGPT's options (n = 10, 17.5%)
- **Complete Dependence**
  - Have ChatGPT generate every aspect of the idea (n = 2, 3.5%)

Text Analysis Tasks

- **Independent Thinking**
  - Read and analyze (n = 30, 52.6%)
- **Involvement of ChatGPT**
  - Analyze the version shortened by ChatGPT (n = 17, 29.8%)
- **Complete Dependence**
  - Have ChatGPT analyze the text without one's involvement (n = 4, 7.0%)

* The above data present the most important findings, revealing that all three types of critical tasks were performed both independently and with varying degrees of ChatGPT involvement.

To summarize, the findings, it could be concluded that while users value ChatGPT's support, they predominantly engage with it as a collaborative tool rather than relying on it entirely, maintaining a meaningful level of autonomy in their critical thinking processes; specifically,

a) many students use ChatGPT just as an assistant to refine*, guide, or simplify tasks* especially evident from problem-solving (56.2%) and creative tasks (49.1%).
b) Only a very small number of students (with the highest being only 7.0%- in text analysis) are completely dependent on ChatGPT in terms of Critical Thinking abilities;

DISCUSSION

The findings confirm that AI's impact is ambivalent: it improves formal aspects of writing while undermining deeper cognitive and ethical dimensions.

- **Benefits**: AI reduces barriers to fluency, allowing students to express complex ideas more clearly. It also improves time efficiency, enabling students to meet deadlines more consistently.
- **Risks**: However, the cost is a weakened ability to synthesize diverse sources, overreliance on AI's structural defaults, and occasional academic misconduct. Importantly, some students equated "good writing" with AI-polished texts, raising questions about their perception of authorship.

This suggests that AI integration should not be dismissed but rather reframed as an opportunity for guided learning. If used reflectively, AI can provide scaffolding that empowers rather than diminishes student agency.

Pedagogical implications on how to amplify benefits and reduce hazards of using AI in writing

Instructional Strategies

Emphasize Process Over Product.

One of the most pressing risks of generative AI in writing is that it produces highly polished final texts that obscure the underlying cognitive and rhetorical work of composing. As Flower and Hayes (1981) remind us, writing is a recursive process of planning, drafting, and revising, not simply the production of a finished artifact. When students rely on ChatGPT, there is a temptation to bypass this process, limiting opportunities for genuine problem-solving and meaning-making. To counter this, instructors should emphasize drafting, peer review, and iterative revision as core course components. Such process-oriented approaches align with evidence from Stadler et al. (2024), who warn that AI's reduction of mental effort can lead to superficial engagement with academic tasks.

Structured Reflection.

Metacognitive engagement is key to preventing overdependence on AI. Zimmerman (2002) demonstrates that self-regulated learning improves when students explicitly monitor their strategies and decision-making. Requiring reflective journals, annotations, or short process reports about how AI shaped their choices can ensure accountability and cultivate awareness of authorship. Nebieridze and Jojua (2024) further argue that reflection reduces "AI anxiety" by prompting students to evaluate when and why they rely on technology, which supports the development of resilient and adaptive learning habits.

AI-Resistant Prompts.

Assignment design also plays a crucial role. Cotton et al. (2023) suggest that tasks requiring local context, personal experience, or integration of classroom dialogue are less easily replicated by AI and therefore more authentic assessments of student knowledge. In STEM writing courses, prompts might ask students to connect technical principles to real-world case studies from their own laboratories or communities. Tran (2024) notes that such personalization not only discourages generic AI responses but also deepens student investment in the writing process.

Blended Assessment.

Finally, a balance of AI-assisted and unaided writing tasks allows instructors to assess both skill sets. Farrokhnia et al. (2023) recommend hybrid assessment designs where take-home essays are complemented by in-class writing sessions. This approach ensures that students practice using AI as a supportive tool while also demonstrating independent competence. In STEM contexts, where

time efficiency is highly valued, such blended models preserve the advantages of AI while protecting against the erosion of essential cognitive and communicative skills.

Ethical and Responsible Use

Explicit Guidelines.

Ambiguity around acceptable AI use can foster misuse. Okaiyeto et al. (2023) emphasize the need for institutional policies that distinguish between legitimate support (e.g., grammar correction) and academic misconduct (e.g., submitting AI-generated essays as original work). Clearly articulated course policies, alongside discussions about citation practices for AI outputs, help students navigate these ethical boundaries.

Scaffolded Use.

Rather than banning AI, a more effective approach is to scaffold its use for specific subtasks. For example, students might use ChatGPT to brainstorm ideas or check grammar, but then be required to rewrite outputs independently. This aligns with constructivist approaches to learning, where external supports serve as temporary scaffolds to build internal competence (Vygotsky, 1978). Such strategies mitigate overdependence and promote gradual autonomy.

Critical Comparison.

Encouraging students to critique AI-generated texts can transform potential risks into teachable moments. By comparing AI drafts with their own work, students learn to identify issues such as shallow argumentation, fabricated references, or lack of disciplinary depth. This practice echoes Doshi and Hauser's (2024) concern that AI tends to homogenize ideas, highlighting the importance of cultivating critical judgment. Instructors can leverage this by assigning "AI critique" tasks where students analyze, improve, and humanize machine-generated drafts.

CONCLUSION

AI tools like ChatGPT are here to stay, and their role in education is growing. For STEM students, these tools can either widen or narrow the gap between technical competence and communicative ability. This study underscores that the challenge is not AI itself, but how educators choose to integrate it: adopting the hybrid collaborative approach, such as intentional and reflective AI integration in teaching writing highlights the potential for AI tools *to augment, rather than replace*, *human thinking* fostering a learning environment where critical thinking, creativity, and independent academic voice can be preserved on the one hand, while technological strengths harnessed; and transforming traditionally used teaching methods into more digitally friendly learning strategies. So with reflective practice, scaffolded pedagogy, and ethical awareness, AI can

be harnessed to to maximize the benefits and overcome the challenges potentially posed by AI in writing.

REFERENCES

Doshi, A. R., & Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances, 10,* eadn5290. https://doi.org/10.1126/sciadv.adn5290

Nebieridze, M., & Jojua, N. (2024). Why do students use ChatGPT and why does it matter? *Online Journal of Humanities ETAGTSU, (9).* https://doi.org/10.52340/PUTK.2024.2346-8149.09.12

Okaiyeto, A., Bai, J., & Xiao, H. (2023). Generative AI in education: To embrace it or not. *International Journal of Agricultural and Biological Engineering, 16*(3), 285–286. https://doi.org/10.25165/j.ijabe.20231603.8486

Ray, P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations, and future scope. *Internet of Things and Cyber-Physical Systems.* https://doi.org/10.1016/j.iotcps.2023.04.003

Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior, 160,* 108386. https://doi.org/10.1016/j.chb.2024.108386

Tran, T. T. H. (2024). AI tools in teaching and learning English academic writing skills. *Proceedings of the AsiaCALL International Conference, 4,* 170–187. https://asiacall.info/proceedings/index.php/articles/article/view/86/68