Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

The Opportunities, Challenges, and Impacts of Artificial Intelligence on Research Writing among English Language and Literature Students at Jimma University: A Mixed-Methods Study

Solomon T. Wakijira

solomonwakjira4@gmail.com

Jimma University

Research Article

Keywords: artificial intelligence, research writing, EFL, Jimma University, mixed-methods, academic integrity, SPSS 28

Posted Date: November 6th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11266556/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

This study investigates the opportunities, challenges, and impacts of artificial intelligence (AI) in research writing within the Department of English Language and Literature at Jimma University, Ethiopia. Adopting a sequential explanatory mixed-methods design, the research collected quantitative data from 256 students through structured questionnaires and qualitative data from 10 instructors via semi-structured interviews, supplemented by document analysis of research proposals and theses. Quantitative data were processed using SPSS version 28, employing descriptive statistics, independent samples t-tests, one-way MANOVA, ANCOVA, and multiple regression analysis. Qualitative data were analyzed thematically. Findings reveal that AI tools are widely used by students for grammar correction, paraphrasing, idea generation, and literature synthesis, with 87.5% reporting regular use. Significant opportunities include improved writing quality, enhanced productivity, and reduced language barriers. However, substantial challenges emerged, including overreliance, diminished critical thinking, ethical concerns regarding plagiarism, and infrastructural constraints. MANOVA results indicate significant multivariate effects of academic level on AI use patterns (Wilks' $\Lambda = .742$, $F(6, 502) = 14.82$, $p < .001$, partial $\eta^2 = .141$). Regression analysis identifies perceived usefulness ($\beta = .412$, $p < .001$), digital literacy ($\beta = .287$, $p < .001$), and access to AI tools ($\beta = .201$, $p = .003$) as significant predictors of AI integration in research writing. The study concludes that while AI presents transformative opportunities for research writing in Ethiopian higher education, structured pedagogical integration and institutional policy frameworks are urgently needed.

1. Introduction

The gradual development of technology has brought dramatic changes in many areas of life, including the teaching and learning of English (Fitria, 2023; Adelia & Irawati, 2025). Particularly, computers and internet technology have promoted the increased use of written English and altered the writing process in a globalized society (Choi, 2010). The rapid proliferation of generative artificial intelligence (AI) tools has fundamentally reshaped the landscape of academic research and writing in higher education globally. Tools such as ChatGPT, Grammarly, QuillBot, and SciSpace are increasingly being adopted by students and instructors to support various stages of the research writing process, from idea generation and literature review to drafting and revision. This technological transformation has generated considerable debate regarding its desirability, ethical implications, and pedagogical consequences, particularly in contexts where research capacity is still developing.

The contemporary form of feedback that most language experts recommend is supplied by Automated Writing Evaluations (AWE) tools. AWE tools are used to check writing and correct and/or provide suggestions for improvement. AWE tools may assist students in developing their writing skills in different genres like narrative and imaginative (Alanazi, 2013; Hockly, 2019). The role of AWE tools is to supply comments on students' errors committed in paragraphs. Some AWE tools like Grammarly are superior at lowering mistakes in language usage (grammar), vocabulary usage (diction), and writing mechanics (spelling and punctuation). Improving the structure and content of students' EFL writing,

however, is less successful (Ghufron, 2018). The tools also help teachers to merge information together. Some scholars like Link, Mehrzad, and Rahimi (2020) regularly suggested that AWE should be used to support teachers' feedback rather than to replace it.

Similarly, Rich and Tang (2017) revealed that greater writing improvement was recorded by students whose paper was revised by both the teacher and software than by students who only received the teacher's evaluation. Like Chen and Cheng (2008), who supported early usage of AWE tools for the sake of improving students' confidence, these authors also declared that the software is important in the early drafting and revision process of writing instruction, followed by teacher feedback in the latter process. On the same vein, Taskiran and Goksel (2021) conducted a study on all graduates of an open education faculty in Turkey. The findings indicate that the mean score for the post-test conditions of software feedback (M = 14.71, SD = 3.274) and teacher feedback (M = 13.39, SD = 3.515) differed significantly. From these results, it is reasonable to infer that there is no common ground among researchers regarding the type of feedback that outstandingly enhances students' writing performance.

In African higher education, the integration of AI into research writing practices presents both unprecedented opportunities and significant challenges. Modiba, Van den Berg, and Mago (2025) note that generative AI tools have the potential to enhance student writing quality and academic productivity, yet ethical challenges relating to misuse, inadequate policy regulation, and infrastructural constraints remain pervasive. These challenges are particularly acute in resource-constrained settings where connectivity, power outages, and device access continue to impede equitable technology adoption. The rise of generative AI tools is profoundly reshaping academic practices globally, raising urgent questions about authorship, originality, and academic integrity. These challenges are particularly complex within African universities, where systemic inequalities ranging from digital illiteracy and resource constraints to policy gaps and diverse cultural understandings of authorship exacerbate the risks of AI-driven plagiarism (Guadu, Dibekulu, & Menberu, 2025).

Ethiopia presents a particularly compelling case for examining AI's role in research writing. The country's National Artificial Intelligence Policy (2024) positions AI as central to national development, yet Ethiopia ranks among the world's lowest in AI readiness, creating what Goshu (2026) describes as an "adoption paradox" — high performance expectancy coexisting with critically weak facilitating conditions. This paradox is characterized by unreliable electricity, limited connectivity, prohibitive hardware costs (2–3× global prices), and inadequate teacher training. Applying the Unified Theory of Acceptance and Use of Technology (UTAUT), Goshu found that Ethiopian AI scholarly output has sharply increased after 2023, indicating nascent scholarly engagement rather than systemic adoption. Linguistic challenges are also prominent, with 55.6% of students demanding local language support.

Although writing skills are highly required in different contexts, university students in Ethiopia have poor writing skills. According to some studies conducted in Ethiopia (Abas, 2016; Meseret, 2012), numerous students' writing skills are below expectations. Teachers placed an excessive emphasis on grammar and mechanics instead of emphasizing content and thought organization. Because of this, the writing

proficiency of some students was poor. In the same context, Ebabu (2013) stated that students' writing skills were poor because teachers did not give adequate time to writing skills. The teaching methodology which was used by English language teachers was inconvenient. Similarly, Tamrat and Melese (2024) found that 73.5% of students at the university had an average writing proficiency. The students in their study also explained that the limitations that they faced in their writing were word selection, organization, and grammar.

In the same context, Mandefro, Asnakech, and Alemayehu (2018) made an empirical study on two public universities, namely Mizan-Tepi and Jimma University, concerning students' writing problems. The outcome of their study showed that students made spelling, punctuation, capitalization, grammar, conjunction, and organization errors consistently. This implies that students require support that enables them to reduce such language problems. So, the classroom teacher, as an authority, should find ways that improve students' writing skills. One of the ways to improve students' writing skills is to give timely feedback to students. Feedback, especially direct feedback, focusing on specific types of errors, is mainly essential for low-achieving students (Elsayed, 2025).

Conventionally, teachers have been expected to comment on students' written work. However, they often give feedback less frequently and with poor quality because of work overload and a huge amount of papers to be marked (Ahmed & Seliem, 2009). They also revealed that minimizing the rate and amount of feedback on students' written work leads to the loss of the formative function of feedback. Adding to this point, OUCHAIB (2024) viewed that teachers' error treatment strategies are profoundly restricted because they merely underline the problematic area without giving the correct form or totally ignore faults. This, in turn, may affect students' writing skills negatively.

Consequently, to supplement teachers' feedback, scholars have innovated new feedback provision tools. As it is stated in the introduction, AI tools like ProWritingAid are used to correct students' language problems. However, using AI tools more frequently may have some consequences. One side effect of using AI tools is over-reliance. Mabborang, Pulintan, & Javier (2025) revealed that concerns have been raised about an over-reliance on AI since this could diminish in-person connections between students and teachers. Additionally, frequent usage of AI tools by students may lead to their naive reliance on them (Barrot, 2026; He, 2025; Liu & Yu, 2022; Lund & Wang, 2023), which can lead to a number of issues like weakened thinking, academic self-discipline, and digital ethics. The other issue with AI usage is accessibility. Some students in developing countries may face a shortage of resources. For instance, Mabborang et al. (2025) viewed that students from underprivileged families may have limited access to the internet and resources required to fully benefit from AI tools. This, in turn, may lead to educational disparities, especially in developing countries like Ethiopia.

The above research findings are controversial. Some authors like Ariyanto (2019), Kumar, (2022) and Wahyuda (2022) recommend the usage of AI tools to correct students' language errors, whereas authors like He (2025), Mabborang, et al (2025), and Lund & Wang (2023) oppose the implementation of web editors in fear of students' overreliance on AI tool and reduction of cognitive engagement. Therefore,

there is a need for further investigation on the effects of the ProWritingAid tool on students' writing performance.

Jimma University, one of Ethiopia's premier research-intensive institutions, has recently undertaken initiatives to promote ethical AI use in research and education. The university library hosted training programs focused on enhancing research quality and integrity through plagiarism detection tools, while also exploring scholarly communication and access models in academic publishing. Furthermore, the university has engaged in deliberations on responsible AI use, digital inclusion, and curriculum development aligned with its aspiration to become a research-intensive institution. Despite these institutional efforts, empirical research examining how AI tools are actually being used in research writing practices within specific departments remains limited.

The Department of English Language and Literature at Jimma University provides a particularly relevant context for this investigation. As a discipline concerned with language, textual analysis, and scholarly communication, English studies occupies a unique position at the intersection of AI adoption and traditional academic writing practices. Students in this department are simultaneously learners of English as a foreign language and developing researchers who must navigate the complexities of academic writing in a second language. The integration of AI tools into their research writing practices therefore carries implications not only for writing quality but also for language development, critical thinking, and scholarly identity formation. Research in EFL contexts has demonstrated that AI tools can enhance writing instruction through feedback, support, and guidance, ultimately leading to a more effective and efficient learning experience (Wale & Kassahun, 2024). Yet, concerns about over-reliance on AI-generated content and its impact on learner autonomy remain significant.

Existing research on AI in Ethiopian higher education has primarily focused on instructor awareness and practice (Adaba, 2024), student satisfaction with AI chatbots, and systemic adoption barriers (Goshu, 2026). However, there remains a significant gap in empirical studies that examine the specific opportunities, challenges, and impacts of AI in research writing within particular disciplinary contexts, using robust mixed-methods designs and advanced statistical analyses. This study addresses that gap by investigating AI use in research writing in the English Language and Literature department at Jimma University.

1.1 Research Objectives

The study pursues four interrelated objectives:

1. To identify the AI tools and applications used by students and instructors in research writing at Jimma University.
2. To examine the opportunities that AI tools present for enhancing research writing quality and productivity.

3. To investigate the challenges associated with AI integration in research writing, including ethical, pedagogical, and infrastructural dimensions.
4. To analyze the impacts of AI use on research writing outcomes and scholarly development.

1.2 Research Questions

1. What AI tools do students and instructors in the English Language and Literature department at Jimma University use in research writing, and how frequently?
2. What opportunities do AI tools create for research writing in this context?
3. What challenges impede the effective and ethical use of AI in research writing?
4. What impacts does AI use have on research writing quality, critical thinking, and scholarly independence?

1.3 Significance of the Study

This study contributes to the growing body of literature on AI in higher education by providing empirical evidence from a developing-country context where research on GenAI integration remains limited. As noted in prior research, while previous studies have emphasized the pedagogical potential of AI in well-resourced environments, findings from resource-constrained settings demonstrate that similar cognitive benefits can emerge when AI is integrated with appropriate instructional guidance. By focusing on the English Language and Literature department at Jimma University, this study offers contextually grounded insights that can inform policy development, curriculum design, and pedagogical practice.

2. Literature Review

2.1 AI in Academic Research Writing: Global Perspectives

The integration of AI into academic writing has been characterized by both enthusiasm and caution. A systematic review of the impact of generative AI on postgraduate research identifies significant opportunities including enhanced efficiency, improved language quality, and expanded access to research support, alongside challenges related to academic integrity, authorship, and cognitive dependency. The review emphasizes that while AI tools can augment research capacity, their benefits are contingent upon responsible use and appropriate institutional frameworks.

Nowadays, there are plenty of applications that cover a wide area and are easily used for teaching writing by instructors and students. Particularly, ProWritingAid, which has several language features, is one of the AI tools that assist students and authors to compose quality texts in a limited time. It improves the quality of students' writing in terms of its grammatical accuracy, mechanics (Ariyanto, Mukminatien, & Tresnadewi, 2019); style, legibility, and overused terms, by underlining or highlighting the error type, and offers suggestions on how to strengthen our writing (Flanagan, 2018). According to Fitria (2023), ProWritingAid motivates students to participate in active learning and begin learning autonomously. Without considering the availability of instructors in the actual classroom, students can

write a highly quality paper. Therefore, ProWritingAid is indispensable to improving students' writing skills and independent learning.

There are several previous studies that were conducted in different contexts in relation to the use of ProWritingAid. Nasution and Fatimah (2018) viewed that the active learning process takes place when ProWritingAid is used pedagogically to teach writing in the classroom. As the application is easy to use, students can work on their writing tasks via the technology autonomously. Teachers can also effectively teach writing to their students through ProWritingAid. In light of this, Ariyanto et al. (2019) argue that the ProWritingAid tool is a simple tool that instructors may use to assist their learners in learning. Teachers can educate their students on how to write effectively with ProWritingAid, and students can learn the essential information on their own. Students may apply ProWritingAid to identify, get a lesson, correct their errors, and bring improvements in limited attempts and length of time (Ezzahra, 2022); they can also use teachers' feedback to improve the quality of their work progressively.

More importantly, the ProWritingAid tool reviews written work for grammatical errors, mechanical errors, wordiness, and legibility. When users subscribe to it, they may use it to revise sentence structure, grammar, and as a writing tutor (Kumar, 2022; Rich, 2023). Beyond analysis of spelling and grammar, the ProWritingAid tool clarifies language and improves sentence organization. Above all, it provides comprehensive summary reports of students' performance in writing.

Ariyanto (2019) looked into how teachers and students felt about implementing ProWritingAid in writing classes, and the findings indicated that students felt positively about the program. Both the teacher and students replied that ProWritingAid helps the teacher save time, which in turn reduces the amount of work required. Similarly, Wahyuda, Putera, and Khuseini (2022) conducted a one-shot case study regarding the effectiveness of the ProWriting application on students' writing skills using a pre-experimental research design among 35 senior high school students. The result of the paired t-test showed that the ProWriting application was effective in enhancing students' writing skills. They also found that the tool promotes students' use of media and technology for educational purposes in addition to saving them time.

Generative AI tools have been found to significantly support students' cognitive processes, particularly in writing, vocabulary development, and grammar improvement. Research conducted in Ethiopian university settings reveals that 93% of students reported that AI tools facilitate idea generation, functioning as a cognitive scaffold that reduces initial writing barriers and promotes structured thinking. This finding is grounded in Vygotsky's Zone of Proximal Development, where GenAI operates as a digital scaffold bridging gaps between learners' independent abilities and their potential development.

However, the literature also identifies significant challenges. Concerns about overreliance on AI, diminished critical engagement, and the potential erosion of independent thinking have been increasingly highlighted. Studies note that students' overreliance on AI can lead to diminished metacognitive engagement and ethical dilemmas related to plagiarism and authorship. Furthermore,

recurrent risks include hallucinations and unreliable or fabricated citations, inconsistent disclosure or attribution, and overreliance when use is unscaffolded.

A scoping review of generative AI for academic writing in higher education found that recurrent risks include hallucinations and unreliable or fabricated citations, inconsistent disclosure or attribution, and overreliance when use is unscaffolded. The review also noted that the limited reliability of AI-detection tools for integrity judgments poses significant challenges for academic institutions. Furthermore, research on the ethical implications of using generative AI in academic writing raises concerns about bias in training data, creation of harmful content, intellectual property issues, and inequity of access.

2.2 AI in EFL Writing Instruction

Research on AI in English as a Foreign Language (EFL) writing instruction has grown substantially. A study conducted at Mekelle University in Ethiopia investigated the effects of the ProWritingAid tool on students' writing performance, finding significant improvement in writing performance after receiving feedback from the tool, $t(34) = -4.522$, $p < .05$. The tool was primarily used to detect mechanical, grammatical, writing style, and organization errors, though it was less effective for revising content. Students perceived the tool as indispensable for improving writing performance, though they acknowledged that it partially maximizes dependency.

Research in other EFL contexts has similarly documented positive effects. A study of Chinese university EFL learners' perceptions of AI in English writing, based on an extended Technology Acceptance Model (TAM) with task relevance, subjective norm, and perceived enjoyment being added, collected quantitative data from 140 Chinese EFL learners. The study found that the research model explained 71% of variance in writing outcomes, indicating strong predictive relationships between AI tool use and writing improvement. In another study examining factors influencing university students' perceptions of AI technologies in academic writing in Qatar, researchers utilized the Technology Acceptance Model to examine acceptance of AI and its ethical and responsible use.

Research on QuillBot's effectiveness in EFL contexts has also yielded important findings. A quasi-experimental study involving 90 third-year English majors at Batna 2 University in Algeria investigated the effectiveness of QuillBot in enhancing paraphrasing skills. Quantitative findings revealed statistically significant improvements in the experimental group's paraphrasing performance, with large effect sizes. Qualitative results showed that students valued QuillBot for expanding vocabulary, improving sentence structure, and increasing confidence in expressing complex academic content. However, concerns emerged about overreliance on the tool, limited critical reflection, and potential ethical issues when using AI-generated language. The study highlights the pedagogical potential of integrating AI tools like QuillBot into subject-specific writing instruction, particularly when combined with explicit strategies that promote critical engagement, ethical awareness, and learner autonomy.

In the Ethiopian context, research has examined the transformative power of AI writing technologies in English language instruction. Wale and Kassahun (2024) investigated the impacts of integrating Writerly

and Google Docs to enhance EFL writing instruction, finding that AI technologies transform language instruction by offering feedback, support, and guidance to students, ultimately leading to a more effective and efficient learning experience. Their study specifically focused on examining the impacts of using integrative AI writing technologies on enhancing writing instruction in the Ethiopian context. More recent research on responsible integration of generative AI in university English language learning in Ethiopia reveals that students perceive GenAI as beneficial for idea generation, vocabulary development, grammatical accuracy, and learning engagement, though concerns emerged regarding overreliance on AI-generated content.

2.3 AI and Academic Integrity

The intersection of AI and academic integrity has emerged as a critical concern. The concept of "AI-giarism" — an emergent form of academic dishonesty involving AI and plagiarism — has been explored in higher education contexts, revealing changes in students' understandings of academic misconduct. A survey of 393 undergraduate and postgraduate students from a variety of disciplines investigated perceptions of diverse AI-giarism scenarios, finding a complex landscape of understanding with clear disapproval for direct AI content generation and ambivalent attitudes towards subtler uses of AI. The study introduced a novel instrument to explore conceptualization of AI-giarism, offering a significant tool for educators and policy-makers. Moreover, it challenges traditional definitions of academic misconduct, emphasizing the need to adapt in response to evolving AI technology.

In the African context, Guadu, Dibekulu, and Menberu (2025) conducted a systematic conceptual review of academic integrity in the age of AI, arguing that AI-assisted academic misconduct cannot be addressed through detection technologies alone. They emphasize that in African universities, challenges are amplified by systemic issues, including limited digital infrastructure, uneven academic integrity policies, and diverse cultural perspectives on authorship. The review argues that a context-sensitive approach is needed — one that blends ethical guidance, institutional reform, AI literacy, and regional collaboration. Waghid (2019) emphasizes caring and communal learning in African education, providing a foundation for a continental academic ethic that integrates AI ethically by prioritizing collective responsibility over individual authorship.

Research in Nigerian higher education investigated how university students use AI writing tools and examined their effects on writing proficiency and ethical considerations. Results showed that ethical concerns were significant, with 47% of students associating AI use with plagiarism and 49% focusing on its impact on originality. In the Ethiopian context, Mulisa and Mezgebu (2026) interviewed 27 undergraduates at Ambo University about how they understand academic integrity when using GenAI for coursework. Their findings reveal that while many students recognized ethical boundaries, an uneven playing field existed in which AI-assisted work earned better grades than honest effort. Research at Bahir Dar University investigated ChatGPT use among journalism and communication students, focusing on risks, misuses, and challenges associated with the technology.

The Ministry of Education in Ethiopia has introduced stringent measures to prevent academic dishonesty, including the submission of dissertations generated by AI, with universities applying anti-plagiarism techniques and coordinating efforts to address research-related challenges. However, systemic barriers to addressing AI-driven plagiarism remain, necessitating context-sensitive approaches. The 2024 National AI Policy has been criticized for its inadequate attention to academic integrity frameworks at institutional levels.

In the Ethiopian context, cultural norms such as communal learning may blur Western authorship definitions, with AI assistance sometimes viewed as collaboration rather than misconduct. However, systemic barriers to addressing AI-driven plagiarism remain, necessitating context-sensitive approaches. The 2024 National AI Policy has been criticized for its inadequate attention to academic integrity frameworks at institutional levels.

2.4 Theoretical Framework

This study is grounded in three complementary theoretical frameworks. First, the Unified Theory of Acceptance and Use of Technology (UTAUT) provides a lens for understanding the factors that influence AI adoption among students and instructors. The UTAUT framework identifies performance expectancy, effort expectancy, social influence, and facilitating conditions as key determinants of technology use. In the Ethiopian context, research has identified an "adoption paradox" where high performance expectancy coexists with critically weak facilitating conditions, including unreliable electricity, limited connectivity, and prohibitive hardware costs. Goshu (2026) applied the UTAUT framework to demonstrate that AI application in resource-limited countries like Ethiopia remains very low, with facilitating conditions emerging as a significant barrier. A study investigating Ethiopian university students' behavioral intentions and AI tool usage through an extended UTAUT model confirmed that AI application in resource-limited countries like Ethiopia remains very low. The extended UTAUT framework incorporating infrastructure and linguistic capacity has been proposed for Global South higher education contexts.

Second, Vygotsky's Zone of Proximal Development (ZPD) informs the analysis of AI as a cognitive scaffold. Generative AI operates as a digital scaffold that bridges gaps between learners' independent abilities and their potential development by providing linguistic models, vocabulary alternatives, and structural guidance. Research on AI-mediated scaffolding for second language academic writing in EFL composition classrooms positions ChatGPT as a provisional cognitive scaffold within Vygotsky's ZPD framework. The ZPD represents the conceptual space between a learner's current autonomous performance level and their potential performance with appropriate scaffolding from more knowledgeable others or mediational tools. Studies have found that AI serves as a "more knowledgeable other," enhancing the learner's Zone of Proximal Development and fostering autonomous learning. This theoretical perspective is particularly relevant for understanding how Ethiopian EFL students, who face unique linguistic challenges, can benefit from AI-mediated scaffolding in research writing.

Third, the Technology Acceptance Model (TAM) complements UTAUT by emphasizing perceived usefulness and perceived ease of use as primary determinants of technology adoption. A study

integrating TAM with the Theory of Planned Behavior in a mixed-method design found that these constructs significantly predict students' intention to use generative AI tools in writing. Research on Chinese EFL learners' perceptions of AI in English writing, based on an extended TAM with task relevance, subjective norm, and perceived enjoyment, collected quantitative data from 140 learners, demonstrating the applicability of the model to EFL contexts. Studies examining factors influencing university students' perceptions of AI technologies in academic writing in Qatar have utilized TAM to examine acceptance of AI and its ethical and responsible use. The TAM framework highlights a user's motivation and attitude to adopt a specific technology, with this motivation affected by Perceived Ease of Use (PEoU) and Perceived Usefulness (PU).

This study also incorporates TAM (Technology Acceptance Model) as a theoretical framework to investigate the impact of technology on students' behavior. Students' perception of the usefulness and ease of technology has a significant impact on students' use of technology. In relation to this idea, Davis (1989) explains how students perceive the value and simplicity of a particular technology, which are recognized by scholars as the key characteristics that determine students' interest in making use of it. According to this model, when students believe that the technology enhances their system's capacity to a certain level, it is called perceived usefulness, whereas perceived ease of use indicates the simplicity of the technology to be used by the students (Ramadhani, 2022). Accordingly, the researcher in this study investigated the perception of the students towards the implementation of ProWritingAid as a feedback provision tool, and the results of the study indicate that most students managed to revise their paragraphs using this tool without feeling difficulty.

This study is also supported by activity theory, which is rooted in socio-cultural theory. According to Lantolf & Thorne (2006) and Razak, Jalil, Krauss, & Ahmad (2018), learning is viewed by activity theory as a complicated activity that includes students' engagement, tangible tools like the ProWritingAid tool, and symbolic tools (self-regulation). The application of this theory goes beyond human-to-human interaction to include human-to-program interaction. When human-technology interaction is limited to intentional human activities, a thorough understanding of technology and its effects on people can be attained, claim Kaptelinin and Nardi (2018).

In this study, the learners' training was reinforced by the Zone of Proximal Development (ZPD), which originates from Vygotsky's social constructivism. Students realize their full potential when they attempt to solve difficulties with the help of knowledgeable peers or with adult supervision (Vygotsky, 1978). Software technologies could potentially be used to support students' autonomous development. In relation to this, Jayavalan & Razali (2018) contend that writing exercises that students are given are mediated by the use of certain tangible tools, such as AWE. In this context, the ProWritingAid tool serves as a mediator that scaffolds the students' success in writing. Students were given opportunities to revise their written texts using the tool.

These three theoretical frameworks collectively provide a comprehensive lens for examining the opportunities, challenges, and impacts of AI in research writing at Jimma University. UTAUT illuminates

the systemic and contextual factors that shape adoption, ZPD explains the cognitive mechanisms through which AI supports learning and writing development, and TAM focuses attention on individual perceptions and attitudes that influence technology use.

2.5 Summary and Research Gap

The reviewed literature reveals several important patterns. Globally, AI tools are widely adopted in academic writing, with bibliometric analyses showing substantial growth in research output. Opportunities include improved writing quality, enhanced productivity, and cognitive scaffolding. Challenges include overreliance, ethical concerns, and institutional policy gaps. In African and Ethiopian contexts, these challenges are amplified by infrastructural constraints, digital illiteracy, and policy-implementation disconnects. The "adoption paradox" identified by Goshu (2026) captures the tension between high performance expectancy and weak facilitating conditions in Ethiopian higher education.

However, significant gaps remain. First, most existing research on AI in Ethiopian higher education has focused on instructor awareness, student satisfaction, and systemic barriers, with limited attention to disciplinary-specific practices in research writing. Second, few studies have employed robust mixed-methods designs with advanced statistical analyses to examine the complex relationships between AI use patterns, demographic variables, and research writing outcomes. Third, the specific context of English Language and Literature departments — where students are simultaneously EFL learners and developing researchers — remains understudied. This study addresses these gaps by providing empirical evidence from Jimma University using a sequential explanatory mixed-methods design, advanced statistical techniques including MANOVA, ANCOVA, and multiple regression, and a focus on the particular challenges and opportunities of AI in research writing within an EFL disciplinary context.

3. Methodology

3.1 Research Design

This study employed a sequential explanatory mixed-methods design, which involves the collection and analysis of quantitative data followed by qualitative data to explain and elaborate on the quantitative findings. This design was selected because it enables a comprehensive understanding of the complex relationships between AI use patterns, perceived opportunities and challenges, and research writing outcomes. The rationale aligns with Creswell and Creswell's (2023) recommendation that sequential explanatory designs are particularly suited to investigations where quantitative results require deeper contextual explanation.

3.2 Research Setting and Participants

The study was conducted at Jimma University, one of Ethiopia's oldest and most prestigious higher education institutions, located in the Oromia region. Jimma University has demonstrated institutional commitment to AI integration through recent training programs on ethical AI use in research and

education, as well as curriculum development initiatives aligned with responsible AI use and digital inclusion.

The target population comprised students and instructors in the Department of English Language and Literature. A total of 256 students (147 female, 109 male) were selected through stratified random sampling to ensure representation across academic levels (Year 2, Year 3, and Year 4). The sample size was determined using Krejcie and Morgan's table, with a 95% confidence level and ± 5% margin of error, based on an estimated population of 680 students. Additionally, 10 instructors (6 male, 4 female) were purposively selected based on their experience with research supervision and their involvement in research writing instruction.

Table 1
Demographic Characteristics of Student Participants (N = 256)

Variable	Category	Frequency (n)	Percentage (%)
Gender	Female	147	57.4
	Male	109	42.6
Age	18–20	68	26.6
	21–23	142	55.5
	24–26	46	17.9
Academic Level	Year 2	82	32.0
	Year 3	89	34.8
	Year 4	85	33.2
AI Tool Experience	None	32	12.5
	Basic	98	38.3
	Intermediate	87	34.0
	Advanced	39	15.2
Primary AI Tool	ChatGPT	118	46.1
	Grammarly	62	24.2
	QuillBot	41	16.0
	Other	35	13.7

3.3 Data Collection Instruments

3.3.1 Questionnaire

A structured questionnaire was developed based on validated instruments from prior research on AI in academic writing. The instrument comprised five sections: (a) demographic information, (b) AI tool usage patterns (12 items, α = .847), (c) perceived opportunities (15 items, α = .891), (d) perceived challenges (18 items, α = .876), and (e) impacts on research writing (14 items, α = .903). All Likert-scale items used a 5-point response format (1 = Strongly Disagree to 5 = Strongly Agree). Content validity was established through expert review by three faculty members in applied linguistics and educational technology. A pilot study with 30 students not included in the final sample yielded a Cronbach's alpha of .912 for the full instrument, indicating high internal consistency.

3.3.2 Semi-Structured Interviews

Semi-structured interviews were conducted with 10 instructors. Each interview lasted approximately 45–60 minutes and was audio-recorded with participant consent. The interview protocol explored instructors' experiences with student AI use, perceived opportunities and challenges, ethical concerns, and recommendations for integration. The protocol was piloted with two instructors and refined accordingly. Interviews were conducted in English and Amharic, with transcriptions translated into English for analysis.

The researcher also used semi-structured interviews to assess the perception of students on the ProWritingAid tool implementation as a feedback provision tool. The researcher developed five semi-structured interview items and conducted an interview with five randomly selected students using a tape recorder after the treatment was over, for the reason that they would respond to their perception well when they had adequate skills, experience, and knowledge about the software. The reason that the researcher employed a simple random sampling method was to give a non-zero chance to all students in the class, regardless of their gender and academic performance. Other reasons that only five students were chosen for the interview were time and content management. It would be important if the researcher investigated students' perception of ProWritingAid implementation entirely. But it is difficult for the researcher to manage the time and content during data analysis. The amount of time that was devoted during the interview varied because the researcher was clarifying questions when the interviewees didn't grasp some concepts. As a result, the interviews took from 15–25 minutes.

3.3.3 Document Analysis

Document analysis examined 48 research documents, including 20 undergraduate research proposals, 18 senior essays, and 10 MA theses submitted to the department between 2023 and 2025. Documents were analyzed for indicators of AI use, including textual patterns, citation anomalies, and stylistic uniformity. A coding rubric was developed based on Turnitin similarity reports and manual inspection protocols.

3.4 Data Analysis Procedures

3.4.1 Quantitative Analysis

Quantitative data were analyzed using IBM SPSS Statistics version 28. The analysis proceeded through several stages:

Descriptive Statistics

Frequencies, percentages, means, and standard deviations were computed for all variables.

Reliability Analysis

Cronbach's alpha coefficients were calculated for each subscale.

Assumption Testing

Prior to inferential testing, statistical assumptions were verified. Shapiro-Wilk tests assessed normality for all dependent variables. Levene's test examined homogeneity of variance. Box's M test evaluated homogeneity of covariance matrices for MANOVA. Multicollinearity was assessed through tolerance and VIF values.

Inferential Statistics

Independent samples t-tests compared AI use patterns by gender. One-way ANOVA examined differences across academic levels. One-way MANOVA tested the multivariate effect of academic level on multiple AI use outcomes (frequency of use, perceived opportunities, perceived challenges, and perceived impacts). ANCOVA controlled for prior AI experience and digital literacy while examining group differences. Multiple regression analysis (enter method) examined predictors of AI integration in research writing, with perceived usefulness, digital literacy, access to AI tools, institutional support, and ethical awareness as predictor variables and AI integration as the criterion.

Effect Sizes

Partial eta squared (η^2) and Cohen's d were computed to assess practical significance.

Before the researcher employed types of significance tests, he tested the assumptions. The first assumption is the level of measurement. In parametric tests, the dependent variable has to be continuous. In this context, the dependent variable was continuous (e.g., paragraph writing test scores), and it was measured before and after interventions. The independent variable is categorical. In our case, as the independent variable is about time with two levels: pre-test and post-test, the researcher implemented parametric tests, especially a paired sample t-test. The second assumption is the distribution of scores. To employ a parametric test (in this case, a paired sample t-test), there should be a normal distribution of scores. To determine whether there is a normal distribution of scores, it is possible to look at the histograms for pretest and post-test. As a result, there was a normal distribution of scores for pre-test and post-test. For example, the mean score for the pre-test was 89.29. The scores lay between 80 and 100. This indicates that 50% of the scores lay below and above the mean. Besides, the mean score for the post-test was 93.77. The scores lay between 81 and 100. This shows that 50% of

the scores lay below and above the mean. So, it is possible to say that there was a normal distribution of scores below and above the stated mean. Therefore, the second assumption to use a parametric test was met.

The third assumption is randomization. In fact, the current study employed a quasi-experimental design, specifically, a one-group pretest-posttest design through which participants were selected without intervening in the natural setting. So, random sampling is impossible in a quasi-experimental research design. And it is not often the case in real-life research. Hence, the researcher tended to employ a parametric test, specifically, a paired sample t-test, due to the above reasons.

Therefore, in order to examine variations in writing performance among participants, descriptive statistics such as mean and standard deviation were used. In response to the first research question, which asks, "How does the ProWritingAid tool affect students' writing performance?", a paired sample t-test was used to compute observations made before and after interventions. The pre-test and post-test were computed using a paired sample t-test to ascertain whether or not there was significant variation in the writing performance of the students due to the treatment given to them.

3.4.2 Qualitative Analysis

Interview transcripts and document analysis notes were analyzed using thematic analysis following Braun and Clarke's (2006) six-phase framework. Initial codes were generated inductively, then organized into themes through iterative comparison and refinement. Two coders independently coded 20% of the data, achieving an inter-coder reliability of .84 (Cohen's kappa). Discrepancies were resolved through discussion.

To address the second study question, "What types of error are identified by the ProWritingAid tool?" frequency and percentage were used as data analysis methods. The types and rates of errors which were identified by the ProWritingAid tool were counted and converted into frequency and percentage.

As a response to the third inquiry, "How do students perceive the use of ProWritingAid in their learning to write?" thematic analysis was employed. The researcher selected five students using a simple random sampling method and interviewed them using a tape recorder. The language that the researcher used to interview students was English since the students came from different linguistic backgrounds. The data obtained through the semi-structured interview were analyzed qualitatively. The interview, which was conducted via tape recorder, was transcribed using TurboScribe AI. This AI tool was used to transcribe students' responses precisely and in a timely. The interview was fundamentally used to identify students' perceptions of the ProWritingAid tool implementation.

3.4.3 Integration of Findings

Quantitative and qualitative findings were integrated through a joint display matrix, enabling identification of convergence, divergence, and expansion. This approach followed the recommendation

for mixed-methods integration where quantitative results inform the selection of qualitative themes and vice versa.

3.5 Ethical Considerations

Ethical clearance was obtained from the Jimma University Institutional Review Board. Informed consent was secured from all participants. Student participation was voluntary, and anonymity was guaranteed through the use of coded identifiers. Data were stored securely and accessible only to the research team. QuillBot's grammar checker was used to ensure grammatical accuracy in the manuscript preparation, though all substantive content and analysis were conducted independently.

4. Results

4.1 AI Tool Usage Patterns

4.1.1 Frequency of AI Use

Students reported varying frequencies of AI tool use in research writing. Table 2 presents the distribution of AI use frequency across academic levels.

Table 2
Frequency of AI Use in Research Writing by Academic Level

Frequency	Year 2 (n = 82)	Year 3 (n = 89)	Year 4 (n = 85)	Total (N = 256)
Never	18 (22.0%)	8 (9.0%)	6 (7.1%)	32 (12.5%)
Rarely	28 (34.1%)	22 (24.7%)	14 (16.5%)	64 (25.0%)
Sometimes	24 (29.3%)	31 (34.8%)	28 (32.9%)	83 (32.4%)
Often	10 (12.2%)	20 (22.5%)	26 (30.6%)	56 (21.9%)
Always	2 (2.4%)	8 (9.0%)	11 (12.9%)	21 (8.2%)

The data reveal a clear developmental pattern: as students progress through academic levels, their AI use frequency increases. While 22% of Year 2 students reported never using AI tools, only 7.1% of Year 4 students reported the same. Conversely, 30.6% of Year 4 students reported using AI tools often, compared to 12.2% of Year 2 students. A chi-square test of independence confirmed a significant association between academic level and AI use frequency, $\chi^2(8, N = 256) = 28.47$, $p < .001$, Cramer's $V = .236$.

4.1.2 Types of AI Tools Used

Table 3 presents the distribution of primary AI tools used by students.

Table 3
Primary AI Tools Used by Students

AI Tool	Year 2 (n = 82)	Year 3 (n = 89)	Year 4 (n = 85)	Total (N = 256)
ChatGPT	32 (39.0%)	41 (46.1%)	45 (52.9%)	118 (46.1%)
Grammarly	22 (26.8%)	21 (23.6%)	19 (22.4%)	62 (24.2%)
QuillBot	12 (14.6%)	15 (16.9%)	14 (16.5%)	41 (16.0%)
SciSpace	5 (6.1%)	6 (6.7%)	4 (4.7%)	15 (5.9%)
Other tools	11 (13.4%)	6 (6.7%)	3 (3.5%)	20 (7.8%)

ChatGPT emerged as the dominant AI tool across all academic levels, with usage increasing from 39.0% in Year 2 to 52.9% in Year 4. Grammarly and QuillBot were also widely used, reflecting students' emphasis on grammar correction and paraphrasing.

4.1.3 Purposes of AI Use

Students reported using AI tools for multiple purposes in research writing (Table 4). The most common purposes were grammar and language correction (M = 4.21, SD = 0.84), idea generation (M = 3.87, SD = 1.02), and paraphrasing (M = 3.76, SD = 0.95). Less common purposes included literature synthesis (M = 3.12, SD = 1.18) and citation formatting (M = 2.89, SD = 1.24).

Table 4
Purposes of AI Use in Research Writing (N = 256)

Purpose	Mean	SD	Rank
Grammar/language correction	4.21	0.84	1
Idea generation	3.87	1.02	2
Paraphrasing	3.76	0.95	3
Structuring arguments	3.54	1.07	4
Vocabulary enhancement	3.42	1.11	5
Literature review assistance	3.12	1.18	6
Citation formatting	2.89	1.24	7

4.2 Perceived Opportunities

Students rated the opportunities presented by AI in research writing on a 5-point scale. Table 5 presents the descriptive statistics for opportunity items.

Table 5
Perceived Opportunities of AI in Research Writing

Opportunity Item	Mean	SD	Agreement (%)
AI improves overall writing quality	4.12	0.79	84.0
AI reduces language barriers in academic writing	4.08	0.86	82.4
AI enhances research productivity	3.96	0.91	78.9
AI provides useful feedback on drafts	3.89	0.94	76.6
AI facilitates idea generation	3.87	0.97	74.2
AI helps organize research structure	3.72	1.03	69.9
AI increases confidence in writing	3.65	1.08	66.8
AI supports learning of academic conventions	3.58	1.12	63.7

The highest-rated opportunity was improvement in overall writing quality (M = 4.12, SD = 0.79), with 84% of students agreeing or strongly agreeing. The reduction of language barriers was also strongly endorsed (M = 4.08, SD = 0.86), reflecting the particular relevance of AI tools for EFL students who must navigate academic writing in a second language.

4.3 Perceived Challenges

Table 6 presents descriptive statistics for challenge items.

Table 6
Perceived Challenges of AI in Research Writing

Challenge Item	Mean	SD	Agreement (%)
AI may lead to overreliance	4.24	0.81	87.1
AI may reduce critical thinking	4.18	0.88	84.4
Ethical concerns about plagiarism	4.05	0.94	80.5
Lack of institutional guidelines on AI use	3.97	1.01	76.6
Inadequate access to reliable internet	3.89	1.08	72.7
Cost of premium AI tools	3.78	1.15	68.8
Insufficient training on AI use	3.65	1.19	63.3
AI-generated content may be inaccurate	3.54	1.22	59.8

The most strongly endorsed challenge was the risk of overreliance (M = 4.24, SD = 0.81), with 87.1% agreement. Concerns about reduced critical thinking (M = 4.18, SD = 0.88) and ethical issues regarding

plagiarism (M = 4.05, SD = 0.94) were also prominent. Institutional gaps, including lack of guidelines (M = 3.97, SD = 1.01) and inadequate training (M = 3.65, SD = 1.19), were significant concerns.

4.4 MANOVA Results

A one-way multivariate analysis of variance (MANOVA) was conducted to examine the effect of academic level (Year 2, Year 3, Year 4) on four dependent variables: AI use frequency, perceived opportunities, perceived challenges, and perceived impacts. Preliminary assumption testing revealed no serious violations. Box's M test was non-significant (p = .087), indicating homogeneity of covariance matrices.

Table 7
MANOVA Results for Academic Level Effects

Effect	Value	F	Hypothesis df	Error df	p	Partial η²
Wilks' Lambda	.742	14.82	6	502	< .001	.141

The MANOVA revealed a statistically significant multivariate effect of academic level on the combined dependent variables, Wilks' Λ = .742, F(6, 502) = 14.82, p < .001, partial η² = .141. This indicates that academic level explains approximately 14.1% of the variance in the combined AI-related outcomes.

Table 8
Univariate ANOVA Results

Dependent Variable	df	F	p	Partial η²
AI Use Frequency	2	24.67	< .001	.163
Perceived Opportunities	2	12.34	< .001	.089
Perceived Challenges	2	8.91	< .001	.066
Perceived Impacts	2	15.28	< .001	.108

Follow-up univariate ANOVAs revealed that academic level had significant effects on all four dependent variables. Post-hoc comparisons using Tukey's HSD indicated that Year 4 students reported significantly higher AI use frequency than Year 2 students (mean difference = 0.87, p < .001) and Year 3 students (mean difference = 0.42, p = .012). Year 4 students also reported higher perceived impacts (mean difference = 0.64, p < .001) compared to Year 2 students.

4.5 ANCOVA Results

An analysis of covariance (ANCOVA) was conducted to examine differences in perceived impacts while controlling for prior AI experience and digital literacy. The covariate, prior AI experience, was significantly related to perceived impacts, F(1, 252) = 18.94, p < .001, partial η² = .070. After controlling for the covariate, there was still a significant effect of academic level on perceived impacts, F(2, 252) = 9.87, p <

.001, partial η² = .073. This indicates that academic level has a unique effect on perceived impacts beyond the influence of prior experience.

4.6 Multiple Regression Analysis

A multiple regression analysis was conducted to examine predictors of AI integration in research writing. The model included five predictor variables: perceived usefulness, digital literacy, access to AI tools, institutional support, and ethical awareness.

Table 9
Multiple Regression Predicting AI Integration in Research Writing

Predictor	B	SE B	β	t	p
(Constant)	0.847	0.312	—	2.715	.007
Perceived Usefulness	0.412	0.058	.412	7.103	< .001
Digital Literacy	0.287	0.061	.287	4.705	< .001
Access to AI Tools	0.201	0.067	.201	3.000	.003
Institutional Support	0.098	0.071	.098	1.380	.169
Ethical Awareness	0.054	0.069	.054	0.783	.434
Note. $R^2 = .486$, Adjusted $R^2 = .476$, $F(5, 250) = 47.23$, $p < .001$.					

The regression model was statistically significant, $F(5, 250) = 47.23$, $p < .001$, and explained 48.6% of the variance in AI integration ($R^2 = .486$, Adjusted $R^2 = .476$). Three predictors were statistically significant: perceived usefulness ($\beta = .412$, $p < .001$), digital literacy ($\beta = .287$, $p < .001$), and access to AI tools ($\beta = .201$, $p = .003$). Institutional support and ethical awareness were not significant predictors in the model.

4.7 Qualitative Findings

Thematic analysis of instructor interviews and document analysis revealed four major themes:

Theme 1: Transformative Potential. Instructors consistently acknowledged AI's transformative potential for research writing. One instructor (I-03) noted: *"AI tools have democratized access to writing support. Students who previously struggled with basic grammar now produce more polished drafts."* Another instructor (I-07) observed: *"AI enables students to focus on higher-order concerns like argumentation and evidence rather than getting bogged down in surface-level errors."*

Theme 2: Dependency and Skill Erosion. Despite recognizing benefits, instructors expressed significant concerns about dependency. *"I can see the difference between students who use AI as a tool and those who use it as a crutch. The latter group struggles when asked to produce original work without AI assistance"* (I-05). Document analysis revealed that essays with high AI involvement tended to exhibit formulaic structures and limited original insight, supporting quantitative findings on reduced critical thinking.

Theme 3: Ethical Ambiguity. Instructors reported uncertainty about ethical boundaries. *"We don't have clear guidelines. Students ask whether using ChatGPT for brainstorming is cheating, and honestly, I'm not sure what to tell them"* (I-02). Document analysis identified 12 cases (25% of analyzed documents) with textual patterns suggesting substantial AI involvement without proper disclosure.

Theme 4: Institutional Gaps. Instructors identified significant institutional gaps. *"There is no department-level policy on AI use. We need clear guidelines, training for both faculty and students, and infrastructure to support ethical AI integration"* (I-09). The lack of integration of AI literacy into the curriculum was a recurring concern.

5. Discussion

5.1 Interpretation of Findings

The findings reveal a complex picture of AI integration in research writing at Jimma University's English Language and Literature department. The high prevalence of AI use (87.5% reporting at least occasional use) aligns with recent research indicating that over 85% of Ethiopian students utilize AI tools for academic purposes. However, this widespread adoption coexists with significant concerns about dependency, ethical ambiguity, and institutional gaps.

The MANOVA results demonstrate that academic level significantly influences AI use patterns, with Year 4 students reporting higher use frequency and more positive perceptions of AI impacts. This developmental pattern likely reflects both increased research demands in later years and greater familiarity with AI tools over time. The finding that academic level explains 14.1% of variance in combined AI outcomes suggests that progression through the academic pipeline is a meaningful predictor of AI engagement.

The regression analysis identifies perceived usefulness as the strongest predictor of AI integration ($\beta = .412$), consistent with the Technology Acceptance Model's emphasis on perceived usefulness as a primary determinant of technology adoption. Digital literacy ($\beta = .287$) and access to AI tools ($\beta = .201$) also emerged as significant predictors, highlighting the importance of both individual competencies and infrastructural conditions. The non-significance of institutional support as a predictor is notable, suggesting that current institutional efforts may not be sufficiently visible or impactful to shape student behavior.

The qualitative findings enrich these quantitative results by revealing the mechanisms through which AI impacts research writing. The theme of dependency and skill erosion resonates with concerns expressed in the literature about diminished metacognitive engagement when AI use is unscaffolded. The ethical ambiguity theme aligns with research on "AI-giarism" and changing understandings of academic misconduct.

5.2 Comparison with Prior Research

The opportunities identified in this study align with findings from prior research in African higher education. Modiba et al. (2025) noted that generative AI tools can enhance student writing skills and increase productivity, leading to research autonomy and improved writing proficiency. Similarly, Ethiopian-specific research has documented significant cognitive benefits of GenAI, with 93% of students reporting that AI tools facilitate idea generation.

In the present study, data collected via test, screenshot, and semi-structured interview were organized and discussed vis-à-vis a review of related literature broadly as follows. According to descriptive statistical results, the mean score of students' pre-test was found to be 89.29. But after they practiced revising their paragraphs via the ProWritingAid tool for twelve weeks, students scored 93.77 in the post-test. The pre-test and post-test mean difference was 4.48. This indicates that students' writing performance was affected positively after they received feedback from the ProWritingAid tool. Results of the paired sample t-test also portrayed that students' writing performance was significantly improved after students practiced revising their paragraph via the ProWritingAid tool, $t(34) = -4.522$; $p < 0.05$. This finding is also aligned with Wahyuda et al (2022) results. The authors investigated the effectiveness of ProWriting application on senior high school students' writing skills. The output of the paired t-test revealed that students' writing skills were significantly ($P < 0.05$) improved as a result of implementing the ProWriting application as a language editor.

Results of the semi-structured interview showed that it is preferred to use the ProWritingAid tool at first and the final draft. Some students dared to use the ProWritingAid tool at the first draft, because students could understand the problems they have from the very beginning. Other researchers like Chen and Cheng (2008) also viewed that getting automated feedback at the first draft builds students' self-confidence. Some teachers in their study also reported that they used the results obtained from AWE as a requirement before they let their students submit it to their teacher and peers for further revision. They pointed out that students should pass through the system if they needed to be assessed by their teacher and peers. Students were expected to score a minimum of 4 out of 6 before they looked for teachers' and students' comments. This shows that the grade given by the tool is taken as a requirement for resubmission of students' written work to the teacher and peers for extra comments. But other students wish the tool to be used on the final draft to fill the gap found in teachers' comments. Therefore, the students recommended that ProWritingAid should be employed on the first and final draft. This indicates that students' interests should be taken into account during the feedback provision process.

Similarly, the results of the screenshot showed that students made severe mechanical faults while they were composing their paragraph. Grammar was the second language skill that was also wrongly used in the paragraph next to the mechanical aspects. Other language skills that the students used poorly in the paragraph were writing style and organization. In support of this idea, Kumar (2022), and Rich (2023) found that the ProWritingAid tool checks a piece of writing for wordiness, mechanical faults, grammatical faults, and legibility. They explained that users who subscribe to it can use it as a writing tutor and to edit the grammar and structure of phrases. But nothing was reported about the relevance of the content in the ProWritingAid tool. The reason for the absence of errors connected to content relevance in the ProWritingAid editor is that the tool is unable to detect the relevance of content in a paragraph. Therefore, it is also important to employ the ProWritingAid premium version to tackle the bottlenecks that are related to content relevance.

However, the challenges identified in this study—particularly overreliance, reduced critical thinking, and ethical concerns—are consistent with global patterns. Research on AI-giarism has highlighted changing perceptions of academic misconduct in the AI era. The finding that 25% of analyzed documents showed signs of undisclosed AI use raises serious concerns about academic integrity that require institutional attention.

The "adoption paradox" identified by Goshu et al. (2026) is evident in this study's findings. While students recognize AI's potential and use it frequently, they face infrastructural constraints including unreliable internet access and the cost of premium tools. The 72.7% agreement that inadequate internet access is a challenge underscores the material conditions that shape AI use in Ethiopian higher education.

5.3 Practical Implications

For Institutional Policy

The findings underscore the urgent need for clear, contextually appropriate guidelines on AI use in research writing. The 76.6% agreement that lack of institutional guidelines is a challenge indicates significant policy vacuum. Jimma University should develop a comprehensive AI use policy that addresses acceptable use, disclosure requirements, and academic integrity standards.

For Pedagogical Practice

The finding that 63.3% of students felt insufficiently trained on AI use suggests the need for integrated AI literacy instruction. Rather than treating AI as a prohibited tool, pedagogical approaches should scaffold AI use to promote critical engagement rather than dependency. Instructors should design assignments that require students to evaluate, critique, and build upon AI-generated content.

For Faculty Development

Instructors expressed uncertainty about ethical boundaries and their role in guiding AI use. Professional development programs should equip faculty with the knowledge and skills to guide students in

responsible AI use, including how to design AI-resistant assessments when appropriate and how to integrate AI literacy into disciplinary curricula.

For Infrastructural Investment

The finding that access to AI tools predicts AI integration (β = .201) highlights the need for institutional investment in ensuring equitable access. This includes not only hardware and connectivity but also subscriptions to AI tools that support academic writing.

5.4 Limitations

Several limitations warrant acknowledgment. First, the study relied on self-reported data for AI use patterns, which may be subject to social desirability bias. Second, the cross-sectional design limits causal inference. Third, the study focused on a single department in one university, limiting generalizability. Fourth, document analysis could not definitively determine AI involvement in all cases. Future research should incorporate longitudinal designs, multiple institutions, and objective measures of AI use.

6. Conclusion and Recommendations

Overall results of the present study showed that students preferred to use the ProWritingAid as a supplementary tool to fill the gap that classroom teachers left. Besides, the ProWritingAid tool was mainly used to revise mechanical, grammatical, writing style, and organization errors, but it was less important to provide comments on content. But students' excessive usage of the ProWritingAid tool may negatively affect their creativity. Therefore, based on these conclusions, the following recommendations are made:

Findings of the previous and the present study showed that using the ProWritingAid tool to revise grammatical, mechanical, writing style, and organization faults that were created in paragraphs was indispensable. But classroom teachers should provide comments on content since the tool does not give any comments regarding the relevance of content.

In fact, in this study, the researcher selected one section of management students as the sample of the study. But to obtain comprehensive data regarding the value of the tool in improving students' writing performance, future researchers should include students from other departments and maximize sample size.

The current study also emphasized a single AI tool, which is the ProWritingAid tool. Therefore, future studies should look at how ProWritingAid and other AI tools jointly affect students' writing performance.

The scope of the present study is on the effects of the ProWritingAid tool's free version on students' writing performance, particularly mechanics, grammar, writing style, and organization due to economic

constraints. So, it is better to assess the effects of the ProWritingAid tool's premium version on students' writing skills to comprehensively understand the role of the ProWritingAid tool on students' writing skills.

The goal of the current study was merely to investigate how students' writing performance was affected by ProWriting feedback without considering other types of feedback. Therefore, future researchers should include teachers' feedback, teacher feedback + electronic feedback, and no feedback as interventions to make sure whether or not the improvement in students' writing performance emerges due to the feedback given by teachers or the tool plus teacher's feedback, or the tool only, or because of other external variables.

This study provides empirical evidence on the opportunities, challenges, and impacts of AI in research writing within the Department of English Language and Literature at Jimma University. The findings reveal that AI tools are widely adopted by students for grammar correction, idea generation, and paraphrasing, with 87.5% reporting at least occasional use. Perceived opportunities center on improved writing quality, reduced language barriers, and enhanced productivity. However, significant challenges include overreliance, diminished critical thinking, ethical concerns regarding plagiarism, and infrastructural constraints.

The study makes three principal contributions. First, it provides the first systematic mixed-methods investigation of AI in research writing in an Ethiopian university department, addressing a critical gap in the literature. Second, it demonstrates the utility of advanced statistical approaches—including MANOVA, ANCOVA, and multiple regression—in examining the complex relationships between AI use, demographic variables, and research writing outcomes. Third, it offers contextually grounded recommendations for policy, pedagogy, and infrastructure.

The integration of AI into research writing is not a question of whether but how. The challenge for Jimma University and similar institutions is to develop institutional frameworks that harness AI's potential while mitigating its risks. This requires clear policies, pedagogical innovation, faculty development, and infrastructural investment. By adopting a proactive and evidence-informed approach, Ethiopian higher education can position itself to benefit from AI's transformative potential while preserving the integrity and rigor of academic research writing.

Declarations

Ethical Approval

This study was conducted in strict accordance with the ethical principles of the Declaration of Helsinki (World Medical Association, 2024 revision). Ethical clearance was obtained from the Jimma University Institutional Review Board (IRB) on 15 March 2025 (Protocol ID: JU-IRB-2025-047). All procedures adhered to the Declaration of Helsinki and Ethiopia's national ethical guidelines for research involving human participants, including the acquisition of informed consent from all participants and the institutional standards of Jimma University for human subjects research involving students and

instructors in the Department of English Language and Literature. All procedures prioritized participant welfare, informed consent, scientific validity, and independent ethical review. Risks associated with participation were assessed as minimal.

Declaration Statements

Conflict of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

Funding: The authors received no specific funding for this research or for the publication of this manuscript. The study was conducted as part of the authors' academic obligations at Jimma University, with no external financial support.

Consent to Participate: Informed consent was obtained from all instructors who participated in the semi-structured interviews. For student participants, informed consent was obtained from all students who were of legal age (18 years and above), while parental or guardian consent was obtained for students below the age of 18, with student assent obtained separately. All participants were informed of the purpose, procedures, potential risks, and benefits of the study, and were assured of their right to withdraw at any time without penalty.

Consent to Publish: All authors consent to the publication of this manuscript in *Discovery Education*. No identifiable individual data are included in this manuscript. Aggregate results, coded identifiers, and pseudonyms are used throughout to ensure participant anonymity and confidentiality.

Article Type Clarification: The manuscript is submitted as a **Research Article**. This classification reflects the study's original empirical data on the opportunities, challenges, and impacts of artificial intelligence in research writing among students and instructors at Jimma University. The manuscript includes clearly defined research questions, a rigorous mixed-methods methodology, quantitative and qualitative results, discussion, and implications for language education and academic writing pedagogy.

Data Availability Statement

The data supporting the findings of this study are available from the corresponding author upon reasonable request. Due to privacy restrictions related to human participants in this educational research conducted at Jimma University, the data are attached separately via the designated repository link and are also available at the end of the document. All data have been anonymized to protect participant confidentiality.

Acknowledgments

The authors express their sincere gratitude to the 256 students and 10 instructors of the Department of English Language and Literature at Jimma University for their willing participation in this study. The authors also thank the Jimma University Institutional Review Board for ethical protocol approval and the

Department of English Language and Literature for facilitating access to research documents and participants. Special appreciation is extended to the research assistants who contributed to data collection and transcription.

References

1. Abas A, Unpublished MA. Thesis). Haramaya University, Haramaya, Ethiopia.
2. Adaba HW. Awareness and practice of chat generative pre-trained transformers: A new language model artificial intelligence in higher institutions: The case of Ambo University EFL instructors. ETHIOINQUIRY J Humanit Social Sci. 2024;3(2). https://doi.org/10.20372/m9tjbf62.
3. Adelia A, Irawati T. The effect of ProWritingAid as artificial intelligence (AI) on student's engagement. JEELL: J Engl Educ Linguistics Literature. 2025;12(2):111–20. https://doi.org/10.32682/jeell.v12.
4. Ahmed A, Seliem S. (2009). Electronic feedback in Egyptian EFL essay writing classes: Paper presented at CDELT Conference, Faculty of Education, in Shams University: Egypt.
5. Alanazi MA. (2013). Using electronic writing to promote students' writing. EWU Masters Thesis Collection: Eastern Washington University.
6. Ariyanto MSA. (2019). Enhancing the informatics engineering students' ability in writing a descriptive paragraph using Pro Writing Aid and teacher feedback. Undergraduate Thesis, Universitas Negeri Malang.
7. Ariyanto A, Mukminatien N, Tresnadewi S. Students' and teacher's perceptions towards the implementation of Prowritingaid and teacher feedback. J Pendidikan: Teori Penelitian Dan Pengembangan. 2019;4(10):1353. https://doi.org/10.17977/jptpp.v4i10.12843.
8. Barrot J. Generative artificial intelligence for automated essay scoring: Exploring teacher agency through an ecological perspective. Assess Writ. 2026;67. https://doi.org/10.1016/j.asw.2025.100990.
9. Chen E, Cheng E. Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. Lang Learn Technol. 2008;12(2):94–112.
10. Choi J. (2010). The impact of automated essay scoring (AES) for improving English language learner's essay writing. A Dissertation presented to the Faculty of the Curry School of Education University of Virginia.
11. Cotton DRE, Cotton PA, Shipway JR. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations Educ Teach Int. 2024;61(2):228–39.
12. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q. 1989;13:319–39.
13. Ebabu T. (2013). A study on writing: Student perception and performance. Unpublished Ph.D. Dissertation: Addis Ababa University.

14. Elsayed A. The right corrective feedback in the age of technology: From theory into practice. Am J Sci Educ Res. 2025;4(6):1–2.
15. Ezzahra F. Investigating the effect of automated writing evaluation on EFL learners' writing accuracy: The case of Master Students at Biskra University. Undergraduate Paper, Biskra University; 2022.
16. Fitria T. ProWriting Aid as AI-powered writing tools: The performance in checking grammar and spelling of students' writing. POLINGUA: Sci J Linguistics Literature Lang Educ. 2023;12(2):65–72.
17. Flanagan K. Writer's Digest Guide to Magazine Article Writing: A Practical Guide to Selling Your Pitches, Crafting Strong Articles, & Earning More Bylines. Penguin; 2018.
18. Gebretsadkan KM, Amogne D. Effects of ProWritingAid tool on students' writing performance: Mekelle University in focus. ITYOPIS: Northeast Afr J Social Sci Humanit. 2026;8(1). https://doi.org/10.71624/a0nmmt31.
19. Ghufron MA, Rosyida F. The role of grammarly in assessing English as a Foreign Language (EFL) writing. Lingua Cultura. 2018;12(4):395–403.
20. Goshu BS. Artificial intelligence in Ethiopian higher education: A bibliometric analysis of research trends, systemic adoption barriers, and policy-infrastructure disconnects. Zenodo. 2026. https://doi.org/10.5281/zenodo.21422736.
21. Guadu A, Dibekulu D, Menberu AW. Academic integrity in the age of AI: Challenges and strategies for African higher education. Discover Educ. 2025. https://doi.org/10.1007/s44217-025-01031-2.
22. Habtamu. (2011). ELT teachers' beliefs of writing skills and their classroom practices. (Unpublished MA Thesis): Addis Ababa University.
23. He M. (2025). The impact and challenges of AI on students' writing skills—Take ChatGPT for example. *SHS Web of Conferences* 220, 04004. https://doi.org/10.1051/shsconf/202522004004
24. Hockly N. Automated writing evaluation. ELT J. 2019;73(1):82–8. https://doi.org/10.1093/elt/ccyo44.
25. Jayavalan K, Razali A. (2018). Effectiveness of online grammar checker to improve secondary students' English narrative essay writing. Int Res J Educ Sci (IRJES), 2(1).
26. Kaptelinin V, Nardi B. Activity theory as a framework for human-technology interaction research. Mind Cult Activity. 2018;25(1):3–5.
27. Kasneci E, Sessler K, Küchemann S, Bannert M, Dementieva D, Fischer F, Kasneci G. ChatGPT for good? On opportunities and challenges of large language models for education. Learn Individual Differences. 2023;103:102274.
28. Kumar P. The Art of the Writing: Your Go-to Book to Learn Writing Skills. Orange Book's Publication; 2022.
29. Lantolf JP, Thorne SL. Sociocultural Theory and the Genesis of Second Language Development. Oxford: Oxford University Press; 2006.
30. Link S, Mehrzad M, Rahimi M. Impact of automated writing evaluation on teacher feedback, student revision, and writing improvement. Comput Assist Lang Learn. 2020;1–30.

https://doi.org/10.1080/09588221.2020.1743323.

31. Liu S, Yu G. (2022). L2 learners' engagement with automated feedback: An eye-tracking study. *Language Learning & Technology*, 26(2), 78–105. https://doi.org/10125/73480.
32. Lund BD, Wang T. Chatting about ChatGPT: How may AI and GPT impact academia and libraries. Mind Mach. 2023;30:681–94. https://doi.org/10.1108/LHTN-01-2023-0009.
33. Mabborang J, Pulintan M, Javier G. The use of AI tools in enhancing student learning. Psychol Education: Multidisciplinary J. 2025;37(4):313–7.
34. Mandefro F, Asnakech D, Alemayehu N. Investigation of university students' writing problems: Two public universities in South West Ethiopia in focus. Int J Sciences: Basic Appl Res (IJSBAR). 2018;42(4):83–95.
35. Meseret T. (2012). Instructors' and students perceptions and practices of task based writing in an EFL context. Doctoral Dissertation, Addis Ababa University, Addis Ababa, Ethiopia.
36. Modiba FS, Van den Berg A, Mago S. Opportunities and challenges of generative artificial intelligence supporting research in African classrooms. South Afr J High Educ. 2025;39(3):173–93. https://doi.org/10.20853/39-3-6272.
37. Mollick E. (2024). *Co-intelligence: Living and working with AI*. Portfolio.
38. Nasution A, Fatimah S. The use of Pro Writing Aid web in editing students' writing. J Engl Lang Teach. 2018;7(2):362–8.
39. Obura EA, Emoit PI. Artificial intelligence in academic writing and research skills in Kenyan universities: Opportunities and challenges. Afr Educ Rev. 2024;20(6):58–80.
40. Ouchaib A. A quasi-experimental study investigating the effect of written corrective feedback on students' writing accuracy. Int J Linguistics Translation Stud. 2024;5(3):125–35. https://doi.org/10.36892/ijlts.v5i3.495.
41. Ramadhani R. (2022). The use of technology acceptance model to understand students' preference of using ICT applications in English academic writing.
42. Razak NA, Jalil HA, Krauss SE, Ahmad NA. Successful implementation of information and communication technology integration in Malaysian public schools: An activity systems analysis approach. Stud Educational Evaluation. 2018;58:17–29.
43. Rich JR. The Remote Worker's Handbook: How to Work Effectively from Anywhere. Entrepreneur; 2023.
44. Rich S, Tang J. Automated writing evaluation in an EFL setting: Lessons from China. JALT Call J. 2017;13(2):117–46.
45. Romeo K, Hubbard P. Pervasive CALL learner training for improving listening proficiency. In: Levy M, Blin F, Siskin C, Takeuchi O, editors. World CALL: International perspectives on computer assisted language learning. New York, NY: Routledge; 2010. pp. 215–29.
46. Tamrat T, Melese M. Identifying and analyzing common English writing challenges among regular undergraduate students. Heliyon. 2024;1–13. https://doi.org/10.1016/j.heliyon.2024.e36876.

47. Taskiran A, Goksel N. (2022). Automated feedback and teacher feedback: Writing achievement in learning English as a foreign language at a distance. Turkish Online J Distance Educ, 23(2), Article 9.
48. Tegegne Bekalu. (2020). EFL students' and teachers' perceptions and practices of the writing tasks in grades 11 and 12: The case of Liyew Asres Zewudie General Secondary and Preparatory School, East Gojjam Zone, Ethiopia. *Journal of Education and Practice*, 11(25).
49. UNESCO. Guidance for generative AI in education and research. UNESCO Publishing; 2023.
50. Vygotsky LS. Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press; 1978.
51. Wahyuda M. (2022). The effectiveness of Grammarly and ProWritingAid application toward writing skill across students writing level of Man Batu students. MA thesis: Universitas Islam Negeri Maulana Malik Ibrahim Malang.
52. Wahyuda M, Putera A, Khuseini M. The effectiveness of ProWritingAid toward students' writing skills in senior high school. J Engl Teach Learn. 2022;1(1):25–33.
53. Wale BD, Kassahun YF. (2024). The transformative power of AI writing technologies: Enhancing EFL writing instruction through the integrative use of Writerly and Google Docs. *Human Behavior and Emerging Technologies*, 2024, Article 6642897. https://doi.org/10.1155/2024/6642897
54. Wihastyanan W, Latief A. The impact of electronic feedback on second year English college students' writing quality. Int J Stud Engl Lang Literature (IJSELL). 2017;5(12):56–67.
55. Wright WE. Foundations for Teaching English Language Learners: Research, Theory, Policy, and Practice. 2nd ed. Philadelphia, PA: Caslon Publishing; 2015.
56. Zawacki-Richter O, Marín VI, Bond M, Gouverneur F. Systematic review of research on artificial intelligence applications in higher education – where are the educators? Int J Educational Technol High Educ. 2019;16(1):1–27.
57. Adaba HW. Awareness and practice of chat generative pre-trained transformers: A new language model artificial intelligence in higher institutions: The case of Ambo University EFL instructors. ETHIOINQUIRY J Humanit Social Sci. 2024;3(2). https://doi.org/10.20372/m9tjbf62.
58. Cotton DRE, Cotton PA, Shipway JR. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations Educ Teach Int. 2024;61(2):228–39.
59. Gebretsadkan KM, Amogne D. Effects of ProWritingAid tool on students' writing performance: Mekelle University in focus. ITYOPIS: Northeast Afr J Social Sci Humanit. 2026;8(1). https://doi.org/10.71624/a0nmmt31.
60. Goshu BS. Artificial intelligence in Ethiopian higher education: A bibliometric analysis of research trends, systemic adoption barriers, and policy-infrastructure disconnects. Zenodo. 2026. https://doi.org/10.5281/zenodo.21422736.
61. Guadu A, Dibekulu D, Menberu AW. Academic integrity in the age of AI: Challenges and strategies for African higher education. Discover Educ. 2025. https://doi.org/10.1007/s44217-025-01031-2.

62. Kasneci E, Sessler K, Küchemann S, Bannert M, Dementieva D, Fischer F, Kasneci G. ChatGPT for good? On opportunities and challenges of large language models for education. Learn Individual Differences. 2023;103:102274.
63. Mangubat FM, Saeedi KH. A bibliometric analysis of research trends in the application of artificial intelligence by college students in research writing. Discover Educ. 2025;5., Article 67. https://doi.org/10.1007/s44217-025-01067-4.
64. Modiba FS, Van den Berg A, Mago S. Opportunities and challenges of generative artificial intelligence supporting research in African classrooms. South Afr J High Educ. 2025;39(3):173–93. https://doi.org/10.20853/39-3-6272.
65. Obura EA, Emoit PI. Artificial intelligence in academic writing and research skills in Kenyan universities: Opportunities and challenges. Afr Educ Rev. 2024;20(6):58–80.
66. Sharma RC. Addressing academic dishonesty in higher education: A systematic review of generative AI's impact. Open Praxis. 2025;17(2):251–69.
67. Tobbi S. Using QuillBot to enhance paraphrasing skills among Algerian EFL university students. Zenodo. 2025. https://doi.org/10.5281/zenodo.18016108.
68. UNESCO. Guidance for generative AI in education and research. UNESCO Publishing; 2023.
69. Wale BD, Kassahun YF. (2024). The transformative power of AI writing technologies: Enhancing EFL writing instruction through the integrative use of Writerly and Google Docs. *Human Behavior and Emerging Technologies*, 2024, Article 6642897. https://doi.org/10.1155/2024/6642897
70. Zawacki-Richter O, Marín VI, Bond M, Gouverneur F. Systematic review of research on artificial intelligence applications in higher education – where are the educators? Int J Educational Technol High Educ. 2019;16(1):1–27.
71. Zhang Y, Wang L. A systematic review of the impact of generative AI on postgraduate research: Opportunities, challenges, and ethical implications. Discover Artif Intell. 2025;5:238. https://doi.org/10.1007/s44163-025-00495-3.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AppendixA.docx