

The Effect of Generative AI Tools (Grammarly and ProWritingAid) on Students' Writing Skill

Rosi Anjarwati¹  Universitas PGRI Jombang, Indonesia¹

rosi.stkipjb@gmail.com¹

Daning Hentasmaka² Universitas PGRI Jombang, Indonesia²

d.hentasmaka@gmail.com²

 <https://doi.org/10.58194/eloquence.v4i3.3064>

Corresponding Author:  Rosi Anjarwati

Article History	ABSTRACT
Received 13-10-2025	Background: The integration of Generative Artificial Intelligence (GAI) into English language education has transformed writing instruction, yet empirical comparisons of its tools remain limited.
Accepted: 11-11-2025	Purpose: This study investigates the comparative effect of Grammarly and ProWritingAid on EFL students' academic writing skills.
Published: 20-12-2025	Method: A quasi-experimental design was employed with three intact groups (n=37): Grammarly (n=13), ProWritingAid (n=11), and electronic dictionary control (n=13). Writing pre-test and post-test data were analyzed using the non-parametric Kruskal-Wallis test due to violations of normality and homogeneity assumptions.
	Results and Discussion: A significant difference was found among groups ($\chi^2=10.544$, $p=0.005$). ProWritingAid yielded the highest mean rank (24.73), followed by Grammarly (21.81), and electronic dictionary (11.35). Post-hoc Mann-Whitney tests confirmed significant gains for both AI tools over the control ($p<0.05$), but no significant difference between Grammarly and ProWritingAid ($p=0.792$).
	Conclusions and Implications: Both AI tools effectively enhance writing skills, with ProWritingAid showing slight superiority in stylistic and structural feedback.
Keywords:	<i>Generative AI; Grammarly; Prowritingaid; Writing Skill; Quasi Experimental Design</i>
ABSTRAK	
	Latar Belakang: Integrasi Generative Kecerdasan Buatan (GAI) ke dalam pendidikan bahasa Inggris telah mengubah pengajaran menulis, tetapi perbandingan empiris terhadap alat-alatnya masih terbatas.
	Tujuan: Studi ini menyelidiki efek komparatif Grammarly dan ProWritingAid terhadap keterampilan menulis akademis siswa EFL.
	Metode: Desain kuasi-eksperimental digunakan dengan tiga kelompok utuh (n=37): Grammarly (n=13), ProWritingAid (n=11), dan kontrol kamus elektronik (n=13). Data pra-tes dan pasca-tes menulis dianalisis menggunakan uji Kruskal-Wallis non-parametrik karena terdapat pelanggaran asumsi normalitas dan homogenitas.
	Hasil dan Pembahasan: Perbedaan signifikan ditemukan antar kelompok ($\chi^2=10,544$, $p=0,005$). ProWritingAid menghasilkan peringkat rata-rata tertinggi (24,73), diikuti oleh Grammarly (21,81), dan kamus elektronik (11,35). Uji Mann-Whitney post-hoc mengonfirmasi peningkatan signifikan untuk kedua perangkat AI dibandingkan kontrol ($p<0,05$), tetapi tidak ada perbedaan signifikan antara Grammarly dan ProWritingAid ($p=0,792$).

Kesimpulan dan Implikasi: Kedua alat AI tersebut secara efektif meningkatkan keterampilan menulis, dengan ProWritingAid menunjukkan sedikit keunggulan dalam umpan balik gaya dan struktural.

Kata Kunci

AI Generatif; Grammarly; ProWritingAid; Keterampilan Menulis; Desain Kuasi Eksperimental



Copyright: © 2025 by the author(s).

This is open access article under the

[Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.](#)

INTRODUCTION

The rapid development of digital technology in recent years has impacted various areas of life, including education. Artificial Intelligence (AI) has given a new color to the learning process, especially at the higher education level, where students are required to apply High Order Thinking Skills (HOTS). A kind of AI technology known as "generative AI" may produce original material by learning from data and employing sophisticated algorithms and neural networks to produce text, graphics, and music that resemble those of a person.^{[1], [2]} One of the most significant goals of AI is to build automated systems that can understand their surroundings and perform tasks as people do. ^[3] Therefore, lecturers and students need to be able to utilize AI wisely.

A critical aspect of learning English is developing writing skills. Writing skills are a form of literacy skill that plays a vital role in both academic and professional life.^[4] However, in practice, writing skills often pose a challenge for many students.^[5] In many cases, students lack confidence in writing, which impacts their writing. This is where the role of self-efficacy or self-confidence becomes essential, because students who have high self-efficacy will be more confident in completing their writing tasks.^{[6], [7], [8]}

As technology develops, the use of Artificial Intelligence (AI) in the world of education is starting to become a serious concern. The technology of AI has brought various innovations, including AI offering various tools and platforms that can help students improve their writing skills, both in terms of writing quality and in increasing self-confidence or self-efficacy. One example of the application AI in learning to write is the use of application-based AI, which can provide automatic and personalised feedback, which can help students understand mistakes and how to correct them.^[9] Furthermore, teaching writing skills in the era of Artificial Intelligence (AI) has changed the traditional method pen-and- paper becoming a much more creative and dynamic pedagogical experience with the involvement of a wide variety of AI in the writing process.^[10] These programs employ advanced algorithms to detect typical problems in grammar, punctuation, and syntax and make recommendations for enhancing clarity and style. These technologies also include unique features, such as paraphrase and sentence completion, to improve efficacy.

Implementation of AI in writing classes offers a variety of interesting possibilities. One example of a writing-based tool AI that is often used is Grammarly, an application that uses AI technology to provide suggestions on grammar, spelling, and sentence structure.^{[11], [12]} The tool can also provide more specific feedback, such as analysing the strength of arguments, vocabulary usage, and consistency of writing style. This feedback can help students understand their writing's strengths and weaknesses, which in turn can increase their confidence in writing. Furthermore, Ahmadzade & Farahian ^[11] also found that Grammarly can be used as an additional tool in EFL classes, especially to help students become more accurate with grammar.

The use of AI can also improve students' writing skills, particularly in continuous feedback-based learning. AI technology can analyze students' writing and provide relevant suggestions tailored to their skill level. This allows students to learn in a more adaptive and tailored way, without feeling burdened by time constraints or the judgment of others. While AI can provide numerous

benefits in writing instruction, several challenges must be addressed. One is the disparity in access to technology that may exist in some regions or between different socioeconomic groups. This can lead to inequities in the use of AI in education. Furthermore, there are challenges in ensuring that the AI tools used provide accurate and relevant feedback and do not replace teachers' roles in delivering holistic, in-depth instruction.

Based on this background, this study is essential to understand the extent to which the use of Generative AI-based technologies, such as Grammarly and ProWritingAid, can influence students' writing skills. Most previous studies focused on the use of a single AI tool, but comparative evidence between Grammarly and ProWritingAid remains limited. While Wahyuda (2022) compared Grammarly and ProWritingAid, the study was limited to high school learners and did not assess argumentation quality or lexical sophistication. This study fills that gap by examining university-level EFL students' development across content, organization, and style, using a validated analytical rubric. This is where the research gap in this study lies—the need to examine the influence of generative AI tools on the cognitive and creative dimensions of writing skills. Thus, the novelty of this study lies in its approach, which assesses how the use of Grammarly and ProWritingAid not only corrects linguistic errors but also has the potential to enrich students' argumentation and writing style in the digital age. Particularly, this research aims to investigate the following questions: (1) Is there a difference in the writing skill of students who use Generative Artificial Intelligence (GAI): Grammarly, ProWritingAid and electronic dictionaries? And (2) Which among Grammarly, ProWritingAid, and electronic dictionaries has a higher influence on students' writing skills?

LITERATURE REVIEW

Generative AI in Language Education

A subfield of artificial intelligence known as "generative artificial intelligence" (GAI) focuses on systems that can exploit patterns discovered from vast amounts of training data to produce fresh, realistic, and unique material (such as text, photos, audio, code, or video).^[13] Unlike discriminative AI, which only classifies or predicts, GAI uses generative models to understand the underlying data structure, enabling the generation of new outputs with characteristics similar to the original data. These increased capabilities have fueled an AI "boom" since the 2020s, bringing popular tools such as ChatGPT, Grammarly, and ProWritingAid, which have transformed various industries, especially in the education field.

Intelligent Tutoring Systems (ITS) and adaptive learning platforms are examples of AI systems that may modify teaching strategies, pace, and materials to meet the individual needs and learning preferences of every student.^[14] This allows for the provision of real-time feedback that is more accurate and detailed, a key factor that, according to Azenoud ^[15], can directly improve student learning outcomes. In addition, AI automates administrative tasks and formative assessments, substantially reducing teachers' workload and allowing them to focus on deeper pedagogical interactions and effective curriculum development.^{[16], [17]} Among many AI that have been integrated into the teaching and learning process, Grammarly and ProWritingAid are commonly used. Grammarly and ProWritingAid have become an integral Artificial Intelligence (AI) based writing aid in writing classes, with a primary focus on improving mechanical accuracy and quality of writing style for students. Research consistently shows that using Grammarly can significantly improve students' writing skills, particularly in grammar, spelling, and punctuation, thanks to its ability to provide direct corrective feedback.^[12] However, some studies, such as Wahyuda's ^[18], give more nuance: Grammarly was found to be more effective for students with low writing skills because of its focus on quick, basic corrections, thereby increasing student confidence and motivation. Meanwhile, ProWritingAid is often seen as offering more comprehensive analysis and depth, with dozens of special reports covering aspects such as sentence

length variation, use of the passive voice, and word repetition. This ProWritingAid approach is more educational and is oriented towards developing writing skills systematically, making it more effective for students with high writing skill or creative writers who want to perfect their style.[18], [19] Both tools also help reduce teachers' workload for basic corrections, allowing them to focus more on higher-level content and writing structure.

AI in Writing Instruction

The integration of Artificial Intelligence (AI) in writing instruction has evolved from a simple evaluative function to an adaptive learning system that supports the entire writing process. AI now serves not only as a correction tool but also as a facilitator in the prewriting, drafting, and revision stages. Several studies have suggested that AI supports a process-based writing approach by providing continuous feedback and increasing student engagement in repetitive writing practices.[20], [21] Pedagogically, AI is positioned as a scaffolding tool that strengthens students' metacognitive processes in planning, monitoring, and evaluating their writing.[22], [23]

One of the most widely researched forms of AI implementation in writing instruction is Automated Writing Evaluation (AWE). AWE systems provide automated feedback on grammar, vocabulary, text organization, and sentence coherence. Several meta-analyses have found that AWE significantly improves students' linguistic accuracy and mechanical awareness of written language.[24] However, several studies highlight that AWE is less than optimal for assessing complex semantic and argumentative aspects, so the teacher's role remains essential in providing conceptual and reflective feedback.[25]

In the context of second language writing, the use of AI has been shown to positively impact vocabulary mastery, writing fluency, and grammatical accuracy.[26], [27] However, the literature also identifies the potential for over-reliance on AI, which can reduce students' cognitive engagement in the independent writing process [28]. Furthermore, ethical issues such as plagiarism, algorithmic bias, and data privacy pose serious challenges to the integration of AI in education.[29] Therefore, educators need to instill AI literacy and academic ethics to ensure critical and responsible use of technology.

Based on the results of the research synthesis, AI makes a significant contribution to improving the quality of writing learning when implemented strategically and pedagogically. AI should function as a supporting tool (instructional aid), not a substitute for the role of teachers, especially in aspects of critical thinking and writing creativity.[20], [23] Future research should focus on longitudinal studies to measure the long-term impact of AI on writing quality, as well as the development of AI-based pedagogical frameworks oriented towards reflective and humanistic learning.[21], [29]

Grammarly vs ProWritingAid: Functional and Pedagogical Differences

Artificial Intelligence like Grammarly and ProWritingAid are considered bridges between autocorrect and writing instruction; experimental research and systematic reviews suggest that AI can reduce mechanical errors and increase learner awareness of language aspects, but their pedagogical effects depend on how they are integrated into teaching practices.[27], [30] Furthermore, AI is effective as supplementary feedback but are not a substitute for targeted instruction.

Functionally, Grammarly stands out for its real-time grammar and spelling checks optimized for standard language, tone/clarity checks, and plagiarism checks easily accessible through browser extensions and Office/Google Docs integration; company documentation and user studies emphasize its ease of use and 'out-of-the-box' recommendation settings.[31] In contrast, ProWritingAid offers in-depth reports (20+ reports on readability, overused words, pacing, consistency, and chapter/manuscript analysis) and a rephrasing tool designed for stylistic editing

and substantive revision, making its functionality more geared toward intensive editing than quick corrections.[18], [32]

These feature differences have pedagogical consequences: Grammarly, with its concise feedback and concise explanation options, is better suited for daily self-correction exercises, quick formative assessments, and helping students correct surface errors; however, several studies note its limited contextual explanations and tendency to “recommend immediate improvements,” risking students accepting suggestions without in-depth reflection.[31], [33] In contrast, ProWritingAid—due to its analytical reporting—encourages reflective processes and awareness of genre/structure (e.g., pacing, overused words), thus better supporting teaching activities that emphasize layered revision and style development.[34]

Empirical testing and literature reviews have yielded mixed results: some studies report reduced grammatical errors and improved writing performance after using Grammarly or ProWritingAid, while others caution against inaccurate detection of contextual constructs and variations in English varieties (e.g., British vs. American) and the need for teacher assistance in interpreting feedback.[27], [31] Research also suggests that student uptake is high provided the tools are paired with clear learning tasks and instructional scaffolding.

Based on functional and pedagogical evidence, practical recommendations are: (1) use Grammarly for quick corrective exercises, mechanical checks, and as a preliminary layer before teacher assessment; (2) utilize ProWritingAid when the learning goal is in-depth revision, style development, or analysis of large texts (chapters/manuscripts); and (3) in both cases, instructional facilitators should teach students how to read and evaluate automated feedback (metacognitive scaffolding) so that feedback becomes a learning resource, not just a direct improvement tool.[27], [30]

Academic Writing and the Role of Automated Feedback

Academic writing is defined as a formal and structured writing style used in educational environments, with the primary purpose of presenting ideas, arguments, and research results clearly and convincingly so that they can be studied and developed by the scientific community.[35] This activity is not just a process of transferring information, but a manifestation of critical thinking and in-depth writing, in which writers are required to present ideas systematically and logically, supported by valid empirical or theoretical evidence. Academic writing encompasses various types of written work, such as essays, research papers, proposals, theses, and dissertations, all of which are oriented towards advancing knowledge in a particular field.

In this era, the study of writing underwent a significant shift, driven primarily by technological advances and the need for digital literacy. Recent research highlights the central role of Artificial Intelligence (AI) as an aid and the ethical challenges it poses in both academic and non-academic writing. For example, studies have shown that AI-based feedback systems (such as Grammarly or ChatGPT) significantly improve students' grammatical accuracy, vocabulary, and coherence, offering instant, personalized feedback.[6], [15], [36], [37] However, there have been worries expressed about the possibility of over-reliance, which can impede the growth of creativity and critical thinking. In order to guarantee the moral and responsible application of AI and uphold academic integrity, the study also highlights the significance of digital literacy and explicit regulations from educational establishments.

English writing skills are one of the most essential productive skills in learning a foreign language. According to Chauhan [35], writing is not simply about putting words on the page, but also involves complex thought processes such as planning, organizing ideas, editing, and revising. This skill requires mastery of linguistic elements such as grammar, vocabulary, punctuation, and correct spelling. Furthermore, writing serves as a means of expressing ideas and demonstrating

critical thinking skills in the target language. Factors influencing English writing skills are diverse, ranging from linguistic factors to psychological ones. Hyland [38] emphasized that motivation, self-confidence, and learning strategies significantly impact the quality of students' writing. Writing difficulties often arise from limited vocabulary, poor skills to organize ideas, and a lack of exposure to good text models. Therefore, effective writing instruction should provide explicit support for the writer's thinking process, rather than just assessing the final product.

Recent research has consistently highlighted the various obstacles EFL learners face in developing their writing skills. The main challenge often lies in limited mastery of grammar and vocabulary.[39], where structural differences between the mother tongue and English make it difficult to express ideas in writing. Furthermore, psychological factors such as writing anxiety and low self-efficacy have also been shown to hinder student performance significantly.[40] Lack of motivation and limited exposure to English outside the classroom exacerbate these problems.[41] These findings emphasize that developing writing skills requires holistic interventions that address not only linguistic aspects but also emotional support and effective pedagogy.

In an effort to overcome these barriers, recent studies have focused on the effectiveness of pedagogical innovation and technology integration. The use of learning strategies such as self-editing and self-correction based on analysis of grammatical and vocabulary errors has been reported to improve students' writing skills significantly.[42] Furthermore, the emergence of advanced technology has become a significant research trend, with the integration of Artificial Intelligence (AI) tools such as ChatGPT into writing instruction shown to provide instant feedback, reduce anxiety, and foster student creativity.[43] Structured peer feedback, particularly that supported by AI tools, is also an effective strategy for improving grammatical accuracy and vocabulary development through collaborative learning.[44] These trends indicate a shift toward hybrid writing processes that leverage technology to improve the quality and efficiency of learning.

Research Gap and Theoretical Positioning

Table 1. Synthesis of previous studies

Study	Focus	Finding	Relevance
Ahmadzade & Farahian (2025)	Grammarly for EFL	Improving grammatical accuracy	Support the use of Grammarly
Wahyuda (2022)	Comparing Grammarly vs. ProWritingAid	ProWritingAid is more effective for advanced writers	Justifying the selection of both tools
Azennoud (2024)	Real-time AI feedback	Improve learning outcomes	Support intervention approaches

These studies collectively confirm the efficacy of AI writing tools but reveal a need for controlled comparative research in EFL academic contexts—precisely the gap this study addresses

METHOD

The research design used in this study is a quasi-experimental, namely an experimental type, where the researcher only has partial control (or no control) over the participants of the manipulated variable.[45] A quasi-experimental design is a standard and practical research method in an educational setting, particularly when a true experiment with random assignment of students or classes is not feasible or ethical. This design aims to establish a cause-and-effect relationship for an intervention, such as a new curriculum or teaching method. Still, it uses pre-existing groups (such as intact classrooms or schools) rather than randomly created ones. A quasi-experiment is considered the most appropriate research design for this study because the researcher cannot conduct random assignment, unlike in a True experiment. The subjects involved in the research are already formed into classes and have a schedule that cannot be changed.

This study involved active students in the English Language Education study program at PGRI Jombang University in the 2024-2025 academic year, consisting of eight classes, as the population. The population is all research objects or target groups, including researchers, symptoms, values, and events that serve as data sources for researchers, or the entire subject group.^[46] In this study, sampling was used as a purposive technique, namely a technique with specific considerations. Three groups were selected as samples: 2024 A as experimental group 1, 2023 A as experimental group 2, and 2024 B as the control group. The selection of these classes was based on the consideration that students were taking a writing course (writing).

The instrument used in this study is a writing test consisting of a pre-test and post-test on writing Hortatory Exposition texts, timed to 90 minutes. All groups used the same writing prompt. The writing was scored using Brown's (2007) analytic rubric (content, organization, vocabulary, grammar, mechanics) on a 5-point scale. Inter-rater reliability (Cohen's κ) was 0.82. Further, an expert validator was involved to validate the content of the research instruments.

During data collection, each student in each group was given a 90-minute pre-test on writing hortatory exposition texts during the first meeting. In the second, third, and fourth meetings, students in experimental groups received a 30-minute orientation on the respective AI tool. Over three weekly sessions, they revised hortatory exposition drafts using the tool for 60 minutes per session. The experimental group 1 used the AI-based application Grammarly for their writing, while the experimental group 2 used the AI-based application ProWritingAid. The control group used an electronic dictionary (Cambridge Online Dictionary) under identical conditions. In meeting 5, Students in each group then completed a 90-minute post-test, writing a hortatory exposition text. Both the pre-test and post-test writings were assessed by three raters, who were English Language Education lecturers who had received prior training. To ensure originality, the writing was checked using Turnitin. Before it, the students were informed about ethical use of AI and originality control (Turnitin<20%).

After the data from pre-test and post-test were obtained, the data were analyzed using the non-parametric Kruskal-Wallis. This is due to the assumptions of normality of the data distribution and homogeneity of group variances not being met, making it impossible to carry out the parametric ANOVA. The Kruskal-Wallis test is a nonparametric test used to compare median or mean ranks across three or more independent groups. This test is an alternative to one-way ANOVA when the data do not meet parametric assumptions, such as normality and homogeneity of variance. [47]. The post-hoc test used Mann-Whitney U test with Bonferroni adjusted $\alpha = 0.017$. Further, the effect size was calculated by using Cohen's d (1988) with the following classification:

Table 2. Cohen's effect size classification

Effect Size (ES)	Interpretation
$0,00 \leq ES < 0,20$	Ignored
$0,20 \leq ES < 0,50$	Small
$0,50 \leq ES < 0,80$	Moderate
$0,80 \leq ES < 1,30$	Large
$1,30 \leq ES$	Very Large

RESULT AND DISCUSSION

This section presents the results of data analysis obtained from research on the influence of using Generative AI Grammarly and ProWritingAid on students' writing skills. The research results are presented based on quantitative findings showing changes in writing skills before and after using the two tools. The discussion focuses on the interpretation of the results, their relationship to theory and previous research, and their implications for improving students' academic writing skills in the digital era.

Writing skills of students using Artificial Intelligence (AI) Grammarly, ProWritingAid, and electronic dictionaries.

The data analysis to answer the first research question (is there a difference in the writing skills of students using Artificial Intelligence (AI) Grammarly, ProWritingAid, and electronic dictionaries?) was conducted using the Kruskal-Wallis non-parametric statistical test. This was due to the assumptions of normality of data distribution and homogeneity of group variances not being met, making it impossible to conduct a parametric ANOVA. Prior to the non-parametric test, a descriptive statistics test was implemented on the data set to calculate the median, quartiles, and Interquartile Range (IQR) with the following results.

Table 3. Descriptives

		Percentiles				
Writing Skill	Application	Median	25	50	75	IQR
		23.00	18.50	23.00	24.25	5.75
	ProWritingAid	23.50	22.00	23.50	23.50	1.50
	Electronic Dictionary	18.50	18.00	18.50	19.50	1.50

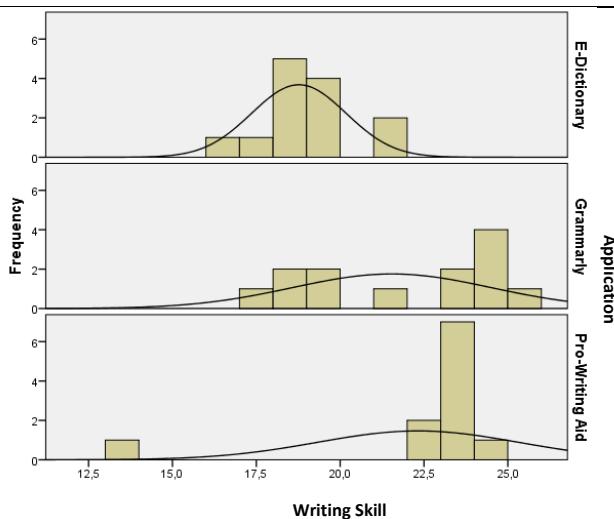
The descriptive statistics presented in Table 3 provide an initial overview of students' writing performance across the three applications (Grammarly, ProWritingAid, Electronic Dictionary) before conducting the non-parametric analysis. The median scores indicate that ProWritingAid yielded the highest central tendency (Median = 23.50), followed closely by Grammarly (Median = 23.00), while the Electronic Dictionary produced the lowest median score (Median = 18.50).

An examination of the IQR shows notable differences in score variability across the applications. Grammarly has the widest IQR (5.75), suggesting greater dispersion and more variability in students' writing scores when using this tool. In contrast, both ProWritingAid and the Electronic Dictionary show much smaller IQR values (1.50), indicating that students' scores were more consistent when using these applications.

The quartile values further support this pattern. For Grammarly, the gap between the 25th percentile (18.50) and the 75th percentile (24.25) is relatively large, again reflecting variability in students' performance. Meanwhile, ProWritingAid displays identical median and 75th percentile values (23.50), suggesting that many participants scored near or at this level. The Electronic Dictionary also shows a narrow range between Q1 (18.00) and Q3 (19.50), indicating that students' performance with this tool was clustered closely around the median.

Overall, the descriptive results suggest that while ProWritingAid produced the highest and most consistent writing scores, Grammarly generated a wider spread of outcomes, and the Electronic Dictionary resulted in the lowest, but more uniform, performance among the students.

Following the descriptive analysis, the non-parametric Kruskal-Wallis test was conducted to compare the mean ranks across three independent groups in this study. The mean ranks comparison, instead of medians, was done since the data distribution in each group has a different shape as presented in Graphics 1.

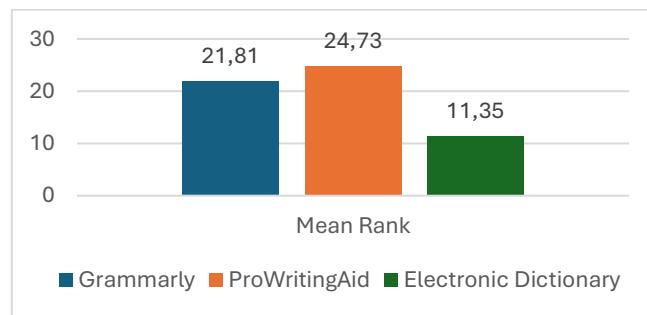
**Graphics 1.** Data Distribution Shapes

The former result of Kruskal-Wallis's test, Table 4, presents the mean ranks for students' writing skill performance across the three applications: Grammarly, ProWritingAid, and the Electronic Dictionary.

Table 4. Rank

	Application	N	Mean Rank
Writing skill	Grammarly	13	21.81
	ProWritingAid	11	24.73
	Electronic Dictionary	13	11.35
	Total	37	

ProWritingAid obtained the highest mean rank (24.73), indicating that students who used this tool generally scored better in writing compared to the other two groups. Grammarly follows with a mean rank of 21.81, showing moderately high performance but still below ProWritingAid. In contrast, the Electronic Dictionary shows the lowest mean rank (11.35), suggesting that students using this tool performed considerably lower than those using the two AI-based writing applications. Graphics 2 gives a clear visualization of the comparison.

**Graphics 2.** Mean Rank Chart

The clear separation between the mean ranks, particularly the gap between ProWritingAid/Grammarly and the Electronic Dictionary, suggests substantial differences in

writing outcomes across the tools. These rank values imply that the AI-based writing applications (especially ProWritingAid) tended to support higher writing performance than the non-AI tool.

Overall, the ranking results indicate that ProWritingAid was associated with the strongest writing performance, followed by Grammarly, while the Electronic Dictionary yielded the weakest performance among the three groups.

Table 5. Test Statistics

	Writing skill
Chi-Square	10.544
df	2
Asymp. Sig.	.005
a. Kruskal Wallis Test	
b. Grouping Variable: Application	

Table 5 presents the results of the Kruskal-Wallis test, which was used to determine whether there are statistically significant differences in writing skill scores among the three groups using different tools. The test produced a Chi-Square value of 10.544 with 2 degrees of freedom (df). Further, the associated Asymp. Sig. value is .005, which is below the alpha value of .05. This indicates that there is a statistical significant difference in writing skill among the three groups. In other words, at least one of the writing tools led to significantly different writing outcomes compared with the others. Given the earlier rank results, this significance suggests that the higher mean ranks for ProWritingAid and Grammarly, and the notably lower mean rank for the Electronic Dictionary, represent real differences rather than random variation. Overall, the Kruskal-Wallis test confirms that the type of AI-based writing applications had a significant effect on students' writing skill scores.

These results align with the findings of Ahmadzade & Farahian [11], who stated that Grammarly can provide automated, personalized feedback, particularly on grammar and spelling, which directly improves writing accuracy. In this context, students who use Grammarly gain a grammar-based learning experience and intelligent feedback, as explained by Azenoud [15], who emphasized the importance of real-time feedback in strengthening student learning outcomes. ProWritingAid, on the other hand, has been shown to provide a more in-depth and educational analysis of writing, covering aspects of style, clarity, and cohesion, making it superior to Grammarly in terms of improving the quality of ideas and writing structure. [38], [48].

From a theoretical perspective, these results reinforce the concept of High Order Thinking Skills (HOTS) explained in the introduction. The use of AI tools like Grammarly and ProWritingAid not only helps students correct mechanical errors but also stimulates higher-order thinking skills, such as analysis, evaluation, and synthesis of ideas. This aligns with Chauhan's [35] view that writing is a complex cognitive activity, encompassing planning, organization, and revision. Students who utilize ProWritingAid perform better because this tool provides not only automatic correction but also contextual analysis that encourages in-depth reflection and revision of the text.

Furthermore, these findings strengthen the argument that AI can act as an effective pedagogical partner in writing learning. Consistent with Intelligent Tutoring System (ITS) theory, AI-based tools can tailor feedback to individual skill levels, enabling adaptive, personalized learning. While electronic dictionaries still enrich vocabulary, they do not provide the reflective or corrective feedback needed to develop holistic academic writing skills. This explains why the mean rank of the electronic dictionary group was significantly lower than that of the other two AI tools.

The effect of Grammarly, ProWritingAid, and electronic dictionaries on students' writing skill.

The results of the Kruskal-Wallis data analysis presented in Tables 1 and 2 show significant

differences in student writing skills across the three observation groups using three different applications, with a mean rank of 21.81 for the group using Grammarly, 24.73 for the group using ProWritingAid, and 11.35 for the group using the Electronic Dictionary.

To identify these significant differences, further analysis was needed by comparing the mean ranks of students' writing skills in each group using the Mann-Whitney U-Test, with the following results:

Table 6. Test Statistics Grammarly vs ProWritingAid

	Writing skill Grammarly vs ProWritingAid
Mann-Whitney U	67.000
Wilcoxon W	158.000
Z	-.263
Asymp. Sig. (2-tailed)	.792
Exact Sig. [2*(1-tailed Sig.)]	.820 ^b

a. Grouping Variable: Application
b. Not Corrected for Ties

The first follow-up analysis aimed to compare the mean rank of students' writing ability in the groups using the AI-based applications Grammarly and ProWritingAid. The Mann Whitney U-Test output in Table 3 shows an Asymp. Sig. value greater than the Alpha value ($.792 > .05$). These results indicate no significant difference in the mean rank of the two observation groups, meaning there is no significant difference in the writing ability of students using the AI-based applications Grammarly or ProWritingAid.

The second analysis further compared the mean ranks of the groups using the AI-based application Grammarly and the Electronic Dictionary.

Table 7. Test Statistics Grammarly vs Electronic Dictionary

	Writing skill Grammarly vs Electronic Dictionary
Mann-Whitney U	43.500
Wilcoxon W	134.500
Z	-2.113
Asymp. Sig. (2-tailed)	.035
Exact Sig. [2*(1-tailed Sig.)]	.034 ^b

a. Grouping Variable: Application
b. Not corrected for ties.

The data analysis results indicated a significant difference in writing ability between students using Grammarly and those using the Electronic Dictionary (Asymp. Sig. = $.035 < .05$), with students using Grammarly having higher mean writing ability ranks than those using the Electronic Dictionary (see Table 1).

Similar results were also found in the third follow-up analysis, which compared the mean ranks of the Grammarly and Electronic Dictionary groups, as shown in Table 5 below.

Table 8. Test Statistics ProWritingAid vs Electronic Dictionary

	Writing skill ProWritingAid vs Electronic Dictionary
Mann-Whitney U	13.000
Wilcoxon W	104.000
Z	-3.424

Asymp. Sig. (2-tailed)	.001
Exact Sig. [2*(1-tailed Sig.)]	.000 ^b
a. Grouping Variable: Application	
b. Not corrected for ties.	

The Asymp. Sig. value shows a significant difference in the writing ability of students using ProWritingAid and the Electronic Dictionary ($.001 < .05$). Referring to the mean rank in Table 1, it can be seen that the mean rank of the group using ProWritingAid is higher than the group using the Electronic Dictionary.

More specifically, ProWritingAid produced the highest mean rank (24.73), followed by Grammarly (21.81), and Electronic Dictionary (11.35). This indicates that although both are AI-based, ProWritingAid has a slightly greater impact on improving students' writing skills. This finding supports the theory put forward by Wahyuda [48] and other studies [18], [19] which state that ProWritingAid is more educational and comprehensive, because it not only emphasizes correcting linguistic errors but also provides in-depth analysis of language style, sentence variation, and paragraph coherence. Thus, students not only correct mechanical errors but also understand the characteristics of academically effective writing.

In contrast, the group using the electronic dictionary demonstrated the lowest writing ability. This can be explained because the electronic dictionary only functions as a translation or vocabulary search tool, without providing contextual or corrective feedback. This finding is in line with Hyland's theory [38] which emphasizes that vocabulary mastery alone is not enough to produce good writing without the support of writing learning strategies that include planning, revision, and editing. Therefore, the role of AI technology in providing direct feedback has proven more effective in developing students' critical and reflective thinking when writing than the use of static electronic dictionaries.

Overall, the results of this study support the view that the use of AI-based tools such as Grammarly and ProWritingAid not only improves the mechanical aspects of writing but also strengthens the cognitive and affective dimensions of the academic writing process. In line with the findings of previous studies [9], [10], [12], the integration of AI technology in writing learning allows students to have a more interactive, personalized, and adaptive learning experience to their needs. In addition to improving linguistic accuracy, using Grammarly and ProWritingAid can increase students' self-efficacy in writing, as they feel more confident and able to control the quality of their writing independently. Thus, the practical implications of this study emphasize the importance of integrating AI technology into writing instruction strategies in higher education.

CONCLUSION AND IMPLICATIONS

Based on the results of the Kruskal Wallis test analysis, it can be concluded that there are significant differences in the writing skills of students who use three types of applications, namely Grammarly, ProWritingAid, and the Electronic Dictionary (Asymp. Sig. = $.005 < .05$). The mean rank value shows that the group using AI-based applications—specifically ProWritingAid (24.73) and Grammarly (21.81)—obtained higher writing ability scores than the group using the Electronic Dictionary (11.35). Further analysis using the Mann-Whitney U test also confirmed significant differences between the Grammarly and Electronic Dictionary groups, and between the ProWritingAid and Electronic Dictionary groups, but no significant difference between Grammarly and ProWritingAid. This indicates that both Grammarly and ProWritingAid significantly improve EFL students' writing skills compared to traditional tools, with ProWritingAid showing marginal superiority in stylistic development.

The implications of these findings suggest that AI-based applications such as Grammarly and ProWritingAid can be practical tools in the academic writing learning process. The

autocorrection, grammatical feedback, and stylistic analysis features provided by these two applications offer a more interactive and reflective learning experience than electronic dictionaries, which offer only word meanings without grammatical context. Institutions should integrate AI writing tools into academic writing courses, accompanied by digital literacy training. Future research should combine quantitative and qualitative methods, extend intervention duration, and explore impacts on critical thinking and originality.

This study is limited by its small sample (n=37), short intervention (3 sessions), and lack of longitudinal data. Generalizability may be affected by the homogeneity of participants.

BIBLIOGRAPHY

- [1] X. Fang, D. T. K. Ng, C. K. L. Leung, and S. K. W. Chu, “A Systematic Review of Artificial Intelligence Technologies Used for Story Writing,” *Educ Inf Technol*, pp. 1–37, 2023.
- [2] A. Han and Z. Cai, “Design Implications of Generative AI Systems for Visual Storytelling for Young Learners,” in *Proceedings of the 22nd Annual ACM Interaction Design and Children Conference*, Chicago IL USA: ACM, June 2023, pp. 470–474. doi: <https://doi.org/10.1145/3585088.3593867>
- [3] N. Nazari, M. S. Shabbir, and R. Setiawan, “Application of Artificial Intelligence Powered Digital Writing Assistant in Higher Education: Randomized Controlled Trial,” *Helijon*, vol. 7, no. 5, p. e07014, May 2021, doi: <https://doi.org/10.1016/j.heliyon.2021.e07014>
- [4] H. H. H. Ali, “The Importance of the Four English Language Skills: Reading, Writing, Speaking, and Listening in Teaching Iraqi Learners,” *Humanit. Nat. Sci. J.*, vol. 3, no. 2, Feb. 2022, doi: <https://doi.org/10.53796/hnsj3210>
- [5] J. M. Gayed, M. K. J. Carlon, A. M. Oriola, and J. S. Cross, “Exploring an AI-based Writing Assistant’s Impact on English Language Learners,” *Comput. Educ. Artif. Intell.*, vol. 3, p. 100055, 2022, doi: <https://doi.org/10.1016/j.caei.2022.100055>
- [6] N. Amani and M. Bisriyah, “University Students’ Perceptions of AI-Assisted Writing Tools in Supporting Self-Regulated Writing Practices,” *Indones. J. Engl. Lang. Teach. Appl. Linguist.*, vol. 10, no. 1, pp. 91–107, 2025. <https://doi.org/10.21093/ijeltal.v10i1.1942>
- [7] J. Huang and A. Mizumoto, “Examining the Effect of Generative AI on students’ Motivation and Writing Self-Efficacy,” *Digit. Appl. Linguist.*, vol. 1, p. 102324, Nov. 2024, doi: <https://doi.org/10.29140/dal.v1.102324>
- [8] I. Perdana, J. Bungai, and I. Ilham, “The Students’ Perceptions of AI-Generated and Teacher Feedback in English Writing Skills: A Comparative Study,” *J. Pendidik. Bhs. Ingg. Undiksha*, vol. 13, no. 1, pp. 22–31, May 2025, doi: <https://doi.org/10.23887/jpbi.v13i1.92668>
- [9] D. T. Y. G. Sumakul, F. A. Hamied, and D. Sukyadi, “Students’ Perceptions of the Use of AI in a Writing Class,” presented at the 67th TEFLIN International Virtual Conference & the 9th ICOELT 2021 (TEFLIN ICOELT 2021), Padang, Indonesia, 2022. doi: <https://doi.org/10.2991/assehr.k.220201.009>
- [10] N. Pellas, “The Effects of Generative AI Platforms on Undergraduates’ Narrative Intelligence and Writing Self-Efficacy,” *Educ. Sci.*, vol. 13, no. 11, p. 1155, Nov. 2023, doi: <https://doi.org/10.3390/educsci13111155>
- [11] M. Ahmadzade and M. Farahian, “At a Crossroad: Comparing ChatGPT and Grammarly in Enhancing EFL Learners’ Use of the Article ‘The,’” *J. Stud. Lang. Learn. Teach.*, vol. 2, no. 1, pp. 47–62, 2025, doi: <https://doi.org/10.2139/ssrn.5699162>

- [12] E. Sulistyowati, “Penerapan Grammarly Tool Untuk Meningkatkan Keterampilan Menulis Teks Analisis Ekspositoris Siswa,” *J. Educ. FKIP UNMA*, vol. 7, no. 2, pp. 559–566, June 2021, doi: <https://doi.org/10.31949/educatio.v7i2.1144>
- [13] S. Huang and P. Grady, “Generative AI: A Creative New World.” Sequoia Capital US/Europe., 2022.
- [14] R. Marrone, A. Zamecnik, S. Joksimovic, J. Johnson, and M. De Laat, “Understanding Student Perceptions of Artificial Intelligence as a Teammate,” *Technol. Knowl. Learn.*, vol. 30, no. 3, pp. 1847–1869, Sept. 2025, doi: <https://doi.org/10.1007/s10758-024-09780-z>
- [15] A. Azenoud, “Enhancing Writing Accuracy and Complexity through AI-Assisted Tools among Moroccan EFL University Learners,” *Int. J. Linguist. Transl. Stud.*, vol. 5, no. 4, pp. 211–226, 2024, doi: <https://doi.org/10.36892/ijlts.v5i4.529>
- [16] C. K. Y. Chan, “A Comprehensive AI Policy Education Framework for University Teaching and Learning,” *Int. J. Educ. Technol. High. Educ.*, vol. 20, no. 1, p. 38, July 2023, doi: <https://doi.org/10.1186/s41239-023-00408-3>
- [17] Y. A. Muhie and A. B. Woldie, “Teaching-Learning, Higher Education, AI, AI-Based Education, Implementation Framework,” *Sci. Technol.*, vol. 10, no. 1, pp. 1–7, 2020.
- [18] T. N. Fitria, “ProWritingAid as AI-Powered Writing Tools: The Performance in Checking Grammar and Spelling of Students’ Writing,” vol. 12, no. 2, 2023.
- [19] R. Schmidt-Fajlik, “ChatGPT as a Grammar Checker for Japanese English Language Learners: A Comparison with Grammarly and ProWritingAid,” *AsiaCALL Online J.*, vol. 14, no. 1, pp. 105–119, June 2023, doi: <https://doi.org/10.54855/acoj.231417>
- [20] G. S. Khalifa, A. M. Elshaer, S. A. El-Aidie, A. S. E. Yee, A. Farhan, and E. Gherbi A,H, “ChatGPT’s Challenges and Opportunities in Higher Education: A Technology-user (s)-environment lens.,” *Int. J. Manag. Educ.*, vol. 18, no. 5, pp. 453–480, 2024, doi: <https://doi.org/10.1504/IJMIE.2024.140898>
- [21] X. Zhai, X. Chu, C. S. Chai, M. S. Y. Jong, and A. Istenic, “A Review of Artificial Intelligence (AI) in Education from 2010 to 2020.,” *Complexity*, vol. 2021, no. 1, 2021, doi: <https://doi.org/10.1155/2021/8812542>
- [22] G. J. Hwang, H. Xie, B. W. Wah, and D. Gašević, “Vision, Challenges, Roles and Research Issues of Artificial Intelligence in Education.,” *Comput. Educ. Artif. Intell.*, vol. 1, p. 100001, 2020, doi: <https://doi.org/10.1016/j.caeari.2020.100001>
- [23] R. Luckin and W. Holmes, *Intelligence Unleashed: An Argument for AI in Education*. London, UK: UCL Knowledge Lab, 2016.
- [24] M. Stevenson and A. Phakiti, “Automated Feedback and Second Language writing,” in *Feedback in Second Language Writing: Contexts and Issues*, Second edition., United Kingdom: Cambridge University Press, 2019, pp. 125–142. <https://doi.org/10.1017/9781108635547.009>
- [25] M. Li, “Automated Writing Evaluation,” in *Researching and Teaching Second Language Writing in the Digital Age*, Cham: Springer International Publishing, 2022, pp. 151–181. https://doi.org/10.1007/978-3-030-87710-1_7
- [26] S. Cao and T. Zhou, “Exploring the Efficacy of ChatGPT-Based Feedback Compared With Teacher Feedback and Self-Feedback: Evidence From Chinese-English Translation,” *SAGE Open*, vol. 15, no. 3, p. 21582440251369204, Jan. 2025, doi: <https://doi.org/10.1177/21582440251369204>

[27] P. Wei, X. Wang, and H. Dong, "The impact of automated writing evaluation on second language writing skills of Chinese EFL learners: a randomized controlled trial," *Front. Psychol.*, vol. 14, p. 1249991, Sept. 2023, doi: <https://doi.org/10.3389/fpsyg.2023.1249991>

[28] N. Zhai and X. Ma, "Automated Writing Evaluation (AWE) Feedback: A Systematic Investigation of College Students' Acceptance," *Comput. Assist. Lang. Learn.*, vol. 35, no. 9, pp. 2817–2842, 2022, doi: <https://doi.org/10.1080/09588221.2021.1897019>

[29] W. Holmes, M. Bialik, and C. Fadel, *Artificial Intelligence in Education Promises and Implications for teaching and learning*. USA: Center for Curriculum Redesign, 2019.

[30] G. L. Parra and X. S. Calero, "Automated Writing Evaluation Tools in the Improvement of the Writing Skill," *Int. J. Instr.*, vol. 12, no. 2, pp. 209–226, Apr. 2019, doi: <https://doi.org/10.29333/iji.2019.12214a>

[31] G. Dizon and J. M. Gayed, "A Systematic Review of Grammarly in L2 English Writing Contexts," *Cogent Educ.*, vol. 11, no. 1, p. 2397882, 2024, doi: <https://doi.org/10.1080/2331186X.2024.2397882>

[32] K. D. Damayanti and M. H. Santosa, "The Effect of ProWritingAid Writing Tool on the Writing Skills of English as a Foreign Language (EFL) Students: A Systematic Literature Review," *Indones. J. Multidisciplinary Res.*, vol. 4, no. 2, pp. 469–478, Oct. 2024, doi: <https://doi.org/10.17509/ijomr.v4i2.78943>

[33] D. Miranty, U. Widiati, B. Y. Cahyono, and T. I. S. T. Sharif, "The Effectiveness of Using Grammarly in Teaching Writing Among Indonesian Undergraduate EFL Students," presented at the International Seminar on Language, Education, and Culture (ISoLEC 2021), Malang, Indonesia, 2021. doi: <https://doi.org/10.2991/assehr.k.211212.008>

[34] M. Rahma and E. L. Zen, "Errors in Students' Writing (Reports from ProWritingAid)," *Scope J. Engl. Lang. Teach.*, vol. 7, no. 2, p. 202, Apr. 2023, doi: <https://doi.org/10.30998/scope.v7i2.15528>

[35] P. Chauhan, "Fundamentals of Academic Writing: A Literature Review," *J. NELTA*, vol. 27, no. 1, 2022. <https://doi.org/10.3126/nelta.v27i1-2.53201>

[36] H. Aljuaid, "The Impact of Artificial Intelligence Tools on Academic Writing Instruction in Higher Education: A Systematic Review," *Arab World Engl. J.*, vol. 1, no. 1, pp. 26–55, Apr. 2024, doi: <https://doi.org/10.24093/awej/ChatGPT.2>

[37] P. D. Deep and Y. Chen, "The Role of AI in Academic Writing: Impacts on Writing Skills, Critical Thinking, and Integrity in Higher Education," *Societies*, vol. 15, no. 9, p. 247, Sept. 2025, doi: <https://doi.org/10.3390/soc15090247>

[38] K. Hyland, *Second language writing*. New York: Cambridge University Press, 2019. <https://doi.org/10.1017/9781108635547.009>

[39] S. Millah, "Peningkatan Kemampuan Menulis Deskripsi Bahasa Inggris Melalui Strategi Pembelajaran Partisipatif," *Bahtra J. Pendidik. Bhs. Dan Sastra*, vol. 24, no. 1, pp. 77–85, Jan. 2025, doi: <https://doi.org/10.21009/bahtra.241.08>

[40] Astari, "Barriers to Writing Skill Development in English Learners: A Systematic Review of Quantitative Studies (2020–2024)," *Asian J. Lang. Lit. Cult. Stud.*, vol. 8, no. 2, pp. 432–439, 2025.

[41] Y. Suharyat and L. Lusiana, "The Investigation of Students' Writing Skills in Learning English as a Second Language in Indonesia," *Eur. J. Engl. Lang. Stud.*, vol. 3, no. 1, pp. 47–57, June 2023, doi: <https://doi.org/10.12973/ejels.3.1.47>

- [42] M. Obisuru and W. Purbani, "Kemampuan Menulis Bahasa Inggris Siswa Melalui self-editing and self-correcting berdasarkan Analisis Kesalahan Gramatikal dan Kosakata," *LingTera*, vol. 3, no. 1, p. 51, May 2016, doi: <https://doi.org/10.21831/lt.v3i1.8473>
- [43] A. Susanti and A. T. Azizan, "Enhancing Writing Skills Through Chat GPT: A Literature Review," vol. 2, no. 2, 2025.
- [44] R. Rahmaniah, "The Effect of Peer Feedback on English Writing Skills Among EFL Students," vol. 13, no. 2, 2025.
- [45] J. W. Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage Publication, 2022.
- [46] A. Asari, Zulkarnaini, A. C. Hartatik, Anam, and Suparto, *Pengantar Statistika*. Solok, Sumatra Barat: PT Mafy Media Literasi Indonesia, 2023.
- [47] J. Pallant, *SPSS Survival Manual: A Step by Step Guide to Data Analysis using IBM SPSS*, 7th ed. London: McGraw-Hill Education., 2020. <https://doi.org/10.4324/9781003117445>
- [48] M. I. Wahyuda, "The Effectiveness of Grammarly and ProWritingAid Application toward Writing Skill across Students writing level of MAN Batu students," Universitas Islam Negeri Maulana Malik Ibrahim, Malang, 2022.