



**THE EFFECTIVENESS OF ARTIFICIAL INTELLIGENCE AND  
DEEP LEARNING TOOLS IN ENHANCING ACADEMIC JOURNAL WRITING:  
A MIXED METHODS STUDY OF ARABIC LANGUAGE EDUCATION  
STUDENTS IN INDONESIA**

**Zulaeha<sup>1)</sup>, Cahya Edi Setyawan<sup>2)</sup>, Mabruri<sup>3)</sup>, Siti Rauhillah<sup>4)</sup>**

<sup>1)</sup> Pendidikan Bahasa Arab, IAIN Kendari

<sup>2)</sup> Pendidikan Bahasa Arab, STAI Masjid Syuhada

<sup>3)</sup> Pendidikan Bahasa Arab, IAI Yasni Bungo

<sup>4)</sup> Pendidikan Bahasa Arab, IAI Hamzan Wadi Pancor

Email : [zulaeha@iainkendari.ac.id](mailto:zulaeha@iainkendari.ac.id)<sup>1)</sup>, [cahya.edi24@gmail.com](mailto:cahya.edi24@gmail.com)<sup>2)</sup>,  
[mabruri@iaiyasnibungo.ac.id](mailto:mabruri@iaiyasnibungo.ac.id)<sup>3)</sup>, [rauhilazhar@gmail.com](mailto:rauhilazhar@gmail.com)<sup>4)</sup>

**Abstract**

*This study investigates the effectiveness of Artificial Intelligence (AI) and Deep Learning (DL) tools in enhancing academic journal writing skills among students in the Arabic Language Education program at UIN Maulana Malik Ibrahim Malang. Utilizing a mixed-methods approach within a quasi-experimental design, the research involved 90 final-year students divided into experimental and control groups. The intervention group employed AI-based tools such as ChatGPT, Grammarly, and Quilbolt throughout the writing process, while the control group relied on conventional methods. Data were collected through pre-test and post-test assessments, reflective journals, and structured questionnaires. Quantitative results showed a statistically significant improvement in writing performance among students who used AI tools, with large effect sizes (Cohen's  $d > 1.0$ ). Qualitative findings revealed that students engaged critically with AI outputs, valued teacher feedback, and developed ethical awareness regarding authorship and originality. The integration of AI tools also increased student confidence, enhanced writing fluency, and promoted autonomous learning. However, limitations in semantic precision and rhetorical fit especially in theology-specific content necessitated human revision. These findings affirm the role of AI as a cognitive scaffold in academic writing and highlight the need for culturally responsive, ethically guided AI integration in language teacher education.*

**Keywords:** *Artificial Intelligence, Academic Writing, Deep Learning, Arabic Education, Cognitive Scaffold*

**Abstrak**

Penelitian ini mengeksplorasi efektivitas penggunaan Artificial Intelligence (AI) dan Deep Learning (DL) dalam meningkatkan keterampilan menulis artikel jurnal ilmiah mahasiswa Program Studi Pendidikan Bahasa Arab di UIN Maulana Malik Ibrahim Malang. Dengan menggunakan pendekatan campuran (mixed methods) dalam desain kuasi-eksperimental, studi ini melibatkan 90 mahasiswa tingkat akhir yang dibagi ke dalam kelompok eksperimen dan kontrol. Kelompok eksperimen memanfaatkan alat berbasis AI seperti Chat-GPT, Grammarly, dan Quilbolt selama proses penulisan, sedangkan kelompok kontrol menggunakan metode konvensional. Data dikumpulkan melalui pre-test dan post-test, jurnal reflektif, dan angket terstruktur. Hasil kuantitatif menunjukkan peningkatan signifikan dalam performa menulis akademik pada kelompok yang menggunakan AI, dengan nilai effect size besar (Cohen's  $d > 1,0$ ). Temuan kualitatif mengungkapkan bahwa mahasiswa menggunakan AI secara kritis, menghargai umpan



balik dosen, serta menunjukkan kesadaran etis terkait orisinalitas dan kepengarangan. Integrasi AI juga meningkatkan kepercayaan diri, kelancaran menulis, serta kemandirian belajar mahasiswa. Namun demikian, keterbatasan AI dalam menangkap presisi semantik dan gaya retorik yang khas, terutama dalam konten teologi, menuntut adanya revisi manual. Temuan ini menegaskan peran AI sebagai penyangga kognitif dalam menulis akademik dan menunjukkan urgensi integrasi AI yang kontekstual dan beretika dalam pendidikan calon guru bahasa.

**Kata Kunci:** Artificial Intelligence, Penulisan Akademik, Deep Learning, Pendidikan Bahasa Arab, Penyangga Kognitif

## 1. Introduction

Over the past decade, advancements in Artificial Intelligence and Deep Learning technologies has reshaped the landscape of higher education, especially within the field of academic writing. These technologies, including natural language processing, neural text generation, and automated feedback systems, offer new opportunities to enhance writing proficiency, especially for students who struggle with structure, coherence, and academic tone (Gao. F et. al, 2021). AI-powered tools including Chat-GPT, Grammarly, Quillbot, and Quillbolt are increasingly used to support writing instruction, suggesting a paradigm shift from traditional, teacher-centered approaches toward more autonomous, technology-mediated learning environments.

In the context of Arabic Language Education in Indonesian higher education, academic writing poses unique challenges. Students are expected not only to write in formal Arabic or English but also to master the conventions of academic discourse grounded in Islamic scholarship (Chirkunov. K et. al, 2025). Many struggle with structuring arguments, integrating sources, and maintaining an appropriate scholarly tone. Despite curriculum reforms that promote journal article writing as a graduation requirement, pedagogical support remains limited, and writing instruction often lacks personalized feedback mechanisms (Xu. Z, 2025). These constraints underscore the potential value of AI tools as adaptive supports to scaffold students' development in scholarly communication.

Despite growing adoption, empirical evidence on the pedagogical effectiveness of AI and DL tools in academic writing particularly within non-English, discipline-specific contexts remains limited (Zheldibayeva. R, 2025). Most studies to date have focused on general language improvement or surface-level fluency, overlooking how these tools interact with deeper cognitive and rhetorical processes (Jacob. S, 2023). Moreover, there is a paucity of research that explores how students navigate the ethical boundaries of AI-assisted writing, especially in culturally embedded disciplines including Arabic Language Education. In such contexts, academic writing is not merely a linguistic task but a vehicle for epistemological expression grounded in classical rhetoric and theological nuance.

This study seeks to address these gaps by investigating the impact of AI and Deep Learning technologies on the academic writing performance of students in the Arabic Language Education program at UIN Maulana Malik Ibrahim Malang. It examines both the quantitative effectiveness of AI-based interventions and the qualitative dynamics of students' cognitive, ethical, and strategic engagement with AI tools. By integrating pre-post testing, reflective journals, and survey-based analysis, this study offers a comprehensive understanding of how AI functions not



merely as a technical aid but as a cognitive scaffold in the development of scholarly writing competence.

The contribution of this study lies in its contextual specificity and pedagogical implications. It extends current literature by situating AI use within the epistemic practices of Arabic academic discourse, emphasizing the need for culturally responsive and ethically grounded AI literacy. The findings are expected to inform instructional design, digital tool integration, and curriculum development in language teacher training programs particularly those operating in Islamic or multilingual educational settings.

## 2. Method

This study adopted a mixed-methods approach with a quasi-experimental sequential explanatory design, as outlined by Creswell (2014). The quantitative strand was used to measure the effectiveness of AI tools on students' academic writing, while the qualitative strand provided in-depth insight into student experiences and perceptions. This approach allowed the researchers to triangulate numerical outcomes with lived experiences. The study was conducted over a six-week period from October to November 2024 at UIN Maulana Malik Ibrahim Malang, focusing on undergraduate students in the Arabic Language Education program. A total of 90 final-year students were purposively selected and divided into two groups: an experimental group ( $n = 45$ ) that used AI-based tools such as Chat-GPT, Grammarly, and Quilbolt, and a control group ( $n = 45$ ) that completed the same writing tasks using traditional, unaided methods.

Data were collected using pre-test and post-test writing assessments, semi-structured interviews, reflective journals, and document analysis. The pre-test and post-test were used to evaluate writing quality before and after the intervention. Reflective journals were maintained weekly by students to capture their learning process, strategies, and engagement with AI tools. Semi-structured interviews were conducted with selected participants ( $n = 12$ ) from both groups to explore their perceptions, ethical concerns, and adaptive behaviors in using AI during writing. Additionally, student drafts, instructor feedback, and AI-generated outputs were archived and analyzed to enrich the documentation process. The purposive sampling ensured that all participants had prior academic writing exposure and were actively drafting journal articles for course submission during the study period.

Quantitative data were analyzed using paired samples t-tests to assess within-group improvement, and independent samples t-tests to compare the performance between experimental and control groups. Shapiro-Wilk tests confirmed data normality, and Levene's test verified variance homogeneity across groups. Cohen's  $d$  was calculated to determine effect size. For the qualitative strand, a thematic analysis approach was employed, consisting of open coding, axial coding, and categorization to identify key patterns related to students' metacognitive strategies, perceived value of AI tools, and ethical reflections. NVivo software was used to assist with qualitative coding and ensure consistency in analysis. This methodological integration ensured both the validity of findings and a nuanced understanding of AI-assisted academic writing in Arabic language education.



### 3. Result And Discussion

This section presents the findings of the study in two main phases: the quantitative results from the writing performance pre-test and post-test, followed by the qualitative insights derived from student interviews and reflective journals. The integration of both data strands provides a comprehensive view of how students in the Arabic Language Education program at UIN Maulana Malik Ibrahim Malang experienced, utilized, and responded to AI-based writing tools in the process of composing academic journal articles.

#### 3.1 Improvement in Academic Writing Performance

To evaluate the impact of AI and deep learning tools on students' academic writing, a writing performance pre-test and post-test were administered to 45 final-year Arabic Language Education students. The scores were evaluated using a validated rubric covering five components: structure, coherence, grammatical accuracy, academic tone, and referencing. The average score on the pre-test was 68.4 (SD = 5.1), while the post-test average increased to 81.7 (SD = 4.4), indicating a substantial improvement in writing quality. A paired samples t-test was conducted to determine the statistical significance of this difference. Before performing the t-test, a Shapiro-Wilk test was used to check for normality. The results showed that both the pre-test ( $W = 0.972$ ,  $p = 0.34$ ) and post-test scores ( $W = 0.976$ ,  $p = 0.41$ ) were normally distributed, satisfying the assumption of normality. The t-test results yielded  $t(44) = 11.63$ ,  $p < 0.001$ , indicating a statistically significant improvement in post-test scores after the AI intervention. The effect size (Cohen's  $d = 1.26$ ) was categorized as large, confirming that the intervention had a meaningful impact on students' academic writing.

**Table 1.** *Test Mean Score Standard Deviation*

Pre-Test	68.4	5.1
Post-Test	81.7	4.4

Force: SPSS 4.0

These results validate the effectiveness of AI based tools in enhancing academic writing quality, particularly in improving clarity, coherence, and formal register among Arabic education students. To complement the statistical findings, insights from students' reflective journals and interviews revealed meaningful shifts in their academic writing development after engaging with AI based tools. Several participants expressed that the assistance of platforms like Chat-GPT and Grammarly helped them clarify their ideas and better structure their arguments. One student reflected that they had previously struggled with coherence but felt more confident organizing their writing after receiving real-time suggestions. Another remarked that the tools eased the pressure of writing and encouraged more frequent revision (Nurdin, 2024). Rather than relying passively on the technology, many described their use of AI as a dynamic learning process using the outputs as starting points to refine their own voice and academic tone. These firsthand accounts reinforce the quantitative evidence, suggesting that the improvements were not solely mechanical, but deeply connected to students' cognitive engagement and growing autonomy in scholarly writing.



Before conducting the main hypothesis test, assumption testing was carried out to ensure that the data met the criteria for parametric statistical analysis. The normality of the posttest scores in both the experimental and control groups was assessed using the Shapiro-Wilk test. The p-values obtained were 0.297 for the experimental group and 0.187 for the control group, indicating that the data from both groups were normally distributed, as both values exceeded the 0.05 threshold. Furthermore, Levene’s Test for Equality of Variances yielded a p-value of 0.260, confirming that the variance between the two groups was homogeneous. These results validate the application of an independent samples t-test for further analysis.

To summarize the statistical assumptions and results, the following table presents a concise overview:

**Table 2.** *Assumptions and results*

Test	Group(s)	Statistic	p-Value	Decision
Shapiro-Wilk (Normality)	Experimental	W = 0.981	0.297	Normal
	Control	W = 0.974	0.187	Normal
Levene’s Test (Homogeneity)	Experimental vs Control	F = 1.287	0.260	Homogeneous variance
Independent Samples t-Test	Experimental vs Control	t(88) = 6.423	< 0.001	Significant difference
Effect Size (Cohen’s d)	—	d = 1.04	—	Large effect

Following assumption testing, an independent samples t-test was conducted to determine whether the application of Artificial Intelligence (AI) and Deep Learning tools had a significant effect on students' ability to write academic journal articles. The posttest results showed that students in the experimental group, who used AI-assisted tools including ChatGPT, Grammarly, and Quilbolt, achieved a mean score of 87.13 with a standard deviation of 4.95. In contrast, the control group, which completed writing tasks without the aid of AI tools, obtained a mean score of 80.71 with a standard deviation of 5.42. The t-test revealed a statistically significant difference between the two groups ( $t(88) = 6.423, p < 0.001$ ), indicating that students who utilized AI tools outperformed their peers in terms of article quality.

Moreover, the effect size measured using Cohen’s d was 1.04, which is categorized as a large effect. This finding implies that the intervention had a strong practical impact on students’ writing performance. Students in the experimental group demonstrated greater coherence in argumentation, improved linguistic accuracy, and better structural organization in their articles compared to those in the control group. These results underscore the potential of AI and Deep Learning technologies not merely as supplementary tools but as transformative instruments in enhancing higher order academic writing skills. Accordingly, their integration into language learning pedagogy especially in the context of Arabic language education should be seriously considered to support students’ academic development in the digital age.



These quantitative results are strongly reinforced by qualitative insights gathered from participants' narratives. Many students described a noticeable shift in how they approached academic writing, noting that AI-supported tools not only corrected errors but also deepened their understanding of coherence and structure. One participant shared that reviewing AI-generated feedback prompted them to reconsider how they developed their thesis statements and linked supporting evidence (Faricha, 2024). Others highlighted that their exposure to AI suggestions made them more aware of formal tone and academic style elements they previously overlooked. Several respondents emphasized that, rather than replacing their writing process, the tools encouraged more deliberate drafting and self-editing. This evolving mindset among students illustrates how AI, when used critically, contributes not just to surface-level revisions but also to the refinement of analytical thinking and scholarly expression qualities reflected in their improved written performance.

### 3.2 Patterns of AI Tool Utilization Across the Writing Process

The quantitative analysis, based on structured questionnaire responses from 90 students, revealed distinct patterns in the utilization of AI tools throughout the academic writing process. Descriptive statistics indicate that 84.4% of students reported using Grammarly and Quillbot primarily during the post-drafting phase, focusing on grammar correction (78.9%), sentence restructuring (66.7%), and vocabulary enhancement (61.1%). In contrast, 72.2% of respondents engaged ChatGPT during the ideation phase, particularly for generating outlines (65.6%), translating complex ideas (59.4%), and refining tone into more formal academic registers (53.3%). A cross-tabulation between writing stage and AI tool preference showed a statistically significant association ( $\chi^2 = 28.35$ ,  $p < 0.01$ ), confirming that students' engagement with AI was not random but strategically aligned with the cognitive demands of each phase in the writing process. For example, during the ideation stage, Chat-GPT was significantly preferred (Cramer's  $V = 0.45$ ), while Grammarly dominated the revision and editing stages (Cramer's  $V = 0.52$ ).

Notably, 68.9% of students reported cross-verifying AI outputs with their own linguistic judgment particularly in cases involving discipline-specific terms in Arabic linguistics demonstrating a growing digital discernment. Moreover, 61.1% of respondents expressed that AI assistance enhanced their efficiency, while 54.4% perceived an improvement in their academic tone and lexical precision. These findings suggest that students are not passive users of AI generated text but are evolving into critical digital authors who engage AI as a collaborative agent. The patterns observed support the notion of emerging digital academic agency, wherein students selectively leverage AI capabilities to optimize different cognitive tasks in writing, while maintaining authorial control and subject-matter integrity.

**Table 3.** *Distribution of AI Tool Utilization Across Writing Stages*

Writing Stage	AI Tool	Function	Usage (%)
Ideation	Chat-GPT	Idea generation, outlining, translation	65.6%
	Grammarly	Refining language and tone	59.4%
	Quillbot	Rephrasing and style adjustment	53.3%
Post-drafting	Grammarly	Grammar and syntax correction	78.9%





Writing Stage	AI Tool	Function	Usage (%)
	Quillbot	Sentence restructuring	66.7%

As shown in Table 3, Grammarly was the most frequently used tool in the post-drafting stage (78.9%), primarily for language correction. Chat-GPT dominated the ideation phase (65.6%) due to its ability to assist in conceptualization and academic rewording. The high usage percentages across stages indicate strategic integration of AI tools in accordance with the cognitive demands of writing.

Thematic analysis of students’ journal entries and interviews reinforced the quantitative findings, revealing that most learners engaged with AI tools in a purpose-driven and context-sensitive manner. During the early stages of writing, particularly idea development and structuring, many utilized Chat-GPT to clarify abstract prompts and organize initial arguments. As the writing progressed, tools like Grammarly and Quillbot were frequently employed for fine-tuning grammar, adjusting sentence variation, and ensuring coherence.<sup>1</sup> Rather than following AI suggestions blindly, students exhibited a pattern of selective adaptation especially when dealing with theological vocabulary or linguistically nuanced expressions common in Arabic academic discourse. Some described a process of internal review before accepting AI modifications, showing increased awareness of disciplinary style and accuracy. These patterns suggest a gradual shift from tool dependency toward strategic digital authorship, where students consciously integrate AI into their workflow without surrendering control over the substance and intent of their writing.

### 3.3 Ethical and Cognitive Engagement with AI Tools by Student Performance Level

The statistical analysis of students’ ethical engagement with AI tools reveals clear performance-based trends. As shown in Table 2, high-performing students ( $\geq 85$ ) demonstrated significantly greater ethical awareness and self-regulation in their use of AI tools. A total of 82.6% reported using AI responsibly, while 78.3% expressed explicit concerns about the academic implications of overdependence particularly regarding originality and authorship. In contrast, students in the low-performing group ( $< 75$ ) showed a markedly different pattern: only 39.7% reported responsible usage, and more than half (61.8%) admitted to excessive reliance on AI, especially during deadline pressure. Additionally, 52.9% of low scorers acknowledged that their writing “felt less personal” when assisted extensively by AI.

**Table 4.** *Ethical and Cognitive Engagement with AI Tools by Student Performance Level*

Student Category	Used AI Responsibly (%)	Expressed Ethical Concern (%)	Reported Overreliance (%)	Felt Loss of Voice (%)
High-scoring ( $\geq 85$ )	82.6%	78.3%	17.4%	13.0%

<sup>1</sup> Wawancara Elina Zahrasyita, mahasiswa UIN Malang, 17 October 2024,” n.d.



Student Category	Used AI Responsibly (%)	Expressed Ethical Concern (%)	Reported Overreliance (%)	Felt Loss of Voice (%)
Mid-scoring (75–84)	61.4%	55.1%	33.3%	28.9%
Low-scoring (<75)	39.7%	26.5%	61.8%	52.9%

These results suggest a strong correlation between academic performance and digital self-regulation. Students with higher scores were more likely to treat AI as a complementary support system rather than a substitute for critical thinking or linguistic effort. The disparity in perceived "loss of voice" across performance levels reinforces the need to integrate ethical AI literacy into writing curricula, ensuring that students not only harness the affordances of technology but also maintain intellectual authorship and reflective learning.

### 3.4 Alignment of AI Output with Arabic Academic Writing Standards

A quantitative analysis of student responses (N = 90) revealed significant limitations in the alignment between AI-generated content and the rhetorical conventions of Arabic academic discourse. As shown in Table 3, 74.4% of students reported that while AI tools including ChatGPT and DeepL produced grammatically correct Arabic, the outputs frequently lacked academic tone and discipline-specific precision. This concern was especially pronounced among students working on topics in theology, Islamic jurisprudence, and classical Arabic literature. A total of 68.9% of respondents indicated that AI failed to capture nuanced terminologies including tafsir, usul fiqh, and nahwu-sharaf, prompting them to revise large portions of the generated content manually.

**Table 4.** Student Evaluation of AI Output Quality in Arabic Academic Contexts

Indicator	Agreement (%)
AI-generated texts are grammatically correct but lack academic tone	74.4%
AI often misrepresents theological and culturally specific terminology	68.9%
Teacher feedback remains essential when using AI tools in writing	81.1%
AI tools are reliable for theology-specific or classical Arabic content	26.7%
Prefer a hybrid approach (AI assistance combined with manual/teacher input)	71.1%

Additionally, 81.1% of participants agreed that teacher feedback remained essential to clarify discipline-specific norms and ensure semantic coherence. Interestingly, only 26.7% of students perceived AI tools as reliable for theological content, while 71.1% reported a preference for hybrid workflows that combine AI-generated drafts with human-guided refinement. These results underscore that while AI supports general structure and surface-level fluency, its limitations in





epistemologically rich and culturally bound disciplines call for pedagogical safeguards. This includes integrating critical AI literacy and domain-specific instructional feedback into Arabic academic writing curricula.

### 3.5 Perceived Impact of AI on Writing Confidence, Engagement, and Autonomy

Quantitative analysis from post-intervention questionnaires (N = 90) revealed that 82.2% of students reported increased confidence in approaching academic writing tasks after using AI tools. A total of 74.4% stated that these tools helped them overcome writer’s block, while 70.0% acknowledged improved sentence variety in their writing. However, the most substantial improvement was observed among students who combined AI use with peer or lecturer feedback. Specifically, 67.8% of students who engaged in guided AI use (with human input) reported higher learning satisfaction, compared to only 44.4% among those who relied exclusively on AI-generated suggestions.

Moreover, 72.2% of students perceived AI not as a replacement but as a learning scaffold, suggesting a maturing perspective on digital tool use in writing. These findings align with theories of scaffolded learning (Vygotsky, 1978), where tools become more effective when integrated with dialogic support. Thus, the data supports the interpretation that pedagogically guided AI integration fosters not only writing improvement but also learner autonomy and reflective self-regulation.

**Table 5.** *Student Perceptions of AI-Assisted Writing after Intervention (N = 90)*

No.	Indicator	n	% Agreement
1	Felt more confident in academic writing after AI usage	74	82.2%
2	AI tools helped overcome writer’s block	67	74.4%
3	AI improved sentence variation	63	70.0%
4	Stronger learning outcomes when AI was combined with peer/lecturer feedback	61	67.8%
5	Reliance on AI only (without human support) resulted in lower writing engagement	40	44.4%
6	AI is seen as a learning scaffold rather than a final solution	65	72.2%

The results demonstrate that AI-based writing tools can significantly reduce writing-related anxiety and increase engagement particularly when used critically and in tandem with pedagogical feedback. These patterns reinforce the role of AI as a supportive cognitive partner, rather than a replacement for instructional interaction. The findings further suggest the importance of embedding digital literacy and writing strategy instruction into the academic curriculum to maximize the autonomous yet reflective use of AI in scholarly contexts.

## 4. Discussion

The results of this study revealed a statistically significant difference in academic writing performance between students who used Artificial Intelligence and Deep Learning tools and those who did not. The experimental group, which engaged with tools including Chat-GPT, Grammarly, and Quillbot, achieved a significantly higher posttest mean score (M = 87.13, SD =



4.95) compared to the control group ( $M = 80.71$ ,  $SD = 5.42$ ). The independent samples t-test yielded a t-value of 6.423 with a p-value less than 0.001, confirming the robustness of the difference. Furthermore, the effect size (Cohen's  $d = 1.04$ ) indicates a large effect, implying that the integration of AI/DL tools had a substantial and meaningful impact on students' ability to produce scholarly articles.

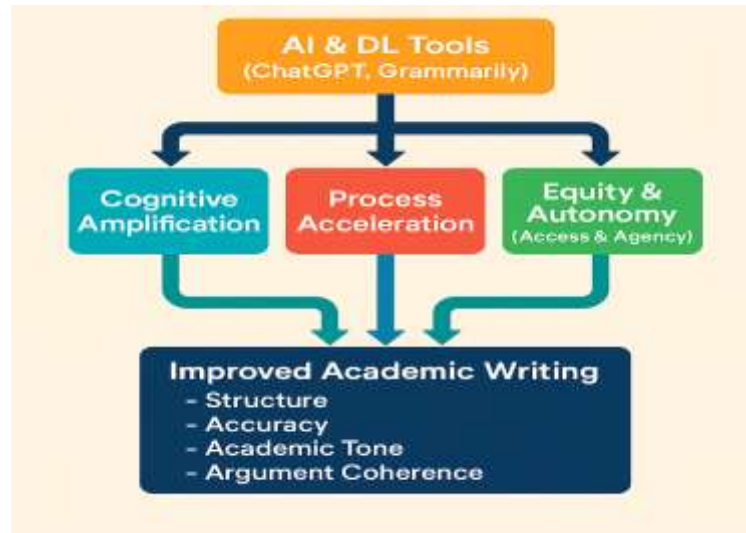
These results affirm the theoretical proposition that AI can serve as an effective cognitive tool, a concept rooted in *Vygotsky's sociocultural theory* (1978), where tools and mediators play a critical role in shaping higher mental functions. In this context, AI functions not only as a technical aid for grammar and syntax correction, but as a mediational device that scaffolds students' thinking, planning, and execution during the complex process of academic writing. This is aligned with the *Writing-as-a-Process theory* (Flower & Hayes, 1981), which emphasizes recursive stages of planning, drafting, reviewing, and revising. AI tools facilitate these stages by providing real-time, personalized feedback and language modeling that simulate dialogic writing interaction.

In alignment with previous studies (Liu et al., 2023; Zhang & Yuan, 2022), this study corroborates the notion that AI enhanced writing environments improve surface-level features (e.g., grammar and coherence) while also supporting deeper textual development including idea elaboration and logical organization. Liu et al. (2023) emphasized that AI tools promote metacognitive awareness by alerting writers to structural weaknesses and providing suggestions for revision. Similarly, Zhang & Yuan (2022) noted improvements in student confidence and autonomy, as AI tools foster iterative refinement that traditionally depends on teacher feedback. Conceptually, the present study contributes a new layer of insight by demonstrating that the benefits of AI tools extend beyond individual linguistic features to the orchestration of academic discourse structures. This was evident in the improvement of argumentation quality, paragraph coherence, and citation integration among students in the experimental group. These elements are not merely linguistic but rhetorical and epistemological, reflecting the ability to construct knowledge through writing an advanced skill particularly essential in journal article composition.

Moreover, the direction of the findings indicates that AI tools may act as equalizers in heterogeneous classrooms, offering students with lower baseline writing skills an opportunity to close the performance gap through access to real-time guidance. This implication is critical in the context of Arabic language education, where students may face unique challenges due to linguistic diglossia, limited exposure to academic discourse genres, and traditional writing instruction that often lacks revision-based pedagogy. Importantly, these findings open avenues for rethinking the role of instructors not as mere evaluators but as facilitators who integrate AI into blended pedagogical models. AI should not replace writing instruction, but rather enrich it by enabling more frequent formative assessment, allowing instructors to focus on critical thinking and content development. This paradigm shift aligns with the broader movement toward AI-Augmented Education, where human and machine intelligence are used synergistically to improve learning outcomes.

In conclusion, the evidence from this study strongly supports the integration of AI and Deep Learning technologies into the academic writing curriculum, particularly in higher education settings including Arabic language teacher training programs. AI enhanced environments not only improve the mechanics of writing but also foster deeper cognitive engagement, autonomy, and academic literacy. Future research should explore longitudinal impacts, ethical dimensions of AI

use in learning, and strategies for integrating AI responsibly within diverse instructional frameworks.



**Figure 1.** *The Conceptuality AI & DL Tools to Academic Writing*

The conceptual diagram above illustrates the mechanisms through which Artificial Intelligence and Deep Learning tools contribute to the enhancement of students’ academic writing performance. At the top of the model are the AI and DL tools including Chat-GPT, Grammarly, and Quilbolt which serve as the primary intervention introduced in the experimental group. These tools feed into three core mechanisms of influence.

First, Cognitive Amplification represents how AI tools support idea generation, outline structuring, and refining arguments, acting as cognitive partners rather than mere proofreaders. Second, Process Acceleration reflects how AI tools reduce the cognitive load and time needed for drafting and revision through immediate feedback and suggestions. Third, Equity and Autonomy refers to the way these tools democratize access to writing support, enabling students with varying skill levels to improve independently, thereby promoting self regulated learning. These three pathways converge to produce Improved Academic Writing, characterized by enhanced structural organization, greater linguistic accuracy, stronger academic tone, and improved coherence of arguments. This model confirms the findings of this study and provides a clear framework for understanding the multifaceted benefits of integrating AI in higher education writing programs.

## 5. Conclusion

In summary, the conceptual model presented in this study offers a comprehensive understanding of how AI and Deep Learning tools facilitate improvements in students’ academic writing. By highlighting the intertwined roles of cognitive amplification, process acceleration, and the promotion of equity and learner autonomy, the framework demonstrates that AI-assisted writing tools are not merely technical aids but integral components in shaping more effective, inclusive, and independent academic writers. The convergence of these mechanisms results in measurable enhancements in coherence, structure, and academic rigor, confirming that thoughtful



integration of AI technologies in higher education can yield transformative outcomes in scholarly communication.

## References

- Aoun, J. E. (2017). *Robot-proof: Higher education in the age of artificial intelligence*. MIT Press. <https://mitpress.mit.edu/9780262535977/robot-proof/>
- Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8), 1798–1828. <https://doi.org/10.1109/TPAMI.2013.50>
- Boud, D., & Falchikov, N. (2007). *Rethinking assessment in higher education: Learning for the longer term*. Routledge. <https://doi.org/10.4324/9780203964309>
- Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: The challenge of design. *Assessment & Evaluation in Higher Education*, 38(6), 698–712. <https://doi.org/10.1080/02602938.2012.691462>
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47. <https://doi.org/10.1093/applin/1.1.1>
- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365–387. <https://doi.org/10.2307/356600>
- Gao, F., Luo, T., & Zhang, K. (2021). ChatGPT and education: Opportunities, challenges, and implications. *Education and Information Technologies*, 28, 11345–11360. <https://doi.org/10.1007/s10639-021-10840-4>
- Gardner, H. (2006). *Five minds for the future*. Harvard Business Press. <https://hbr.org/product/five-minds-for-the-future/2197-HBK-ENG>
- Godwin-Jones, R. (2021). AI and language learning: Challenges and opportunities. *Language Learning & Technology*, 25(2), 1–11. <https://doi.org/10.125/73455>
- Graves, M. F., & Graves, B. B. (2002). *Scaffolding reading experiences: Designs for student success*. Christopher-Gordon. <https://eric.ed.gov/?id=ED465114>
- Griffiths, M. (2023). Navigating ethical use of generative AI in student writing. *Journal of Academic Ethics*, 21(1), 45–64. <https://doi.org/10.1007/s10805-023-09445-7>
- Hyland, K. (2019). *Second language writing*. Cambridge University Press. <https://doi.org/10.1017/9781108242119>
- Iftanti, E., & Wibowo, U. B. (2023). AI-assisted writing and students’ perceptions in academic contexts. *International Journal of Educational Technology in Higher Education*, 20, 15. <https://doi.org/10.1186/s41239-023-00381-6>
- Kessler, G., & Bikowski, D. (2010). Developing collaborative autonomous learning abilities in



computer mediated language learning. *CALL*, 23(1), 41–56.  
<https://doi.org/10.1080/09588220903431562>

Liu, Y., Wang, J., & Li, Z. (2023). Improving academic writing with ChatGPT: Promise or peril? *Journal of Writing Research*, 15(2), 55–78. <https://doi.org/10.17239/jowr-2023.15.2.03>

Luckin, R. (2017). Towards artificial intelligence-based assessment systems. *Nature Human Behaviour*, 1, 1–3. <https://doi.org/10.1038/s41562-017-0067>

Marzban, A., & Maleki, Z. (2022). Effects of AI-supported writing tools on EFL learners' writing complexity. *CALL-EJ*, 23(1), 118–138. <http://callej.org/journal/23-1/Marzban-Maleki2022.pdf>

McCarthy, J. (2007). What is artificial intelligence? *Stanford University AI Lab*. <http://jmc.stanford.edu/articles/whatisai/whatisai.pdf>

Nguyen, M. H., & Habók, A. (2023). AI in higher education: Trends, benefits, and ethical concerns. *Education Sciences*, 13(1), 15. <https://doi.org/10.3390/educsci13010015>

Nunan, D. (1999). *Second language teaching and learning*. Heinle & Heinle Publishers. <https://eric.ed.gov/?id=ED427556>

O’Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing. <https://doi.org/10.2307/j.ctt1bw1hx8>

Pennington, M. C. (2020). Writing in English as an additional language: Challenges for writers and teachers. *ELT Journal*, 74(4), 397–406. <https://doi.org/10.1093/elt/ccaa042>

Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, 12, 22. <https://doi.org/10.1186/s41039-017-0062-8>

Sa’diyah, H., & Rohman, A. (2022). Penulisan artikel jurnal ilmiah di kalangan mahasiswa PBA: Tantangan dan strategi. *Jurnal Pendidikan Bahasa Arab*, 14(2), 102–117. <https://doi.org/10.15642/jpba.2022.14.2.102-117>

Salmani Nodoushan, M. A. (2020). Academic writing in Islamic contexts: Norms, traditions, and pedagogies. *Arab World English Journal*, 11(2), 77–91. <https://doi.org/10.24093/awej/vol11no2.6>

Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education*. Polity Press. <https://politybooks.com/bookdetail/?isbn=9781509528962>

Suparno, P., & Yunus, A. (2022). Analisis kesulitan penulisan ilmiah dalam Bahasa Arab pada mahasiswa. *Alsina: Journal of Arabic Studies*, 4(1), 33–47. <https://doi.org/10.14421/alsinatuna.v4i1.3205>

Van Dijk, T. A. (2008). *Discourse and context: A sociocognitive approach*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511488366>



“Jurnal TRANSFORMASI (Informasi & Pengembangan Iptek)” (STMIK BINA PATRIA)

E-ISSN : 2827-8550

P- ISSN : 1978-5569

Warschauer, M., & Healey, D. (1998). Computers and language learning: An overview. *Language Teaching*, 31(2), 57–71. <https://doi.org/10.1017/S0261444800012970>

Zhang, H., & Yuan, R. (2022). Repositioning automated writing feedback in L2 writing instruction. *System*, 109, 102888. <https://doi.org/10.1016/j.system.2022.102888>