

THE EFFECTIVENESS OF THE CANVA AI PLATFORM IN DIAGNOSTIC ASSESSMENT FOR AL-IMLA' WAL KHAT LEARNING

Muhammad Zaqhlul Rafif¹, Safara Bunaiya Hibda², Tulus Musthofa³, Zamakhsari⁴

^{1,3,4}Universitas Islam Negeri Sunan Kalijaga Yogyakarta; Indonesia

²Universitas Negeri Malang; Indonesia

Correspondence Email; muhammadzaqhlulrafif30@gmail.com

Submitted: 03/06/2026

Revised: 15/06/2026

Accepted: 01/07/2026

Published: 13/07/2026

Abstract

This study systematically evaluates the effectiveness of the Canva AI platform as a diagnostic assessment instrument for Arabic writing proficiency, specifically in the aspects of *Al-Imla'* (orthography) and *Al-Khat* (calligraphy), using the Systematic Literature Review (SLR) method. By adapting the PRISMA 2020 guidelines, this research synthesizes empirical evidence and indexed scientific literature from the 2016–2026 period across the Google Scholar, DOAJ, and Dimensions databases. From the 240 identified articles, nine core studies examining the intersection of artificial intelligence, visual-textual feedback, and Arabic linguistic processing were analyzed in depth. The synthesis results indicate that the visual generative features, automatic pattern recognition, and instantaneous feedback of Canva AI significantly accelerate the identification of granular orthographic errors, such as *hamzah* positioning, *alif layyinah* placement, and letter-joining discrepancies, and map student weakness profiles more intuitively than manual methods. Structurally, this platform functions as a practical "low-code" intermediary that bridges complex deep learning algorithms with pedagogical needs in the classroom. Although AI spatial analytics are effective in measuring the geometric proportions of calligraphy, the purely aesthetic assessment of *Al-Khat* still requires human expert validation. This research constructs a conceptual framework that integrates Information Systems Success Theory and Slavin's QAIT pedagogical model and concludes that Canva AI functions optimally as a secondary diagnostic assistant. This Human-AI collaboration paradigm reduces the administrative burden on teachers, thereby facilitating the design of precise and evidence-based remediation plans.

Keywords

AI-Assisted Language Learning; Arabic Orthography; Diagnostic Assessment; Handwriting Skills.



© 2026 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY NC) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

INTRODUCTION

Mastery of writing skills within the constellation of Arabic language education rests on two complementary academic pillars: orthographic accuracy through *Al-Imla'* rules and structural-aesthetic precision through the art of *Al-Khat* (Setiyadi et al., 2024; Habibah et al., 2024). Ideally, *Al-Imla'*'s competence requires learners to possess cognitive and psychomotor abilities to convert heard phonemes into accurate graphemes in accordance with standard writing rules. On the other hand, *Al-Khat* learning trains the proportions, layout, fine motor coordination, and visual harmonization of Arabic letters (Mohsen et al., 2026). The integration of these two aspects is vital for building functional Arabic literacy, where writing must not only be structurally legible but also representatively presented.

However, the sharp gap between these theoretical expectations and classroom practical realities often triggers deep academic anxiety. Previous research has highlighted several key findings regarding these challenges and the resulting technological interventions. First, various elementary orthographic errors ranging from fatal errors in distinguishing *hamzah washal* and *qatha'*, misplacement of *alif layyinah*, to structural distortions in connecting specific letters still dominate learners' writing (Munas, 2022). Second, the integration of Artificial Intelligence (AI) in language assessment has successfully introduced a promising new paradigm for automated evaluation (Azmi et al., 2019). This automation is further strengthened by recent frameworks that emphasize a tridimensional collaboration between human expertise and generative AI to minimize bias in language evaluation (Zubaidi et al., 2025). Third, automated parameter-efficient approaches are proven to provide detailed diagnoses of grammatical and orthographic errors (Mahmoud et al., 2024), utilizing enhanced neural network architectures to step up the recognition of handwritten inputs (Qasim & Oleiwi, 2024). Fourth, artificial intelligence-based systems effectively provide instant error pattern analysis and orthographic correction (Alamri & Teahan, 2019). Fifth, deep learning models are capable of diagnosing complex orthographic error patterns, reducing human assessment bias, and optimizing administrative efficiency (Almanea, 2024), which fundamentally reconceptualizes how formative feedback shapes the cognitive tracking of second language writing acquisition (Liu & Yu, 2022).

This systemic condition is exacerbated by assessment paradigms in most educational institutions that tend to be summative-oriented and uniform (one-size-fits-all). Consequently, educators face difficulties diagnosing the locations of students' specific cognitive deficits or motor

obstacles early on, resulting in delayed pedagogical interventions or remediation programs. Particularly in Arabic linguistics, AI technology is required to process unique and complex morphological characteristics and visual representations of Arabic writing. In this context, the Canva AI platform offers tactical solutions through an integrated ecosystem that includes Magic Studio, Magic Activities, and Visual Suite.

Empirical evidence shows that tools like Canva Magic Writer significantly assist students in structuring text and enhancing drafting proficiency (Putri et al., 2025). while providing a dynamic visual interface that allows learners to self-correct paragraph alignment and mechanical errors in real time (Utami & Karnedi, 2024). Consequently, these features no longer function merely as commercial graphic design tools; instead, they can be re-engineered into interactive diagnostic instruments capable of providing precise calligraphy modeling, designing self-directed dictation modules, and visually mapping student errors. This transition transforms the platform into a multi-purpose environment that boosts short-story communication and writing proficiency across diverse learning styles (Hafidzin et al., 2024; Ulfayati et al., 2025). Through this dialogic approach to classroom innovation, digital interfaces act as active scaffolds for writing skills development rather than passive display media (Eragamreddy & Joseph, 2025; Hinchcliff & Mehmet, 2023; Le et al., 2023).

To provide a solid scientific foundation for the effectiveness of this technology, the operationalization of Canva AI needs to be analyzed through a comparative theoretical integration. This study combines the Information Systems Success Theory from William (2003) with the QAIT (Quality, Appropriateness, Incentive, Time) Model of Instructional Effectiveness from Slavin (1994). Through the lens of DeLone & McLean (2003), effectiveness is measured by System Quality in reliably processing input data, as well as Information Quality which refers to the precision of AI analytical outputs when detecting errors. Meanwhile, Slavin's QAIT model contextualizes effectiveness from an instructional standpoint: the quality of clear materials (Quality), the appropriateness of materials to the students' diagnostic levels (Appropriateness), visual appeal that motivates students (Incentive), and time efficiency in execution and assessment (Time). This theoretical collaboration provides a robust argument that automated assessment using generative visuals is crucial for detecting fundamental weaknesses in Arabic writing.

Although studies on artificial intelligence in language assessment continue to proliferate, the current literature landscape remains dominated by English-language testing or advanced essay syntax analysis. There is a significant gap in the literature regarding a comprehensive synthesis of

Canva AI applications contextualized specifically for diagnosing basic *Al-Imla' wal Khat* skills. Previous research has not systematically mapped how the generative and analytical features of this platform can be operationalized as valid diagnostic measurement tools. Therefore, this study aims to fill this void while asserting its novelty by dissecting, extracting, and synthesizing current empirical evidence on the functionality of Canva AI as an *Al-Imla' wal Khat* diagnostic assessment instrument through a comprehensive systematic review.

METHOD

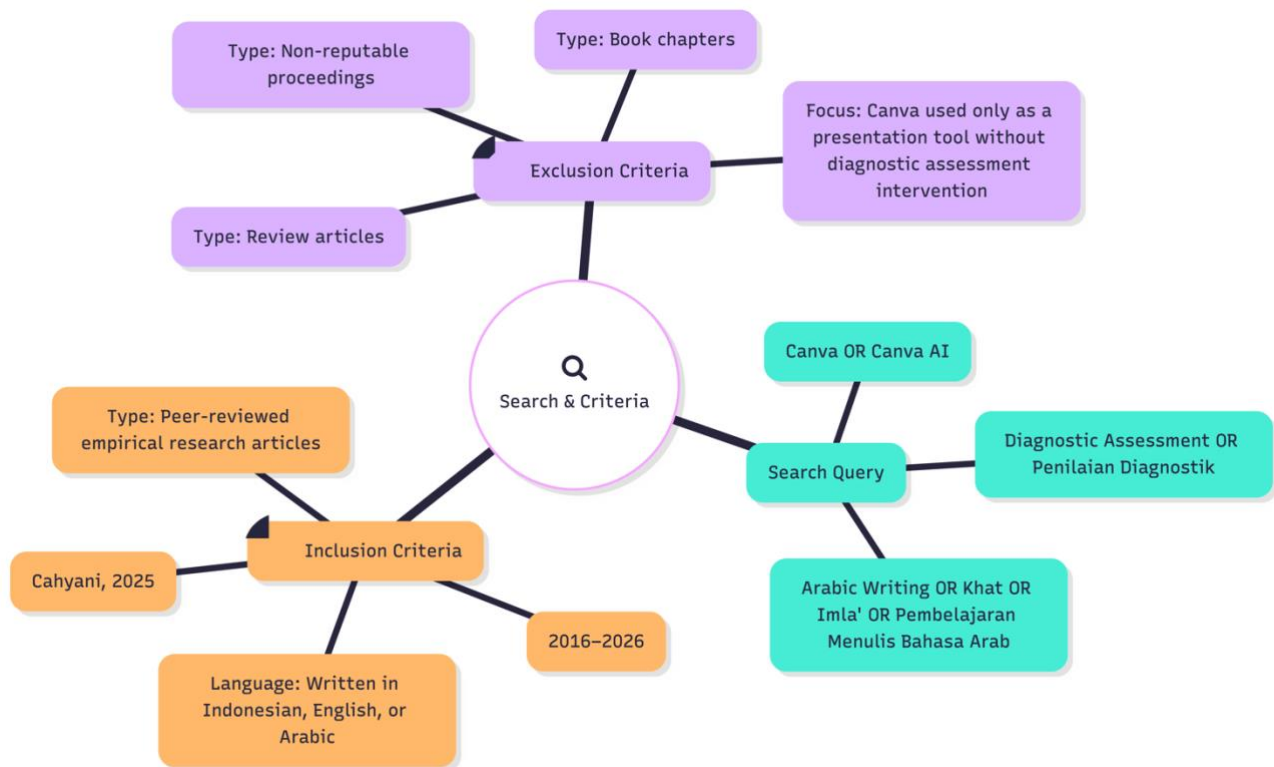
This study employs the Systematic Literature Review (SLR) method by adapting the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 reporting guidelines (Crocker et al., 2025; Grimshaw et al., 2021; Muradi, 2011; Oto-Millera et al., 2025; Page et al., 2021). The application of the SLR method with the PRISMA protocol aims to ensure that the collection, selection, and analysis of the literature are conducted transparently and systematically, and can be replicated by future researchers. This systematic review is focused on mapping the effectiveness of artificial intelligence features on the Canva platform in facilitating diagnostic assessments of *Al-Imla'* (orthography) and *Al-Khat* (calligraphy) in Arabic writing instruction. The data collection process was carried out through a comprehensive search across three main academic databases: Google Scholar, Directory of Open Access Journals (DOAJ), and Scopus. Search strategies were formulated using Boolean operators (AND, OR) to restrict while sharpening the relevance of findings. The search syntax used was: ("Canva AI" OR "Canva") AND ("diagnostic assessment") AND ("khat" OR "imla'" OR "Arabic writing instruction" OR "Arabic writing").

To systematically operationalize the PRISMA protocol within this research, the literature selection process was executed through three concrete stages:

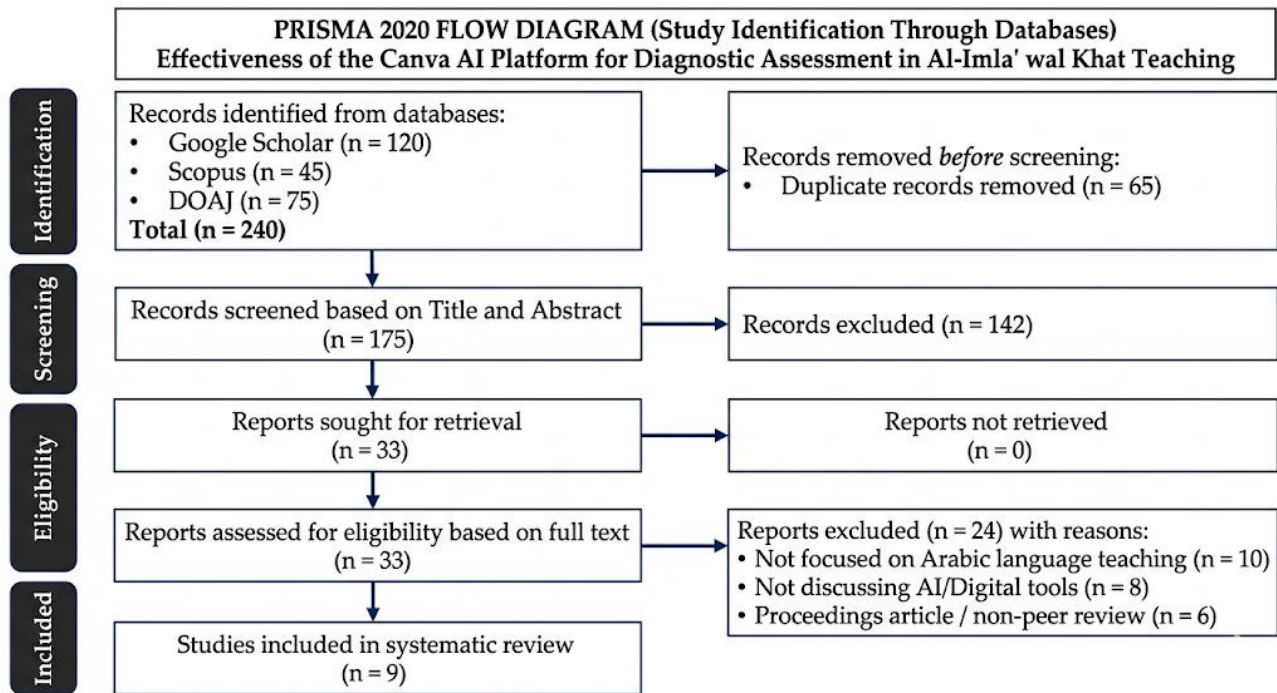
- **Identification:** The initial search across databases yielded a total of 240 records, comprising Google Scholar (n=120), DOAJ (n=75), and Scopus (n=45). Before the screening phase, 65 duplicate records were systematically removed.
- **Screening:** A total of 175 unique records were screened based on the relevance of their titles and abstracts to the primary research objective. This phase resulted in the exclusion of 142 records that did not match the core theme.
- **Eligibility:** The remaining 33 reports were successfully retrieved and assessed thoroughly based on full-text evaluations against the inclusion and exclusion criteria. During this phase, 24 reports

were explicitly excluded with specific reasons: studies not focused on Arabic language teaching (n=10), studies not discussing AI or digital tools (n=8), and non-peer-reviewed proceedings articles (n=6). Ultimately, this rigorous selection process yielded 9 core empirical studies for synthesis.

Figure 1. SLR Search Protocols and Selection Criteria Configuration



The inclusion criteria established were: (1) scientific journal articles based on empirical research that have undergone a peer-review process; (2) published within the last ten years (2016–2026) to maintain relevance with current AI technology developments; (3) written in Indonesian, English, or Arabic; and (4) explicitly discussing the integration of digital platforms or artificial intelligence in writing assessment (Cahyani, 2025). Exclusion criteria included review articles, non-reputable proceedings, book chapters, and articles that discuss Canva solely as a presentation tool and do not include a diagnostic assessment intervention. The literature selection process is presented in detail in the following PRISMA 2020 flow diagram:

Figure 2. PRISMA 2020 Flow Diagram of the Study Selection Process

The details in the diagram present a comprehensive synthesis of the research gap map from previous literature, while also confirming the novelty of Canva AI's contribution to *Al-Imla' wal Khat* diagnostic assessments. The mapping of this matrix empirically demonstrates that a collaborative approach between humans and artificial intelligence can address functional deficits identified in previous assessment models. To structure the entire process of conceptual discovery systematically and measurably, the continuation of the synthesis presentation process is discussed in the following section.

FINDINGS AND DISCUSSION

Findings

Based on the results of searching, screening, and extracting data through the Systematic Literature Review (SLR) procedure, it was found that the implementation of artificial intelligence and digital design platforms has entered the domain of linguistic assessment, although literature specifically integrating Canva AI features for diagnostic assessments of *al-impla'* (orthography) and *al-khat* (calligraphy/typography) remains in its early stages of development. The synthesis of relevant literature indicates that the current focus of technological intervention is on automating the detection of syntactic, morphological, and orthographic errors using Deep Learning, as well as on

using visual anomalies to evaluate text precision. An analysis of data mapping from the extracted main literature is presented in Table 1:

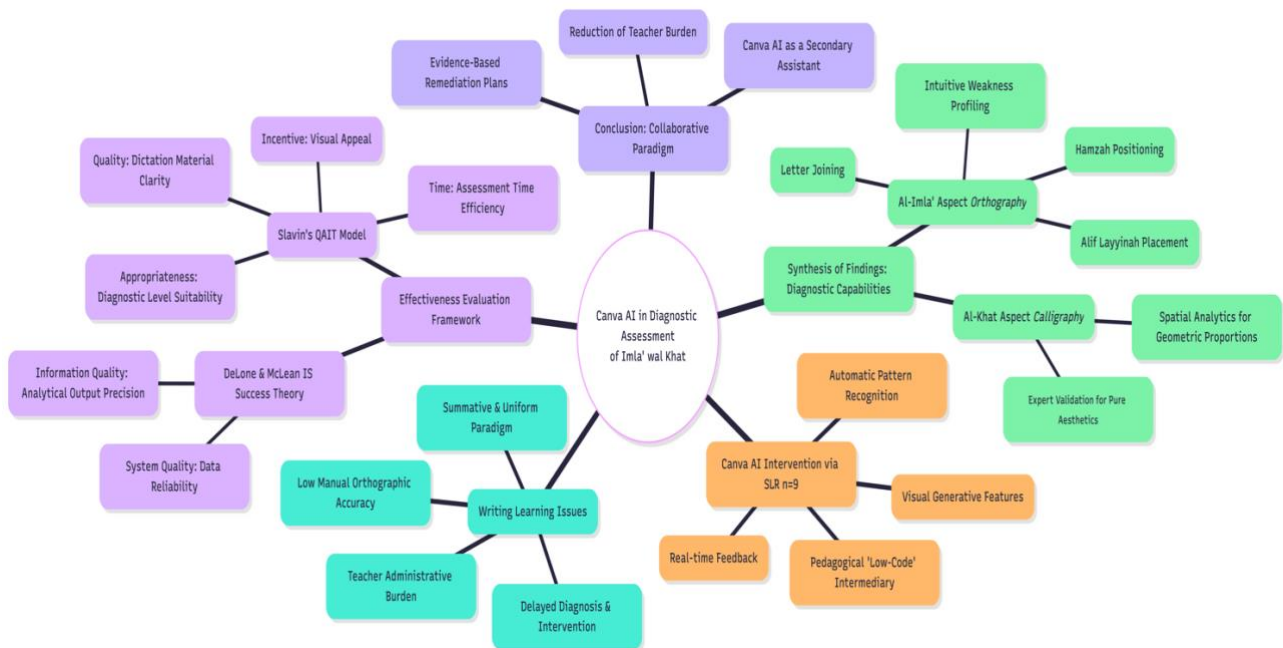
Table 1. Literature Extraction of AI and Digital Platform Integration in Arabic Orthography and Calligraphy Assessment

No	Author & Year	Focus of Canva AI Feature / Technological Intervention	Findings on Imla'/Khat
1	Alamri & Teahan (2019)	Computational-based automatic text recognition and correction for dyslexia.	Effectively identifies and instantly corrects <i>imla'</i> (orthography) errors in Arabic texts written by individuals with specific disorders.
2	Azmi et al. (2019)	Automated essay evaluation (AAEE) system with natural language processing (NLP).	Shows a high correlation with human assessment in detecting word writing errors, basic spelling, and syntactic structure.
3	Michalak & Ellixson (2025)	Human-tool collaboration in AI literacy integration.	Visual and textual feedback from AI-based platforms improves writing accuracy through a self-editing process.
4	Wati & Ahmed (2023)	Diagnostic application-based writing error detection (Sahehly).	Significantly reduces the basic <i>imla'</i> error rate, specifically in diacritic placement and letter joining.
5	Almanea (2024)	Utilization of Deep Learning in linguistic studies and the Arabic writing system.	Artificial neural network models can diagnose complex orthographic error patterns with precision surpassing manual methods.
6	Attar (2025)	Deep Convolutional Neural Network (DCNN) untuk pengenalan karakter tulisan tangan.	Capable of automatically detecting and evaluating isolated handwritten Arabic characters (<i>khat</i>) based on the precision of curves and lines.
7	Mahmoud et al. (2024)	Automatic assessment of Arabic essays with a parameter-efficient approach.	Provides detailed diagnoses of grammatical and orthographic (<i>imla'</i>) errors as a database for learning remediation.
8	Elsayed et al. (2025)	Human-AI collaborative framework (ZaQQ Dataset) for writing assessment.	Cross-validation between AI analytics and human expertise yields a highly accurate diagnostic instrument for assessing Arabic writing quality.
9	Mohammed (2025)	Evaluation of LLM-driven translation and writing text quality.	Indicates that generative AI is effective in diagnosing lexical inconsistencies but requires expert verification for the aesthetic aspects of <i>khat</i> .

To deepen the scientific gap analysis and establish the novelty position of this research among existing literature, an article screening matrix based on research gaps and the solutive contributions offered by the Canva AI diagnostic model was compiled:

Table 2. Screening Matrix Based on Research Gaps and Solutive Contributions

No	Author & Year	Main Focus	Identified Research Gaps	Novelty Contribution (Canva AI)	
1	Alamri & Teahan (2019)	Clinical text correction	Limited to special dyslexia populations and purely clinical.	Adapt orthographic correction functions into an interactive visual design platform for regular classes.	
2	Azmi et al. (2019)	Essay evaluation (NLP)	Hanya berorientasi pada esai teks panjang; mengabaikan estetika visual huruf (khat).	Menyediakan modul penilaian diagnostik awal yang menggabungkan aspek visual huruf (khat) dan teks (imla').	
3	Michalak & Ellixson (2025)	Human-tool collaboration	Focuses on English; has not touched the complexities of Arabic graphemes and ligatures.	Utilizes Canva's visual interface to track alignment and aesthetics of Arabic lettering.	
4	Wati & Ahmed (2023)	Sahehly application	The application is rigid; teachers cannot redesign the assessment instruments.	Provides a flexible no-code platform where teachers can independently customize diagnostic quizzes.	
5	Almanea (2024)	Linguistic Learning	Deep	Highly technical-computational approach; difficult for non-IT teachers to access.	Presents deep learning analysis results in a visually readable Canva dashboard for teachers.
6	Attar (2025)	DCNN characters	hand	Evaluation is limited to isolated single characters, not in flowing text (cursive).	Facilitates a diagnostic assessment instrument covering whole-word writing and layout analysis.
7	Mahmoud et al. (2024)	Automated essay assessment	essay	Only processes text and basic spelling of Arabic grammar; does not assess aesthetic aspects.	Integrates visual assessment of letter proportions through Canva AI's geometric grid and ruler features.
8	Elsayed et al. (2025)	ZaQQ Dataset		Evaluation is static in a lab; untested in dynamic classroom learning interactions.	Presents real-time diagnostic assessment in class using interactive Canva media.
9	Mohammed (2025)	LLM evaluation		LLMs are often biased when assessing the visual-aesthetic proportions of calligraphy (<i>khat</i>).	Formulates a collaborative framework: AI filters spelling errors, and teachers validate the art of calligraphy.

Figure 3. Mind Map of Research Synthesis and Effectiveness Evaluation Framework

The synthesis of findings from the scientific literature indicates that integrating artificial intelligence has a transformative impact on the effectiveness of diagnostic assessment in *Al-Imla' wal Khat* learning. Conventional diagnostic assessments in Arabic writing instruction have long faced obstacles such as prolonged grading time and high evaluator subjectivity, especially when analyzing typographical precision errors in calligraphy (khat) and letter-writing rules involving complex morphemes. The presence of AI-based visual and textual analytical features, as adapted from spatial pattern recognition technology (Attar, 2025) and automated spell correction (Alamri & Teahan, 2019), allows educators to perform real-time and aggregated error mapping.

From the standpoint of Information Systems Success Theory by William (2003), the effectiveness of the Canva AI platform is determined by System Quality, which is reliable in processing input data rapidly, and Information Quality, which is presented as precise visual representations and spelling error maps. The utilization of an AI-backed visual platform facilitates the identification of learners' misconceptions at a highly granular level. Errors in *hamzah* positioning, omission of *alif lam syamsiyah* writing, and inaccuracies in the vertical and horizontal line proportions of *khat* can be automatically detected. The generated diagnostic data is no longer merely a rigid final summative score; rather, it takes the form of a detailed orthographic error map. This shifts the assessment paradigm from mere outcome evaluation to dynamic process intervention.

Pedagogically, analysis using Slavin's QAIT confirms that this technology simultaneously optimizes four dimensions of instructional effectiveness. Teaching Quality (Quality) increases because automated feedback provided by the AI system is proven to accelerate learners' cognitive processes in recognizing and fixing their writing errors independently through a self-editing process (Azmi et al., 2019; Michalak & Ellixson, 2025). Instructional appropriateness (Appropriateness) is achieved because educators can accurately diagnose students' competency levels before designing learning materials. This optimization aligns with the principles of authentic diagnostic assessment in Islamic educational frameworks, which seek to pinpoint individual learning thresholds systematically (Fadlillah & Kusaeri, 2024). By using automated diagnostic data, contemporary educators can overcome core pedagogical challenges in Islamic higher education by moving away from generic instructional modes toward targeted skill remediation (Kamal, 2025). Ultimately, this approach provides a highly structured environment that can be integrated directly into specialized language textbooks to strengthen both receptive and productive language competencies simultaneously (Mufid et al., 2023).

Discussion

Cognitive Mechanisms of Automated Feedback in Reducing Error Fossilization

The most crucial psycholinguistic impact of automated feedback facilitated by Canva AI is the prevention of fossilization of orthographic errors among students. In learning Arabic as a foreign language, writing errors that are not immediately corrected tend to become permanent (fossilized) in students' long-term memory due to continuous incorrect motor repetition. When students write Arabic words in an interactive Canva AI template and receive instantaneous visual feedback or an automated spell check, an immediate cognitive dissonance occurs, prompting students to compare the graphemes they produced with the correct orthographic standards. This repeated self-correcting process (self-editing) strengthens neural pathways in recognizing the relationship between Arabic phonemes and graphemes, thereby significantly increasing *imla'* writing accuracy.

Critical analysis of recent literature indicates that these findings simultaneously reinforce, complement, and offer new viewpoints toward previous research postulates regarding Arabic language assessment. First, these findings strengthen the conventional error analysis theory put forward in applied linguistics literature (Al-Hamad & Mohamed, 2020; Taha et al., 2014). Those studies emphasize that Arabic orthography errors are generally systematic and influenced by the complexity of the script system as well as phonological interference (Almusawi, 2023; Tibi et al.,

2021). In line with this, natural language processing (NLP) studies conducted by Almanea (2024) and Mahmoud et al. (2024), reinforce this premise by proving that these orthographic error patterns can be computationally codified, predicted, and diagnosed with a high degree of accuracy.

This study complements the literature related to digital learning media, which previously only positioned design platforms like Canva merely as one-way visual presentation tools or passive means of knowledge transfer (Kerzel, 2021; Thohir & Muslimah, 2020). Various previous studies focused more on the effectiveness of specific learning applications (Wati & Ahmed, 2023) or text topic modeling (Brahmi et al., 2012). Through this synthesis, the researcher complements the discourse by developing an understanding that AI-integrated commercial visual platforms can be used as two-way diagnostic assessment instruments. AI features on design platforms are capable of bridging the gap between visual aesthetic evaluation (*khat*), analyzed via a Convolutional Neural Network (Attar, 2025) and textual evaluation (*imla'*), processed through natural language processing (Ismail et al., 2026).

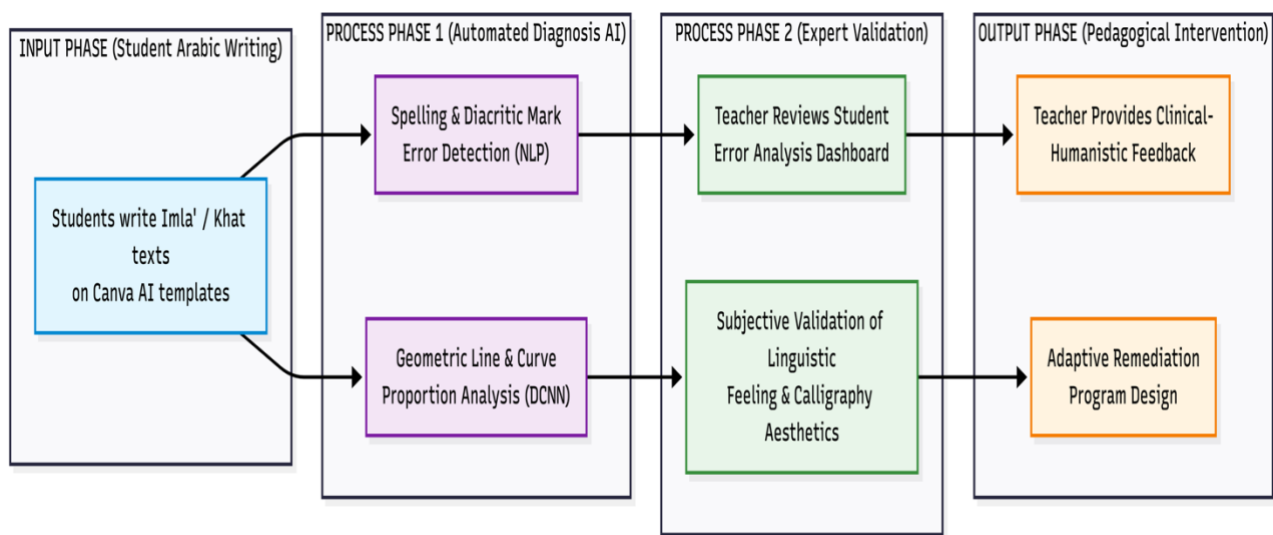
To understand the effectiveness of Canva AI comprehensively, it is important to compare the computational characteristics when assessing *Al-Imla'* and *Al-Khat*. The characteristics of *Al-Imla'* writing are discrete, digital, and rule-based. Arabic spelling rules, such as hamzah positioning, can be codified as logical decision trees or natural language processing (NLP) models with binary certainty levels (correct or incorrect). Conversely, the characteristics of *Al-Khat* writing are continuous, spatial, and geometric. Calligraphy assessment involves pixel analysis, pen (*qalam*) tilt ratio, baseline alignment, and the visual balance of empty spaces. To compute these variables accurately, recent studies have introduced deep learning vectorization models that preserve the high-fidelity curves and strokes of complex Arabic calligraphic scripts (Allaf, 2025). Consequently, while *Al-Imla'* can be entirely delegated to automated text correction algorithms, *Al-Khat* assessment demands more intricate image processing using convolutional neural networks integrated with precise visual tools. This computational precision supports the broader instructional implementation and evaluation efficiency of traditional scripts like Khat Naskhi within higher education classrooms (Widmann et al., 2022).

These findings offer a new perspective on the importance of a collaborative framework between human intelligence and machines (Human-AI Collaboration) in Arabic language education, a concept that is beginning to emerge in contemporary digital pedagogy literature (Elsayed et al., 2025; Windholz & Michalak, 2026). While initial studies tended to view AI automation

as a complete replacement for the teacher's role in the classroom (Azmi et al., 2019), recent literature synthesis explicitly warns of the fundamental limitations of artificial intelligence. Generative AI remains vulnerable to transliteration (Mahmoudi et al., 2026) and has limitations in independently assessing language feeling or pure aesthetic values in calligraphy (Mohammed, 2025).

Therefore, the new perspective offered by this study is that Canva AI is most effective when positioned as a secondary diagnostic assistant (*diagnostic assistant*). This framework harmoniously integrates the machine's computational advantage with the teacher's pedagogical expertise:

Figure 4. Mapping of Research Findings: Diagnostic Assessment Framework for Al-Imla' wal Khat Learning



Through this division of labor, artificial intelligence instantly and tirelessly takes over high-volume mechanical correction tasks. Meanwhile, human educators can allocate their mental capacity and time to focus on higher instructional aspects: providing clinical interpretations of students' learning difficulties, offering emotional motivational encouragement, and designing humanistic, student-centered remediation actions (Michalak, 2025; Mirza et al., 2026). This synergy ensures that the utilization of technology in the era of artificial intelligence continues to preserve the human dimension in Arabic language education.

CONCLUSION

This systematic review concludes that the integration of the Canva AI platform shows promising effectiveness as an auxiliary tool in the diagnostic assessment of *Al-Imla' wal Khat* learning, with an operational mechanism rooted in the convergence between analytical automation and pedagogical validation. A body of evidence from the literature indicates that artificial

intelligence capabilities in natural language processing and visual pattern recognition can identify orthographic anomalies and inaccuracies in the typographical proportions of Arabic letters at a granular level. Through its ability to map misconceptions instantaneously, ranging from omissions of basic morpheme rules to the precision of calligraphy curves, this platform facilitates the transformation of assessment from a mere accumulation of summative scores into structured error diagnostic data. The acceleration of automated feedback provision has been shown to effectively stimulate learners' self-editing capabilities, thereby expediting mastery of proper writing rules without neglecting the aesthetic aspects of writing. Future implications of this study demand a paradigm shift among Arabic language educators, moving from manual evaluators, vulnerable to time constraints and subjectivity, to clinical facilitators skilled in interpreting AI diagnostic data to design adaptive remediation programs. In conclusion, Canva AI is not positioned to replace the humanistic intuition and expertise of teachers; instead, it serves as a catalyst that optimizes error-mapping efficiency, thereby opening a broad avenue for further empirical research to directly test the validity of this human-AI collaboration model in real learning ecosystems.

REFERENCES

- Al-Hamad, M., & Mohamed, Y. (2020). An Orthographic Phonological-Based Error Analysis of the Arabic of English-Speaking Learners. *The Language Scholar*, 6, 8–25.
- Alamri, M. M., & Teahan, W. J. (2019). Automatic Correction of Arabic Dyslexic Text. *Computers*, 8(1), Article 19. <https://doi.org/10.3390/computers8010019>
- Alif Cahya Setiyadi, Afin Al Aufi, Pafi Rahmatullah Al Gifari, & Jalaluddin. (2024). Nāmūdḥaj Ta'allum al-Khaṭ fi Mahārah al-Kitābah fi al-Ma'had al-'Aṣrī Dār al-Salām Kūntūr al-Sābi' Kālīandā Lāmbūng al-Janūbiyyah. *Al Mahāra: Jurnal Pendidikan Bahasa Arab*, 10(2), 316–325. <https://doi.org/10.14421/almahara.2024.0102-07>
- Allaf, S. R. (2025). A Novel Deep Learning Approach for High-Fidelity Vectorization of Arabic Calligraphy. *Journal of Advances in Mathematics and Computer Science*, 40(12), 1–14. <https://doi.org/10.9734/jamcs/2025/v40i122069>
- Almanea, M. (2024). Deep Learning in Written Arabic Linguistic Studies: A Comprehensive Survey. *IEEE Access*, 12, 172196–172233. <https://doi.org/10.1109/ACCESS.2024.3496123>
- Almusawi, H. (2023). Factors Affecting the Writing Performance in Hearing and Deaf Children: An Insight into Regularities and Irregularities of the Arabic Orthographic System. *Language and Speech*, 66(1), 246–264. <https://doi.org/10.1177/00238309221087352>
- Attar, E. T. (2025). Deep Convolutional Neural Network for Isolated Arabic Handwritten Character Recognition: Design, Evaluation, and Comparative Study. *Scientific Reports*, 15(1), Article 42467. <https://doi.org/10.1038/s41598-025-42467-w>
- Azmi, A. M., Al-Jouie, M. F., & Hussain, M. (2019). AAEE–Automated Evaluation of Students' Essays in Arabic Language. *Information Processing & Management*, 56(5), 1736–1752. <https://doi.org/10.1016/j.ipm.2019.05.008>
- Brahmi, A., Ech-Cherif, A., & Benyettou, A. (2012). Arabic Texts Analysis for Topic Modeling

- Evaluation. *Information Retrieval*, 15(1), 33–53. <https://doi.org/10.1007/s10791-011-9170-1>
- Crocker, T. F., Todd, O., & Clegg, A. (2025). Around the Equator with Clin-Star: Systematic Reviews of Intervention Effectiveness; Challenges and Opportunities. *Journal of the American Geriatrics Society*, 73(8), 2356–2365. <https://doi.org/10.1111/jgs.19498>
- DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean Model of Information Systems Success: A Ten-Year Update. *Journal of Management Information Systems*, 19(4), 9–30. <https://doi.org/10.1080/07421222.2003.11045748>
- Elsayed, Y., Nabil, E., Torki, M., Faizullah, S., & Khalafallah, A. (2025). ZaQQ: A New Arabic Dataset for Automatic Essay Scoring Via a Novel Human–AI Collaborative Framework. *Data*, 10(9), Article 148. <https://doi.org/10.3390/data10090148>
- Eragamreddy, N., & Joseph, R. (2025). Digital Tools and Writing Education: A Thematic Analysis of Technology’s Role in Writing Skills Development. *Arab World English Journal*, 16(3), 3–25. <https://doi.org/10.24093/awej/vol16no3.1>
- Fadlillah, N., & Kusaeri, K. (2024). Optimizing Assessment for Learning in Islamic Education through Authentic and Diagnostic Assessment : A Systematic Literature Review. *Jurnal Kependidikan: Jurnal Hasil Penelitian dan Kajian Kepustakaan di Bidang Pendidikan, Pengajaran dan Pembelajaran*, 10(2), 654. <https://doi.org/10.33394/jk.v10i2.11555>
- Febrina Dwi Cahyani, M. S. M. (2025). The Development of Arabic Linguistics in Educational Studies: An Analysis of Ibn Khaldun’s Thought. *Jurnal Islam Nusantara*, 9(3). <https://doi.org/https://doi.org/10.33852/jurnalnu.v9i3.709>
- Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McGuinness, L., McDonald, S., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., Moher, D., Glanville, J., Chou, R., Brennan, S. E., Boutron, I., Akl, E., & Tetzlaff, J. M. (2021). Pravila PRISMA 2020. *Medicina Fluminensis*, 57(4), 444–465. https://doi.org/10.21860/medflum2021_264903
- Habibah, I. F., Rahmayanti, I., Holis, H., & Wargadinata, W. (2024). Analysis of the Psychological Foundation of the Implementation of Learning the Four Arabic Language Skills. *Kitaba*, 2(1), 19–35. <https://doi.org/10.18860/kitaba.v2i1.24933>
- Hafidzin, H. K., Muzammil, L., & Sulistyo, T. (2024). The Effectiveness of the Canva Application in Enhancing Students’ Writing Proficiency Across Learning Styles. *Edulitics (Education, Literature, and Linguistics) Journal*, 9(2), 130–139. <https://doi.org/10.52166/edulitics.v9i2.7958>
- Hinchcliff, M., & Mehmet, M. (2023). Embedding Canva into the Marketing Classroom: A Dialogic and Social Learning Approach to Classroom Innovation. *Higher Education, Skills and Work-Based Learning*, 13(6), 1174–1186. <https://doi.org/10.1108/HESWBL-11-2022-0230>
- Ismail, A. F., Alqudah, R. A., Al-Saliti, R. A. M. N., & Hamid, A. A. (2026). Grammatical error Patterns in ChatGPT-Generated Modern Standard Arabic Texts: A Linguistic Analysis of Recurrent Patterns. *Languages*, 11(5), Article 86. <https://doi.org/10.3390/languages11050086>
- Kamal, H. (2025). Teaching Arabic Today: Challenges, Strategies, and Opportunities in Islamic Higher Education. *International Journal of Learning, Teaching and Educational Research*, 24(10), 644–659. <https://doi.org/10.26803/ijlter.24.10.31>
- Kerzel, U. (2021). Enterprise AI Canvas Integrating Artificial Intelligence into Business. *Applied Artificial Intelligence*, 35(1), 1–12. <https://doi.org/10.1080/08839514.2020.1843231>
- Le, A. N. N., Bo, L. K., & Nguyen, N. M. T. (2023). Canva-Based E-portfolio in L2 Writing Instructions: Investigating the Effects and Students’ Attitudes. *Computer Assisted Language Learning Electronic Journal (Call-Ej)*, 24(1), 41–62.
- Liu, C., & Yu, S. (2022). Reconceptualizing the Impact of Feedback in Second Language Writing: A Multidimensional Perspective. *Assessing Writing*, 53, 100630.

- <https://doi.org/10.1016/j.asw.2022.100630>
- Mahmoud, S., Nabil, E., & Torki, M. (2024). Automatic Scoring of Arabic Essays: A Parameter-Efficient Approach for Grammatical Assessment. *IEEE Access*, 12, 142555–142568. <https://doi.org/10.1109/ACCESS.2024.3468501>
- Mahmoudi, M., Moulahi, W., & Jdey, I. (2026). A Native Supervision Approach to Arabic VLM: Overcoming Transliteration Bias for Semantic Accuracy. *Proceedings of the 18th International Conference on Agents and Artificial Intelligence - Volume 5: ICAART*, 4643–4650. <https://doi.org/10.5220/0014480900003636>
- Michalak, R. (2025). From Canvas to Quartex: The Evolution of art in the Age of AI and Digital Archives. *College & Research Libraries News*, 86(2), 63.
- Michalak, R., & Ellixson, D. (2025). Fostering Ethical AI Integration in First-Year Writing: A Case Study on Human-Tool Collaboration in Artificial Intelligence Literacy. *Journal of Library Administration*, 65(3), 361–377. <https://doi.org/10.1080/01930826.2025.2428751>
- Mirza, A. Z., Zunairoh, Y., & Ahid, N. (2026). Transforming Arabic Academic Writing Competence through ADDIE-based instructional Design: A Developmental Study. *Jurnal Pendidikan Islam*, 12(1), 1–9.
- Mohammed Haneefa Abdul Munas. (2022). Common Spelling Mistakes of Hamza and Alif among Beginners. *International Journal of Linguistics, Literature and Translation*, 5(1), 52–58. <https://doi.org/10.32996/ijllt.2022.5.1.7>
- Mohammed, T. A. (2025). Evaluating Translation Quality: A Qualitative and Quantitative Assessment of Machine and LLM-Driven Arabic–English Translations. *Information*, 16(6), Article 440. <https://doi.org/10.3390/info16060440>
- Mohsen, M., Alsudairy, N. A., Alhamami, M., & Al-Hoorie, A. H. (2026). Exploring cognitive Processes in Arabic Dictation: A Study on Writing Challenges among Dyslexic Children. *Acta Psychologica*, 262, 106118. <https://doi.org/10.1016/j.actpsy.2025.106118>
- Mufid, M., Isnainiyah, I., & Ainiy, N. (2023). Contextual Teaching and Learning Model in the Speaking Skill Textbook for the Arabic Language Education Department. *Journal of Advanced Research in Social Sciences and Humanities*, 8(4). <https://doi.org/10.26500/JARSSH-08-2023-0405>
- Muradi, A. (2011). Bahasa Arab dan Pembelajarannya ditinjau dari Berbagai Aspek. In Muhaimin (Ed.), *Angewandte Chemie International Edition*, 6(11), 951–952. (Vol. 3, Issue 1). Pustaka Prisma Yogyakarta. [https://idr.uin-antasari.ac.id/9179/1/Bahasa Arab %26 Pembelajarannya Baru.pdf](https://idr.uin-antasari.ac.id/9179/1/Bahasa%20Pembelajarannya%20Baru.pdf)
- Oto-Millera, N., Pellicer-Ortín, S., & Bustamante, J. C. (2025). Augmented Reality in English Language Acquisition Among Gifted Learners: A Systematic Scoping Review (2020–2025). *Applied Sciences*, 15(21), 11487. <https://doi.org/10.3390/app152111487>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., & Moher, D. (2021). The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *BMJ*, 372, Article n71. <https://doi.org/10.1136/bmj.n71>
- Putri, H. S., Friatin, L. Y., & Tarawana, W. (2025). Using Canva AI Magic Writer to Assist Students Writing Descriptive Text. *Journal of English Education Program (JEEP)*, 12(2), 115. [https://doi.org/10.25157/\(jeep\).v12i2.19344](https://doi.org/10.25157/(jeep).v12i2.19344)
- Qasim, S. S., & Oleiwi, S. H. (2024). Advancing Arabic Handwritten Digit Recognition with AI-Enhanced Neural Network Architectures. *Babylonian Journal of Artificial Intelligence*, 2024, 146–157. <https://doi.org/10.58496/BJAI/2024/016>
- Slavin, R. E. (1994). Quality, Appropriateness, Incentive, and Time: A Model of Instructional Effectiveness. *International Journal of Educational Research*, 21(2), 141–157.

- [https://doi.org/https://doi.org/10.1016/0883-0355\(94\)90029-9](https://doi.org/https://doi.org/10.1016/0883-0355(94)90029-9)
- Taha, H., Ibrahim, R., & Khateb, A. (2014). Exploring the Phenotype of Phonological Reading Disability as a Function of the Phonological Deficit Severity: Evidence from the Error Analysis Paradigm in Arabic. *Reading Psychology*, 35(7), 683–701. <https://doi.org/10.1080/02702711.2013.790325>
- Thohir, M., & Muslimah, K. C. (2020). Evaluation of Arabic Learning Outcomes using Google Form During School Quarantine Due to COVID-19 Pandemic. *Evaluation*, 4(1), 12–22.
- Tibi, S., Fitton, L., & McIlraith, A. L. (2021). The Development of a Measure of Orthographic Knowledge in the Arabic Language: A Psychometric Evaluation. *Applied Psycholinguistics*, 42(3), 739–762. <https://doi.org/10.1017/S014271642100010X>
- Ulfayati, Z. A., Mustaji, & Aisyah, S. (2025). Development of Canva-Based Teaching Materials to Enhance Writing and Short Story Communication Skills. *At Turots: Jurnal Pendidikan Islam*, 870–882. <https://doi.org/10.51468/jpi.v7i2.1172>
- Utami, S., & Karnedi, K. (2024). Enhancing Students' Writing Paragraphs through Canva Magic AI. *Leksika: Jurnal Bahasa, Sastra dan Pengajarannya*, 18(2), 105. <https://doi.org/10.30595/lks.v18i2.23475>
- Wati, M., & Ahmed, B. M. B. (2023). Utilization of the Sahehly Application as a Learning Media for Arabic Writing. *Jurnal Al-Maqayis*, 10(2), 37–47. <https://doi.org/10.18592/jams.v10i2.9511>
- Widmann, M., Apondi, B., Musau, A., Warsame, A. H., Isse, M., Mutiso, V., Veltrup, C., Schalinski, I., Ndeti, D., & Odenwald, M. (2022). Reducing Khat use Among Somalis Living in Kenya: a Controlled Pilot Study on the Modified ASSIST-Linked Brief Intervention Delivered in the Community. *BMC Public Health*, 22(1), 2271. <https://doi.org/10.1186/s12889-022-14681-w>
- William, H. (2003). *The DeLone and McLean Model of Information Systems Success : A Ten-Year Update*. 19, 9–31.
- Windholz, M., & Michalak, R. (2026). Sound Learning: Partnering with Faculty to Expand Expression in the age of AI. *Journal of Library Administration*, 66(1), 1–13. <https://doi.org/10.1080/01930826.2026.2510301>
- Zubaidi, A., Munip, A., Widodo, S. A., & Zerrouki, T. (2025). Enhancing Arabic Writing Skills Using Chat GPT-based AI Learning Models: A Tridimensional Human-AI Collaboration Framework. *Indonesian Journal of Applied Linguistics*, 15(1), 87–101. <https://doi.org/10.17509/ijal.v15i1.75378>