Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Special Issue on CALL Vol. 17, No. 12, July 2026 Pp.3-22
DOI: https://dx.doi.org/10.24093/awej/call12.1

The Effectiveness of AI-Powered Digital Assessment Tools in Enhancing Academic Writing Skills Among EFL University Students

Ali Albashir Mohammed Alhaj
Center for Human and Philosophical Studies, King Khalid University,
Abha, Saudi Arabia
Corresponding Author: alalhaj@kku.edu.sa

Received: 03/18/2026 Accepted:05/15/2026 Published:07/15/2026

Abstract
The rapid development of artificial intelligence (AI) technologies is transforming contemporary educational practices, particularly in terms of language assessment and academic writing instruction. The study aimed to examine the effectiveness of AI-powered digital assessment tools in improving academic writing skills among university students learning English as a Foreign Language (EFL). The study employed a quasi-experimental research design involving two groups of undergraduate students: an experimental group that received instruction supported by AI-based digital assessment tools and a control group that followed traditional writing instruction. To assess students' writing performance, pre-test and post-test writing assessments were administered before and after the instructional intervention. Statistical analyses, including independent-samples and paired-samples t-tests, were conducted to examine differences both between and within the groups. The findings indicated that there were no statistically significant differences between the experimental and control groups in the pre-test, confirming the initial equivalence of the two groups. However, the post-test results demonstrated a statistically significant improvement in the experimental group compared with the control group. Specifically, students who utilized AI-powered digital assessment tools achieved considerably higher mean scores in academic writing performance. In addition, within-group analysis demonstrated a substantial improvement in the experimental group from the pre-test to the post-test, whereas the control group showed only limited progress. These findings suggest that integrating AI-supported feedback and automated writing evaluation systems can effectively enhance students' writing accuracy, organization, and overall academic writing quality. The study underscores the pedagogical value of integrating AI-powered digital assessment tools in enhancing students' academic writing performance while supporting more efficient and objective assessment practices in EFL learning environments.

Keywords: academic writing skills, artificial intelligence, automated writing evaluation, digital assessment, educational technology, English as a foreign language writing

Cite as: Alhaj, A. A. M. (2026).The Effectiveness of AI-Powered Digital Assessment Tools in Enhancing Academic Writing Skills Among EFL University Students. *Arab World English Journal (AWEJ) Special Issue on CALL 17.* (12)3-22. https://dx.doi.org/10.24093/awej/call12.1

Introduction

Education is widely recognized as a central pillar of national development and human capital formation in the Saudi Arabia. In alignment with the strategic vision articulated in Saudi Vision 2030, the Kingdom has undertaken comprehensive educational reforms aimed at enhancing quality and preparing graduates who are competitive in both local and global labour markets (Saudi Vision 2030, 2016). These reforms place strong emphasis on innovation, digital transformation, and the integration of advanced technologies within teaching and learning environments. Within this national framework, the Ministry of Education has introduced a range of initiatives designed to modernize curricula, strengthen teachers' instructional competencies, and embed digital technologies across higher education institutions. Such efforts seek to promote technology-enhanced learning and support the broader national agenda of educational transformation and knowledge-based development (Ministry of Education, 2020).

A particularly influential development in contemporary education is the integration of artificial intelligence (AI) into learning and assessment processes. AI-powered digital assessment tools offer automated evaluation, intelligent feedback, and data-driven insights that enable more effective monitoring of students' academic progress. In English as a Foreign Language (EFL) context, these technologies are especially valuable for academic writing instruction, as AI systems can analyse linguistic patterns, identify grammatical errors, and generate personalized feedback that supports learners in improving the quality of their writing. Consequently, traditional assessment practices have been progressively reshaped by the adoption of AI-supported tools. Unlike conventional paper-based examinations, AI-driven assessments facilitate continuous performance tracking and provide immediate formative feedback, thereby enhancing both skill development and instructional responsiveness. In contrast, traditional approaches tend to prioritize summative evaluation and numerical scoring, which may not adequately capture the developmental nature of writing or offer actionable feedback for improvement. **(**Al-Ghamdi et al.,2026; Alhaj et al.,2026).

Despite the growing adoption of AI-based assessment technologies in higher education, empirical evidence regarding their effectiveness in enhancing academic writing skills among EFL university students remains limited. While some studies report positive outcomes associated with automated feedback and personalized learning support, others underscore challenges related to technological readiness, pedagogical alignment, and teacher training. This divergence in findings highlights the need for further empirical investigation to determine the extent to which AI-powered digital assessment tools can produce measurable improvements in students' writing performance.

Accordingly, this study addresses the following main research question: What is the effectiveness of AI-powered digital assessment tools in enhancing academic writing skills among EFL university students? To operationalize this inquiry, the study examines three subsidiary questions: whether statistically significant differences exist between experimental and control groups in pre-test scores; whether significant differences emerge in post-test scores following the implementation of AI-powered tools; and whether meaningful improvement occurs within each group between pre-test and post-test measures.

The study is guided by four primary objectives: to investigate the effectiveness of AI-powered digital assessment tools in improving academic writing skills among EFL university students; to establish the equivalence of experimental and control groups prior to the intervention through pre-test comparison; to evaluate the impact of AI-based assessment tools on writing performance by

analysing post-test outcomes; and to assess the magnitude of improvement in students' writing skills over time.

This study contributes to the expanding body of literature on AI-supported language learning and digital assessment in several important ways. It provides empirical evidence on the role of AI-powered assessment tools in enhancing academic writing skills among EFL learners, offers insights for policymakers and institutional leaders regarding effective strategies for integrating AI into language education, and presents practical implications for instructors seeking to improve feedback quality, learning outcomes, and instructional effectiveness in higher education contexts.

Literature Review

Artificial Intelligence and the Transformation of Contemporary Education

AI has rapidly emerged as a transformative force in contemporary education, reshaping teaching, learning, and assessment practices across diverse educational contexts. The rapid evolution of digital technologies has enabled AI to support innovative instructional approaches and data-driven decision-making within educational institutions. Through machine learning algorithms and learning analytics, AI can analyse students' learning behaviours, monitor academic progress, and generate valuable insights that assist educators in designing more effective teaching strategies. These capabilities allow instructors to better understand learners' needs and adapt instructional practices accordingly. As a result, AI contributes to enhancing the quality, operational efficiency, and overall impact of educational processes while supporting the ongoing digital transformation of higher education institutions (Altamimi, 2025; Holmes et al., 2019)

One of the most significant contributions of artificial intelligence in education lies in its capacity to enhance teaching and learning processes by facilitating personalized learning environments tailored to students' individual needs and learning preferences. As such, AI-powered educational systems can adapt instructional materials, learning pathways, and assessment activities according to individual learners' abilities, needs, and learning styles. This adaptive capacity enables students to receive immediate, targeted feedback and customized learning experiences that enhance their engagement and academic performance. In addition, AI-driven tools support automated assessment, intelligent tutoring systems, and predictive learning analytics that assist instructors in monitoring student progress and identifying learning difficulties at an early stage. Consequently, AI technologies play a crucial role in creating more responsive, learner-cantered, and effective educational environments (Luckin et al., 2022).

Digital Assessment and E-Assessment in Higher Education

Digital assessment, commonly referred to as e-assessment, has become a central component of contemporary higher education as institutions increasingly adopt technology-enhanced learning environments. E-assessment involves the application of digital technologies to design, administer, evaluate, and provide feedback on students' academic performance. Compared with traditional paper-based examinations, digital assessment systems offer several advantages, including automated scoring, immediate feedback, and continuous monitoring of learners' progress. These systems support diverse assessment formats such as online quizzes, interactive tasks, and adaptive tests, all of which enable instructors to measure students' knowledge and skills more efficiently and accurately. Consequently, digital assessment improves the reliability, flexibility, and effectiveness of evaluation processes while simultaneously reducing the

administrative workload typically associated with conventional assessment methods (Aljuaid, 2024; Bennett, 2015; Redecker & Johannessen, 2013)

The growing integration of digital technologies in higher education has significantly enhanced the potential and effectiveness of electronic assessment (e-assessment) systems. Indeed, today's digital assessment platforms now incorporate learning analytics and artificial intelligence tools that allow educators to analyse students' responses, identify learning gaps, and provide timely instructional support. In addition, these systems facilitate formative assessment by delivering immediate and personalized feedback, thereby helping students gradually improve their academic performance and develop essential learning skills. In language education, particularly in (EFL) context, digital assessment tools play a significant role in evaluating writing, reading, listening, and speaking skills through automated and data-driven approaches. As a result, e-assessment is increasingly recognized as an effective strategy for improving assessment quality, enhancing learning outcomes, and supporting the ongoing digital transformation of higher education institutions (Nicol & Macfarlane, 2006).

AI-Powered Assessment Tools in Language Learning

AI has increasingly been integrated into language assessment through the integration of advanced technologies, particularly machine learning and natural language processing (NLP)and automated scoring systems, transforming traditional evaluation practices by enabling the development of intelligent assessment tools capable of analysing learners' linguistic performance with greater accuracy and efficiency. AI-powered systems can assess multiple aspects of language proficiency, including grammatical accuracy, vocabulary use, coherence, and syntactic complexity, and provide detailed feedback that supports learners' language development (Chapelle & Voss, 2016; Shermis & Burstein, 2013) In EFL teaching and learning contexts, automated writing evaluation systems and AI-based assessment platforms allow instructors to evaluate large numbers of students more efficiently while maintaining consistency and reliability in scoring. As a result, AI-based assessment tools contribute to more objective, scalable, and data-driven evaluation practices in language education.

Furthermore, AI-driven assessment technologies support both formative and summative evaluation by providing immediate and personalized feedback enabling learners to identify their strengths and areas requiring further development. Machine learning algorithms can analyse learners' responses and adjust assessment tasks according to individual proficiency levels, thereby creating adaptive testing environments that more accurately reflect students' language abilities (Luckin et al., 2022). Moreover, natural language processing technologies enable automated analysis of both spoken and written language, thus facilitating the assessment of pronunciation, fluency, and discourse organization. These innovations enhance the reliability and validity of language assessment while reducing teachers' administrative workload and supporting more data-informed instructional decision-making in language learning environments (Holmes et al., 2019; Redecker & Johannessen, 2013).

Academic Writing Skills in EFL Contexts

Academic writing is widely regarded as a central component of (EFL) education, particularly within higher education institutions where students are required to produce structured, coherent, and evidence-based academic texts. However, effective academic writing extends beyond linguistic competence, requiring the ability to organize ideas logically and coherently,

construct well-supported arguments, and adhere to appropriate academic conventions. In university contexts, these skills enable students to engage in scholarly communication, complete research-based assignments, and present ideas in a systematic and academically acceptable manner. Nevertheless, many EFL learners encounter significant challenges in developing academic writing proficiency due to limited exposure to academic discourse, restricted vocabulary resources, and difficulties with grammatical accuracy and syntactic complexity (Hyland, 2019). Consequently, strengthening students' academic writing competence remains a key objective in EFL teaching and curriculum development.

Despite its acknowledged importance, many university students studying (EFL) continue to experience difficulties with various aspects of academic writing. Common challenges include organizing ideas coherently, maintaining clarity in argumentative writing, using appropriate academic vocabulary, and applying accurate grammatical structures. In addition, EFL learners often experience difficulties in understanding citation practices, avoiding plagiarism, and adhering to the disciplinary writing conventions required in academic contexts (Hyland, 2019). Limited opportunities for guided practice and constructive feedback may further hinder the development of students' writing skills. As a result, higher education institutions are increasingly adopting innovative instructional approaches and digital learning tools designed to support academic writing development and provide continuous feedback that enhances students' writing performance and overall academic achievement (Flowerdew, 2025; Graham, 2019).

AI-Based Feedback Systems for Improving Academic Writing

AI–driven feedback systems are increasingly integrated into language learning environments to facilitate the development of students' academic writing proficiency. These intelligent writing tools leverage natural language processing and machine learning technologies to analyze written texts and provide automated feedback on grammatical accuracy, sentence structure, vocabulary use, and overall textual coherence. By detecting linguistic errors and offering immediate corrective feedback, AI-powered systems enable learners to revise and refine their writing more effectively. Within EFL learning contexts, such feedback is particularly beneficial, as many learners face challenges related to grammatical accuracy, academic vocabulary, and sentence organization due to limited exposure to academic discourse and academic writing conventions. Through timely and personalized feedback, AI-powered writing tools are able to help learners recognize weaknesses in their writing and progressively enhance their language proficiency and academic communication skills (Hyland, 2019).

Furthermore, AI-based feedback systems play a significant role in strengthening formative assessment practices in writing instruction. Unlike traditional teacher-cantered feedback, which may be delayed because of heavy instructional workloads, AI-supported tools provide immediate responses thus allowing students to revise their drafts continuously during the writing process. This interactive feedback cycle encourages learner autonomy, promotes reflective learning, and helps students develop stronger self-editing and revision skills. Empirical research also indicates that automated writing evaluation systems can improve the overall quality of students' academic texts by supporting better organization, coherence, and lexical diversity (Holmes et al., 2019; Shermis & Burstein, 2013). Consequently, integrating AI-driven feedback technologies in higher education offers significant pedagogical benefits by supporting instructors, improving the efficiency of writing assessment, and enhancing students' academic writing performance.

Previous Studies

Recent advances in artificial intelligence (AI) and automated writing evaluation (AWE) technologies have significantly reshaped academic writing instruction in EFL contexts.AI-driven digital assessment tools provide automated feedback that enables learners to detect linguistic errors, revise their texts efficiently, and enhance their overall academic writing performance. Research suggests that these tools contribute to the development of grammatical accuracy, revision strategies, and learner engagement. For instance, Wei et al. (2023) found that students who used AI-based writing feedback tools demonstrated significant improvements in their academic writing performance. receiving AI-supported feedback demonstrated significantly better performance across multiple writing components compared to those receiving traditional instruction. Their study also highlighted that learners' initial proficiency levels and writing self-efficacy significantly influenced outcomes in terms of task achievement, coherence and cohesion, and lexical development.

Similarly, previous studies have highlighted the pedagogical value of AWE tools in enhancing EFL writing instruction. Aldosemani et al. (2023) reported that automated feedback systems provide timely and meaningful responses to students' writing, which helps learners revise their work more effectively and independently. These tools also contribute to reducing teachers' workload by automating certain aspects of writing assessment. Their effectiveness, however, largely depends on learners' engagement with the feedback and their level of language proficiency. Likewise, Karatay and Karatay (2024) found that AWE systems positively influence L2 writing practices and contribute to improvements in writing quality. Despite these benefits, however, the authors noted that students' interaction with AWE tools is influenced by contextual and individual factors, indicating that automated feedback should complement rather than replace teacher guidance.

Experimental research has further highlighted the effectiveness of AI-assisted feedback tools in improving EFL writing performance. Mekheimer (2025) found that students who used Grammarly as an AI-supported feedback tool achieved significantly higher writing scores than those who relied on traditional teacher feedback. The use of such AI features was found to encourage frequent revisions and improvements in content organization and cohesion, while also increasing students' motivation and confidence in writing. Similarly, Wijaya (2025) found that Grammarly enhances writing accuracy and linguistic proficiency through automated corrective feedback. However, these systems are generally more effective in addressing surface-level language issues, such as grammar and mechanics, than in supporting higher-order writing skills, including idea development and coherence. (Karagöz, 2025).

Recent empirical studies further confirm the positive impact of AI-based feedback on EFL writing development. Rahimi et al. (2025) found that learners who received automated written corrective feedback significantly outperformed those who received traditional feedback in terms of writing performance, task achievement, and grammatical accuracy. The study also reported that students demonstrated high levels of behavioural, cognitive, and emotional engagement when interacting with electronic feedback systems. Similarly, Dewi (2022) reported that EFL learners generally hold positive perceptions of Grammarly as an automated writing evaluation (AWE) tool, highlighting its effectiveness in identifying grammatical errors and enhancing vocabulary use. Likewise, Thao et al. (2025) and Murtisari et al. (2025) found that AI-based feedback systems significantly improved learners' writing accuracy, linguistic complexity, and overall fluency. Nevertheless, students have emphasized the continued importance of teacher interaction and

pedagogical support, suggesting that the most effective approach to writing instruction involves integrating AI feedback with human guidance.

Methodology

The current research implemented a quasi-experimental approach to assess the influence of digital assessment tools on students' academic outcomes. This methodological framework is commonly adopted in educational research when the random assignment of participants is not practical, while still allowing for structured comparisons between groups within controlled research settings (Creswell & Creswell, 2018). The study involved two groups: an experimental group, which received instruction supported by digital assessment tools, and a control group, which followed traditional instructional methods. To evaluate the impact of the intervention, pre-test and post-test assessments were administered to both groups. The pre-test measured students' baseline academic performance, while the post-test assessed learning outcomes after the instructional treatment. This approach enabled the researcher to determine changes in students' achievement resulting from the integration of digital assessment tools. Quasi-experimental designs are particularly suitable for classroom-based studies because they allow researchers to investigate causal relationships while maintaining authentic educational settings (Fraenkel et al., 2019). Therefore, this design provided a systematic framework for evaluating the influence of digital assessment tools on students' academic achievement.

Research Population and Sample

The study population was composed of male and female first-level English Language students enrolled at King Khalid University during the academic year 1445–1446 AH. A sample of students was selected from this population to participate in the study it included 80 first-level English Language students divided into two groups: the experimental group, which received instruction supported by digital assessment tools, consisted of 40 students, while the control group, which followed traditional instructional methods, also included 40 students. The use of two comparable groups allowed for a systematic comparison of students' academic performance before and after the intervention. Such group-based sampling procedures are commonly employed in quasi-experimental educational research to ensure balanced comparisons and reliable analysis of instructional effects (Creswell & Creswell, 2018; Fraenkel et al., 2019). Table 1 shows the distribution of the study participants.

Table 1. *Distribution of the research sample across the experimental and control groups*

Group	College	Number of Students	Total Sample Size
Experimental	Applied College in Dhahran Al-Janoub	40 students	80 students
Control	Applied College in Dhahran Al-Janoub	40 students	

In accordance with the objectives of this study, a set of research hypotheses was formulated to examine the effect of digital assessment tools on students' academic achievement. These hypotheses were designed to determine whether the integration of digital assessment tools contributes to measurable improvements in learners' performance compared with traditional assessment approaches.

Research Hypotheses

To evaluate the effectiveness of AI-powered digital assessment tools in improving students' academic writing performance, the study tested several hypotheses. First, it was hypothesized that a statistically significant difference would exist between the experimental and control groups in the post-test scores of the achievement test in Favor of the experimental group. Second, no statistically significant difference was expected between the two groups in the pre-test, indicating their equivalence prior to the instructional intervention. Third, any significant differences observed in the post-test were assumed to be attributable to the use of digital assessment tools integrated into the experimental treatment. Finally, it was hypothesized that both groups would demonstrate improvement between the pre-test and post-test administrations, reflecting the overall impact of the instructional period. "These hypotheses align with established experimental procedures commonly used in educational research to evaluate the effectiveness of instructional interventions." (Creswell & Creswell, 2018; Fraenkel et al., 2019).

Research Instruments

Achievement Test

Widely used in educational research to measure students' mastery of specific learning outcomes and to evaluate the effectiveness of instructional interventions, the study employed an achievement test as the primary research instrument to assess students' academic writing performance. Achievement tests are widely used in educational research to measure students' mastery of specific learning outcomes and to evaluate the effectiveness of instructional interventions (Brown & Abeywickrama, 2019). The test was carefully designed through a series of systematic procedures to ensure its validity, clarity, and suitability for measuring the targeted writing skills among university students learning English as a Foreign Language (EFL).

Purpose of the Achievement Test

The achievement test aimed to measure students' academic performance in academic writing skills, with particular emphasis on three essential dimensions: writing accuracy, linguistic complexity, and writing fluency. These components were selected because they represent core indicators of academic writing proficiency and are widely recognized in second language writing research as critical dimensions of writing quality (Graham, 2019; Hyland, 2019). Assessing these dimensions allows researchers to evaluate both linguistic correctness and the sophistication of students' written production.

Development of the Test

The preliminary version of the test began with an introductory statement addressed to students, which was read aloud by the course instructor prior to administering the assessment. The statement clarified the purpose of the test and informed students that the results would not influence their course grades but would be used exclusively for scientific research purposes and for improving teaching and learning practices. Such procedures help ensure ethical transparency and reduce students' anxiety during testing (Creswell & Creswell, 2018).

Students were instructed to answer all questions carefully and to record their responses on the designated answer sheet according to the example provided. The test consisted of 20 multiple-choice items, each designed to measure specific aspects of academic writing performance, including grammar accuracy, vocabulary usage, sentence structure, and coherence.

Test Validity

To establish content validity, the test instrument was reviewed by a panel of experts specializing in linguistics and English language teaching methodologies. Expert validation is an essential procedure in educational measurement to ensure that test items adequately represent the intended learning objectives (Brown & Abeywickrama, 2019).

The reviewers evaluated the test items based on several criteria:

1. The degree to which each item accurately measured its intended learning objective.
2. The alignment of the test items with appropriate levels of Bloom's taxonomy.
3. The clarity and comprehensibility of the item wording.
4. The linguistic accuracy and appropriateness of the test content.

Based on the feedback provided by the experts, several items were subsequently revised and refined to enhance clarity, linguistic accuracy, and suitability for the targeted group of learners. These revisions ensured that the instrument accurately reflected the learning outcomes related to academic writing skills.

Digital Testing Platform

In alignment with the objectives of the study and the integration of digital assessment tools, the researcher utilized Microsoft Forms to administer both the pre-test and post-test electronically. Digital assessment platforms have become increasingly important in modern education because they allow automated scoring, rapid feedback, and efficient data management (Bennett, 2015; Redecker & Johannessen, 2013).

Microsoft Forms was selected for several practical advantages, including its ease of use, ability to control response timing, automated scoring functions, and efficient data storage. Additionally, the platform allows test results to be exported directly into Microsoft Excel, facilitating accurate statistical analysis. Using a single digital platform also ensured consistency in the assessment environment and minimized potential variability that might arise from using multiple digital assessment systems.

Implementation of the Research Experiment

The achievement test was administered as a pre-test to both the experimental and control groups during the first semester of the 1445–1446 AH academic year to establish baseline equivalence prior to the experimental intervention.

The pre-test results were analysed using descriptive statistics and an independent samples *t*-test to examine whether any significant differences existed between the two groups before the intervention. Establishing baseline equivalence between groups is a fundamental requirement in quasi-experimental research designs, as it helps ensure that any subsequent differences in outcomes can be attributed to the experimental treatment rather than pre-existing disparities between participants (Creswell & Creswell, 2018).

Table 2 presents the means, standard deviations, *t*-value, degrees of freedom, and significance level for both groups in the pre-administration of the achievement test. The statistical analysis revealed no statistically significant differences between the experimental and control groups in the pre-test results, indicating that the two groups were statistically equivalent prior to the implementation of the experimental intervention.

Table 2. *Pre-test results of the achievement test: means, standard deviations, t-test, and significance levels*

Group	N	Mean	SD	*t*-value	df	Sig. (p)
Control	40	10.94	4.75	-1.26	78	0.21
Experimental	40	12.42	5.51			

Table 2 presents the results of an independent samples *t*-test comparing the pre-test scores of the experimental and control groups. The control group (N = 40) obtained a mean score of 10.94 (SD = 4.75), while the experimental group (N = 40) achieved a slightly higher mean of 12.42 (SD = 5.51). However, the statistical analysis revealed that this difference was not statistically significant with $t(78) = -1.26$, $p = .21$, which exceeds the conventional significance level of $\alpha \leq .05$. Therefore, the null hypothesis of no difference between the two groups cannot be rejected. These findings indicate that both groups demonstrated comparable levels of academic achievement prior to the experimental treatment. Establishing such baseline equivalence is essential in quasi-experimental research, as it strengthens the internal validity of the study and ensures that any subsequent differences observed in post-test performance can be attributed to the instructional intervention rather than pre-existing differences between participants (Creswell & Creswell, 2018; Field, 2018).

Reliability of the Research Instrument

To verify the reliability of the achievement test, Cronbach's alpha coefficient was calculated to examine the internal consistency among test items. This statistic is extensively utilized in educational research to determine the degree to which items within a scale consistently measure the same latent construct. (Tavakol & Dennick, 2011).

The results showed that the achievement test had a Cronbach's alpha of $\alpha = 0.87$, indicating high internal consistency. Generally, alpha values above 0.70 are considered acceptable, and values above 0.80 are regarded as good (Field, 2018). Therefore, the obtained coefficient demonstrates that the instrument was sufficiently reliable for measuring students' academic writing performance.

Table 3. *Reliability of the achievement test*

Instrument	Number of Items	0.87
Achievement Test	20	0.87

Statistical Analysis Procedures

The collected data were analysed using the Statistical Package for the Social Sciences (SPSS), a widely recognized statistical software used in educational and social science research for quantitative data analysis (Field, 2018). Descriptive statistical techniques were first utilized to summarize and interpret the research data and included percentages, means, and standard deviations, providing an overall description of students' academic performance and illustrating the distribution of scores across the study groups.

To determine whether statistically significant differences existed between the experimental and control groups, independent samples *t*-tests were conducted, which enabled the comparison of mean scores between the two groups in both the pre-test and post-test measurements. The statistical significance of the results was evaluated at the $\alpha \leq .05$ level, which is widely accepted in educational and behavioural research as the threshold for determining meaningful differences between groups (Creswell & Creswell, 2018).

To ensure reliability, Cronbach's alpha was calculated to assess the internal consistency of the test items. This measure indicates how well the items reflect the same construct and produce consistent results (Tavakol & Dennick, 2011). Additionally, split-half reliability was used to examine the consistency of the instrument by correlating scores from two halves of the test.

Data Analysis

Statistical analyses were performed using SPSS version 25 to evaluate the effectiveness of digital assessment tools in enhancing students' academic writing outcomes. Initially, descriptive statistics—including means, standard deviations, and percentages—were calculated to summarize score distributions and provide an overall picture of students' performance across the study groups. To examine differences between the experimental and control groups, independent-samples t-tests were conducted on both pre-test and post-test scores. In addition, paired-samples t-tests were used to assess the degree of improvement within each group by comparing their performance before and after the intervention. Statistical significance was set at $\alpha \leq .05$, a commonly accepted threshold in educational and social science research for determining meaningful differences between groups. (Creswell & Creswell, 2018). The reliability of the research instrument was assessed using Cronbach's alpha coefficient to measure the internal consistency of the test items, while the split-half reliability method was also applied to confirm the stability of the instrument. Furthermore, the magnitude of the treatment effect was estimated using Eta squared (η^2) to determine the practical significance of the instructional intervention. These statistical procedures are commonly recommended in quantitative educational research to enhance the reliability, validity, and interpretability of research findings (Field, 2018; Tavakol & Dennick, 2011).

Ethical Considerations

Ethical considerations were strictly observed to ensure compliance with established research standards. Participants were clearly informed about the study's purpose, procedures, and their role prior to data collection. Participation was entirely voluntary, and students were assured that their involvement would not affect their academic grades or evaluations. Informed consent was obtained from all participants before administering the research instruments. Confidentiality and anonymity were fully maintained, with no identifying information collected, and all data were securely stored for academic use only. Participants were also informed of their right to withdraw at any time without penalty, in line with recognized ethical guidelines (Creswell & Creswell, 2018; Israel & Hay, 2006).

Results and Findings

Results Related to the Main Research Question

To address the main research question— "What is the effectiveness of AI-powered digital assessment tools in enhancing academic writing skills among EFL university students?"—the corresponding research hypothesis was statistically examined. The hypothesis proposed that a statistically significant difference would exist at the significance level ($\alpha \leq .05$) between the mean post-test scores of the experimental and control groups, favouring the experimental group.

To test this hypothesis, both descriptive and inferential statistical analyses were conducted. Descriptive statistics, including arithmetic means and standard deviations, were calculated to summarize the performance of students in both groups. Subsequently, an independent samples *t*-test was performed to determine whether the observed differences between the groups were

statistically significant. In addition, Eta squared (η^2) was calculated to estimate the effect size of the instructional intervention and to determine the magnitude of the impact of AI-powered digital assessment tools on students' academic writing performance (Field, 2018; Pallant, 2020). The results of these analyses are presented in Table 4.

Table 4: *Independent-Samples* **t**-*Test Results for Post-Test Scores of the Experimental and Control Groups*

Group	N	Mean	SD	*t*-value	df	Sig. (p)
Control	40	8.91	2.79	-11.24	78	< .001
Experimental	40	17.65	1.49			

Table 4 presents the results of the independent samples *t*-test conducted to compare the post-test performance of students in the experimental and control groups. The analysis revealed a statistically significant difference between the two groups at the conventional significance level ($\alpha \leq .05$). Specifically, the experimental group (N = 40) achieved a substantially higher mean score (M = 17.65, SD = 1.49) than the control group (N = 40), which recorded a mean score of (M = 8.91, SD = 2.79). The obtained *t*-value, $t(78) = -11.24$, $p < .001$, indicates a highly significant difference in favor of the experimental group.

These findings indicate that the instructional intervention applied to the experimental group had a substantial and positive impact on students' academic writing performance in the post-test. In practical terms, students who engaged with AI-powered digital assessment tools demonstrated significantly higher levels of achievement compared with those who received traditional instruction. Such outcomes align with the principles of experimental research in education, where statistically significant post-test differences between groups are generally interpreted as evidence of the effectiveness of the implemented instructional strategy (Creswell & Creswell, 2018).

Furthermore, the considerable difference between the mean scores of the two groups highlights the practical educational significance of the intervention in enhancing students' learning outcomes. These results support the study's hypothesis that integrating innovative digital assessment tools into instructional practices can significantly improve students' academic writing skills. Similar findings have been reported in prior research, thereby emphasizing the positive role of technology-enhanced learning environments in promoting student achievement, engagement, and assessment effectiveness (Graham, 2019; Pallant, 2020). Therefore, the results provide strong empirical evidence supporting the effectiveness of AI-powered digital assessment tools in enhancing academic writing performance among EFL university students.

Results Related to Answering the First Sub-Research Question

The first sub-research question examined whether statistically significant differences existed between the experimental and control groups in the pre-administration of the achievement test. Addressing this question is essential for verifying the initial equivalence of the two groups prior to implementing the experimental intervention. Establishing baseline comparability represents a fundamental methodological requirement in quasi-experimental research designs, as it ensures that any subsequent differences in learning outcomes can be attributed to the instructional treatment rather than to pre-existing variations among participants. When both groups demonstrate comparable performance during the pre-test stage, the internal validity of the study is strengthened, and the research design becomes more methodologically robust. Consequently, confirming group equivalence provides a reliable foundation for interpreting the post-test results

and evaluating the effectiveness of the experimental intervention (Creswell & Creswell, 2018; Fraenkel et al., 2019).

To examine differences between the two groups, an independent samples *t*-test was conducted. The results of the statistical analysis are presented in Table 5.

Table 5. Independent-samples t*-test results for post-test scores*

t-value (Post-test)	Degrees of Freedom (df)	Significance Level (p)
-11.24	78	< .001

Table 5 presents the results of the independent samples *t*-test conducted to compare the post-test scores of the experimental and control groups. The statistical analysis revealed a highly significant difference between the two groups, $t(78) = -11.24$, $p < .001$. Because the obtained *p*-value is substantially lower than the conventional significance threshold of $\alpha \leq .05$, the null hypothesis indicating no difference between the groups is rejected.

These findings indicate that the instructional intervention had a significant effect on students' academic writing performance. The large negative *t*-value suggests that the experimental group achieved significantly higher mean scores than the control group in the post-test. This result provides strong empirical evidence supporting the effectiveness of the digital assessment tools used in the study in improving students' academic achievement. In experimental and quasi-experimental educational research, statistically significant post-test differences between groups are generally interpreted as evidence that the implemented instructional strategy contributed to the observed improvement in learning outcomes (Creswell & Creswell, 2018; Field, 2018). Therefore, the results reported in Table 5 confirm the positive impact of the instructional intervention applied to the experimental group.

Results Related to the Second Sub-Research Question

The second sub-research question examined whether statistically significant differences existed between the experimental and control groups in the post-test scores following the implementation of AI-powered digital assessment tools. This question aimed to evaluate the effectiveness of the experimental intervention by comparing the academic performance of both groups after the completion of the instructional treatment period. In experimental and quasi-experimental research designs, comparing post-test scores between groups represents a key methodological procedure for determining the impact of a particular instructional strategy or technological intervention. Such comparisons help establish whether improvements in students' learning outcomes can be attributed to the applied treatment rather than to pre-existing differences among participants (Creswell & Creswell, 2018; Fraenkel et al., 2019). Accordingly, an independent-samples *t*-test was conducted to compare the mean post-test scores of the experimental and control groups.

Table 6. Independent-samples **t***-test results for post-test scores*

Group	N	Mean	SD	*t*	df	P
Control	40	8.91	2.79	-11.24	78	< .001
Experimental	40	17.65	1.49			

Note. $t(78) = -11.24$, $p < .001$.

The statistical analysis revealed a highly significant difference between the two groups, $t(78) = -11.24$, $p < .001$. Because the obtained probability value is substantially lower than the conventional significance level of $\alpha \leq .05$, the null hypothesis indicating no difference between the

groups is rejected. The findings show that the experimental group (M = 17.65, SD = 1.49) achieved markedly higher mean scores than the control group (M = 8.91, SD = 2.79). This outcome indicates that the instructional intervention—implemented through AI-powered digital assessment tools—had a strong positive impact on students' academic writing performance. In experimental educational research, statistically significant differences in post-test scores typically indicate that the observed improvement can be attributed to the implemented intervention rather than to random variation or pre-existing differences among participants (Creswell & Creswell, 2018; Fraenkel et al., 2019). Furthermore, the large magnitude of the obtained *t*-value reflects a substantial disparity between the mean scores of the two groups, highlighting the effectiveness of the AI-supported digital assessment approach. These findings therefore provide strong empirical support for the study's hypothesis that AI-powered digital assessment tools can significantly enhance academic writing skills among EFL university students. Consequently, integrating AI-based assessment technologies into language instruction may improve students' writing proficiency and overall learning outcomes (Field, 2018).

Results Related to the Third Sub-Research Question

The third sub-research question examined whether a statistically significant improvement occurred in students' academic writing skills within each group between the pre-test and post-test administrations of the achievement test. Investigating changes within the same group over time represents an important analytical procedure in experimental and quasi-experimental research designs as it allows researchers to determine whether students' performance improves following an instructional intervention (Creswell & Creswell, 2018; Fraenkel et al., 2019). In the present study, this analysis aimed to assess the extent to which students' academic writing abilities developed during the instructional period.

To address this question, a paired-samples *t*-test was conducted for the experimental and control groups to compare their mean scores in the pre-test and post-test. This statistical procedure is commonly employed to evaluate whether the difference between two related measurements is statistically significant, thereby providing empirical evidence regarding the effectiveness of an instructional treatment (Field, 2018). In this study, data analysis was employed to determine the impact of implementing AI-powered digital assessment tools on students' academic writing performance.

A paired-samples t-test was conducted to assess whether the experimental group demonstrated a statistically significant improvement in scores from pre-test to post-test. The results indicate a clear improvement in students' performance following the instructional intervention. As presented in Table 7, the mean score increased from 12.44 (SD = 5.52) in the pre-test to 17.77 (SD = 1.56) in the post-test, with a mean difference of 5.33 points. This substantial increase suggests that the use of AI-powered digital assessment tools contributed to a notable improvement in students' academic writing performance.

Table 7. *Paired-samples t-test results for the experimental group (pre-test vs. post-test)*

Test Administration	N	Mean	SD
Pre-test	40	12.44	5.52
Post-test	40	17.77	1.56

A paired-samples *t*-test was also conducted for the control group to examine whether any improvement occurred between the pre-test and post-test administrations without the experimental

intervention. As shown in Table 8, the mean score increased from 8.91 (SD = 2.83) in the pre-test to 10.98 (SD = 4.78) in the post-test, with a mean difference of 2.07 points. Although this result indicates a modest improvement in students' performance, the magnitude of change is considerably smaller than that observed in the experimental group, suggesting that the instructional intervention played a significant role in enhancing students' academic writing skills.

Table 8. *Paired-samples t-test results for the control group (pre-test vs. post-test)*

Test Administration	N	Mean	SD
Pre-test	40	8.91	2.83
Post-test	40	10.98	4.78

Discussion

The present study aimed to examine the effectiveness of AI-powered digital assessment tools in enhancing academic writing skills among EFL university students. Using a quasi-experimental research design, it compared the writing performance of students in both experimental and control groups across pre-test and post-test administrations of the achievement test. The results revealed statistically significant differences in favour of the experimental group, indicating that the integration of AI-supported digital assessment tools contributed to substantial improvements in students' academic writing performance. These findings highlight the growing pedagogical relevance of artificial intelligence technologies in language education, particularly within technology-enhanced learning environments. As higher education institutions increasingly adopt digital and AI-based instructional approaches, evaluating the effectiveness of AI-mediated feedback and assessment has become essential for improving teaching practices and student learning outcomes. In line with principles of experimental educational research, statistically significant post-test differences between groups are typically interpreted as evidence of the effectiveness of the implemented instructional intervention (Alhaj, 2024; Creswell & Creswell, 2018; Fraenkel et al., 2019)

The findings of the present study demonstrate that students who engaged with AI-powered digital assessment tools achieved significantly higher writing scores compared with those who received traditional instruction. The independent-samples *t*-test confirmed a highly significant difference between the two groups. In contrast, the paired-samples *t*-test revealed a notable improvement in the experimental group between the pre-test and post-test administrations. Although the control group showed a modest improvement, the magnitude of change was considerably smaller than that observed in the experimental group. This suggests that AI-mediated feedback and automated evaluation systems play an important role in supporting students' writing development. AI-assisted writing tools can provide immediate, individualized feedback thereby helping learners to identify linguistic errors, revise their drafts, and improve the overall quality of their writing. Previous research has similarly shown that automated writing evaluation systems and generative AI tools can enhance writing accuracy, revision behaviour, and learning engagement among EFL learners (Polakova, 2024).

The results of this study are consistent with an expanding body of literature highlighting the positive impact of artificial intelligence technologies on language learning and writing instruction. Recent studies have demonstrated that AI-supported writing tools—such as automated writing evaluation systems and generative AI assistants—can improve students' grammatical accuracy, textual coherence, and revision strategies. These technologies provide continuous formative

feedback enabling students to engage more actively in the writing process and refine their drafts iteratively. Research on generative AI in language education has also shown that AI-mediated feedback can enhance learner autonomy and promote more effective writing practices by encouraging students to analyse and revise their own work. Moreover, studies examining AI-based feedback systems indicate that such technologies can also increase students' motivation and engagement by providing real-time guidance during the writing process, thereby facilitating deeper learning and improved academic writing outcomes.

From a theoretical and pedagogical perspective, the improvement observed in the experimental group can be interpreted through the principles of technology-enhanced learning and formative assessment theory. AI-powered digital assessment tools function as adaptive learning systems that provide immediate and individualized feedback, thereby establishing a continuous feedback loop that supports self-regulated learning. This process enables learners to monitor their progress, identify linguistic weaknesses, and refine their writing strategies in real time. These findings are consistent with recent research highlighting the effectiveness of automated writing evaluation (AWE) systems in enhancing learners' engagement and writing performance (Hyland, 2022; Li, 2023). However, in contrast to earlier studies that primarily emphasize improvements in surface-level linguistic accuracy, the present study extends the literature by demonstrating that AI-mediated feedback can also foster higher-order writing skills, including coherence and argument development. This added value may be attributed to the integration of advanced AI-driven feedback mechanisms that promote deeper cognitive engagement and reflective learning.

Furthermore, the findings are strongly aligned with formative assessment theory, which underscores the importance of timely, specific, and actionable feedback in facilitating meaningful learning (Black & Wiliam, 2009). They also support the growing body of literature on data-driven and personalized learning environments, where AI technologies enable more responsive and adaptive instructional practices (Holmes et al., 2019; Luckin et al., 2022). In this regard, AI-mediated assessment environments allow instructors to complement traditional pedagogical approaches with analytics-based insights, thereby enhancing instructional decision-making and supporting differentiated learning pathways. From a theoretical perspective, these results also resonate with technology adoption frameworks, particularly in relation to performance expectancy and facilitating conditions, which are critical factors influencing the effective integration of educational technologies.

Nevertheless, several limitations should be acknowledged. The relatively small sample size and the short duration of the intervention may limit the generalizability of the findings across diverse educational contexts. Compared with large-scale and longitudinal investigations, the present study offers preliminary yet context-specific insights within an EFL setting, which should be interpreted with caution. Future research is therefore recommended to include larger and more diverse samples and to adopt longitudinal designs in order to examine the sustained impact of AI-powered digital assessment tools on students' writing development over time. In addition, qualitative approaches, such as interviews and think-aloud protocols, could provide deeper insights into learners' perceptions, engagement patterns, and cognitive processes during AI-assisted writing activities (Creswell & Creswell, 2018; Field, 2018).

Conclusion

This study investigated the effectiveness of AI-powered digital assessment tools in enhancing the academic writing skills of EFL university students within a quasi-experimental

research framework. By comparing the performance of students in experimental and control groups across pre-test and post-test administrations of an achievement test, the study aimed to determine whether the integration of AI-supported assessment technologies could significantly improve writing performance. The findings revealed clear and statistically significant improvements in the academic writing abilities of students who engaged with AI-based digital assessment tools compared with those who received traditional instruction. Students in the experimental group demonstrated substantially higher post-test scores and greater progress between the pre-test and post-test administrations, indicating that AI-mediated feedback and automated assessment mechanisms played a meaningful role in supporting writing development. These results suggest that AI-powered technologies can provide valuable pedagogical support by offering immediate, individualized, and data-driven feedback, thereby facilitating students in identifying linguistic errors, revising their drafts, and refining the organization and clarity of their academic writing. Such feedback mechanisms encourage iterative learning and foster greater learner autonomy by allowing students to monitor and improve their writing performance independently. The findings are consistent with recent research highlighting the potential of artificial intelligence tools to enhance writing accuracy, learner engagement, and revision practices in EFL contexts. In addition, AI-based digital assessment environments may contribute to more efficient and responsive assessment processes, supporting both instructors and students in achieving improved learning outcomes. Overall, the results of this study provide empirical evidence that integrating AI-powered digital assessment tools into language instruction can significantly enhance students' academic writing performance in higher education settings. Nevertheless, the use of AI technologies in educational contexts should be implemented thoughtfully and ethically, ensuring that automated systems complement rather than replace human pedagogical guidance. Future research may further explore the long-term effects of AI-assisted assessment on writing development and investigate students' perceptions of AI-supported learning environments across diverse educational contexts.

Study Implications

The findings of this study provide several important implications for language educators, curriculum designers, and educational institutions seeking to integrate artificial intelligence–powered digital assessment tools into EFL writing instruction. The significant improvement observed in the experimental group suggests that AI-assisted feedback can effectively support students' writing development by providing immediate, individualized, and formative feedback which helps learners identify errors, revise their texts, and enhance overall writing quality. Such tools can complement traditional instructional approaches by promoting greater learner autonomy and engagement in the writing process. Furthermore, the integration of AI-driven assessment systems may contribute to more efficient and data-informed formative assessment practices in higher education, enabling instructors to monitor students' progress and provide targeted pedagogical support. Consequently, incorporating AI-based digital assessment tools into EFL writing courses can play a meaningful role in improving academic writing outcomes and modernizing language assessment practices.

Funding: This research is funded by the Center for Human and Philosophical Studies at King Khalid University, Abha, Kingdom of Saudi Arabia, under project number (E 1-1/2026).

Acknowledgments: The author gratefully acknowledges the Center for Human and Philosophical Studies at King Khalid University, Abha, Saudi Arabia, for its generous financial support of this research under Project No. (E 1-1/2026). The author also extends sincere. appreciation to the journal's editorial board and anonymous reviewers for their insightful comments, constructive feedback, and diligent follow-up throughout the review process, which significantly enhanced the quality and rigor of this manuscript.
Conflicts of Interest: The author declares that there are no conflicts of interest related to the publication of this study.
Authenticity: This manuscript represents original work that has not been published previously and is not under consideration for publication elsewhere.
Artificial Intelligence Statement: Artificial intelligence (AI) and AI-assisted technologies were used solely for language editing and proofreading to improve clarity and readability. The author takes full responsibility for the content of the manuscript.

About the Author:
Prof. Drs Ali Albashir Mohammed Alhaj is a University Professor of Translation Studies at King Khalid University, Abha, Saudi Arabia. He holds four PhDs in Translation Studies, Linguistics, and English Literature and is a recipient of an Artificial Intelligence Fellowship. He has published more than 300 research papers and over 60 books, with research interests spanning AI in education, TESOL, translation studies, linguistics, and literature. He received the King Khalid University Excellence Award in 2020 and 2025. He is also a certified editor/reviewer for Taylor & Francis, Frontiers AI/Psychology, and QiST Journal. https://orcid.org/0000-0003-4845-176X

References
Aldosemani, T. I., Assalahi, H., Lhothali, A., & Albsisi, M. (2023). Automated writing evaluation in EFL contexts: A review of effectiveness, impact, and pedagogical implications. *International Journal of Computer-Assisted Language Learning and Teaching, 13*(1), 1–19. https://doi.org/10.4018/IJCALLT.329962

Al-Ghamdi, S. M., Albahiri, M. H., & Alhaj, A. A. (2026). Teachers' perceptions of the feasibility of using generative artificial intelligence in developing multiple-choice questions for formative assessment in Saudi Arabia. *International Journal of Learning, Teaching and Educational Research, 25*(4), 160–184. https://doi.org/10.26803/ijlter.25.4.8

Alhaj, A. A. M. (2024). The impact of machine translation on the development of tourism translation from the perspectives of translators and experts in Saudi Arabia. *Theory and Practice in Language Studies, 14*(4), 1274–1283. https://doi.org/10.17507/tpls.1404.20

Alhaj, A. A. M., Abdelkarim, M. B. A., & Alian, E. M. I. (2026). Artificial intelligence integration in Saudi TESOL: Students' perceptions, ethical readiness, and institutional implementation needs. *International Journal of Learning, Teaching and Educational Research, 25*(3), 19–46. https://doi.org/10.26803/ijlter.25.3.2

Aljuaid, H. (2024). The impact of artificial intelligence tools on academic writing instruction in higher education: A systematic review. *Arab World English Journal (AWEJ), Special Issue on ChatGPT*. https://doi.org/10.24093/awej/ChatGPT.2

Altamimi, D. H. F. (2025). Unlocking potential: Saudi EFL male students' perspectives on AI tools for enhancing English writing proficiency. *Arab World English Journal (AWEJ), Special Issue on Artificial Intelligence*, 40–58. https://doi.org/10.24093/awej/AI.3

Bennett, R. E. (2015). The changing nature of educational assessment. *Review of Research in Education, 39*(1), 370–407. https://doi.org/10.3102/0091732X14554179

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5–31. https://doi.org/10.1007/s11092-008-9068-5

Brown, H. D., & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices* (3rd ed.). Pearson.

Chapelle, C. A., & Voss, E. (2016). 20 years of technology and language assessment in language learning & technology. *Language Learning & Technology, 20*(2), 116–128. http://llt.msu.edu/issues/june2016/chapellevoss.pdf

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage.

Dewi, U. (2022). Grammarly as an automated writing evaluation: Its effectiveness from EFL students' perceptions. *Lingua Cultura, 16*(2), 155–161. https://doi.org/10.21512/lc.v16i2.8315

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). Sage.

Flowerdew, J. (2025). English for research publication purposes. In S. Starfield & C. A. Hafner (Eds.), *The handbook of English for specific purposes* (pp. 421–444). Wiley. https://doi.org/10.1002/9781119985068.ch22

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2019). *How to design and evaluate research in education* (10th ed.). McGraw-Hill Education.

Graham, S. (2019). Changing how writing is taught. *Review of Research in Education, 43*(1), 277–303. https://doi.org/10.3102/0091732X18821125

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

Hyland, K. (2019). *Second language writing* (2nd ed.). Cambridge University Press.

Hyland, K. (2022). *Second language writing* (3rd ed.). Cambridge University Press.

Israel, M., & Hay, I. (2006). *Research ethics for social scientists*. Sage.

Karagoz, I. (2025). AI-generated feedback in English writing instruction for language learners: A systematic review. *The Reading Matrix: An International Online Journal, 25*(1), 51–67.

Karatay, Y., & Karatay, L. (2024). Automated writing evaluation use in second language classrooms: A research synthesis. *System, 123*, 103332. https://doi.org/10.1016/j.system.2024.103332

Li, J. (2023). Artificial intelligence in language assessment: A systematic review. *Computers and Education: Artificial Intelligence, 4*, 100110. https://doi.org/10.1016/j.caeai.2023.100110

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2022). *Intelligence unleashed: An argument for AI in education*. Pearson.

Mekheimer, M. (2025). Generative AI-assisted feedback and EFL writing: A study on proficiency, revision frequency, and writing quality. *Discover Education, 4*(1), 170. https://doi.org/10.1007/s44217-025-00602-7

Ministry of Education. (2020). *Education and training sector strategy*. Ministry of Education, Saudi Arabia.

Murtisari, E. T., Januardi, J. I., Bonar, G., & Kurniawan, D. (2025). Beyond error correction: Lower-proficiency EFL learners' engagement with Grammarly. *Language Awareness*, 1–26. https://doi.org/10.1080/09658416.2025.2585004

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–218. https://doi.org/10.1080/03075070600572090

Pallant, J. (2020). *SPSS survival manual* (7th ed.). McGraw-Hill Education.

Polakova, P. (2024). The impact of ChatGPT feedback on the development of writing skills in foreign language education. *Cogent Education, 11*(1), 2410101. https://doi.org/10.1080/2331186X.2024.2410101

Rahimi, M., Fathi, J., & Zou, D. (2025). Exploring the impact of automated written corrective feedback on the academic writing skills of EFL learners: An activity theory perspective. *Education and Information Technologies, 30*, 2691–2735. https://doi.org/10.1007/s10639-024-12896-5

Redecker, C., & Johannessen, Ø. (2013). Changing assessment—Towards a new assessment paradigm using ICT. *European Journal of Education, 48*(1), 79–96. https://doi.org/10.1111/ejed.12018

Saudi Vision 2030. (2016). *Kingdom of Saudi Arabia Vision 2030*. https://www.vision2030.gov.sa

Shermis, M. D., & Burstein, J. (2013). *Handbook of automated essay evaluation*. Routledge.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Thao, L. T. M., Minh, H. T., Long, D. T., & Anh, N. T. N. (2025). Integrating AI-based feedback into EFL writing instruction: A case study from a Vietnamese university. *Eurasian Journal of Applied Linguistics, 11*(3), 114–127. https://doi.org/10.32601/ejal.11311

Wei, X., Li, Y., & Chen, H. (2023). Automated writing evaluation, self-efficacy, and EFL learners' writing performance. *Frontiers in Psychology, 14*, 1249991. https://doi.org/10.3389/fpsyg.2023.1249991

Wijaya, K. F. (2025). The role of Grammarly as an automated writing evaluation tool in EFL writing learning process. *Datokarama English Education Journal, 6*(2), 61–76. https://doi.org/10.24239/dee.v6i2.127