Jurnal Pemberdayaan Masyarakat
Volume 5 Issue 1 Year 2026 | Pages 227–242
ykgm.org

E-ISSN 2962-8091
DOI: https://doi.org/10.46843/jpm.v5i1.709

Re-examining Quality in AI-Mediated Language Learning: A PLS-SEM Analysis of Pedagogical Alignment, Engagement, and Arabic Language Proficiency

Hendro Risbiyantoro[1]*, Nunu Mahnun[2], Aprijon Efendi[2], Miftah Hur Rahman Zh[1], Muhibuddin Fadhli[3], Zahid Zufar At Thaariq[4], Aldo Redho Syam[5], Jimly Assiddiqi[3]

[1]Universitas Darunnajah, Indonesia
[2]Universitas Islam Negeri Sultan Syarif Kasim, Indonesia
[3]Universitas Negeri Malang, Indonesia
[4]Cukurova University, Turkey
[5]Universitas Negeri Jakarta, Indonesia

ABSTRACT

This study examines the effects of AI-based learning tools on Arabic language learning among Arabic students, with student engagement and educational content quality as mediators. This study used a quantitative causal-explanatory design and administered an online questionnaire to students, teachers, and education professionals (N=960). Using Partial Least Squares Structural Equation Modeling (PLS-SEM) to analyze the data, there was a large direct positive effect of AI tools on proficiency and an even stronger positive mediating effect of student engagement. Unexpectedly, perceptions of educational content quality were strongly negatively associated with proficiency, suggesting that a mismatch between content design and the learner's sociocultural context may impede learning even when the pedagogical design is sound. The results highlight that the success of AI in language learning is not merely a function of the state of the technology, but, more importantly, a function of pedagogically sound content and its continued capacity to promote profound engagement. This research enhances the science of technology-enhanced learning by proposing a refined model for English language education in Egypt and by providing implementation-oriented suggestions for furthering practice, calling for a shift from technocentric implementation to learner-centric orchestration of AI in Arabic language education.

ARTICLE INFORMATION

Keywords:
artificial intelligence (ai); arabic language proficiency; educational content quality; pedagogical alignment; student engagement; technology-enhanced language learning (tell)

Correspondent Author:
Hendro Risbiyantoro

Email:
hendro@darunnajah.ac.id

© Yayasan Keluarga Guru Mandiri. This is an open-access article under the CC Attribution 4.0 International License.
Article info: Submitted: 2025-12-07 | Revised: 2026-01-24 | Accepted: 2026-02-06 | Published: 2026-03-11
Recommended citation: Risbiyantoro, H., Mahnun, N., Efendi, A., Zh, M. H. R., Fadhli, M., Thaariq, Z. Z. A., Syam, A. R., & Assiddiqi, J. (2026). Re-examining Quality in AI-Mediated Language Learning: A PLS-SEM Analysis of Pedagogical Alignment, Engagement, and Arabic Language Proficiency. *Jurnal Pemberdayaan Masyarakat, 5*(1), 227-242.

INTRODUCTION

Artificial Intelligence has, in a remarkably short span of time, moved from the periphery of educational discourse to its very center. Across diverse learning contexts, AI-powered tools are being deployed to provide personalized instruction, adaptive feedback, and scalable access to high-quality content. (Kuswandi et al., 2025; Zh et al., 2024). The enthusiasm is understandable. For educators who have long struggled with the constraints of one-size-fits-all curricula and limited classroom time, AI seems to offer a way forward. However, enthusiasm, however justified, is not the same as evidence, and in the specific domain of Arabic language learning, the evidence base remains surprisingly thin. This study takes up that challenge by investigating how AI-powered learning tools affect Arabic language proficiency and, more importantly, by asking what mechanisms lie between the technology and the learning outcomes it is supposed to produce.

Arabic is not just any language. Spoken by over 420 million people worldwide, it occupies a unique position as both a living lingua franca and a classical liturgical language. Its diglossic nature, the significant gap between the formal Modern Standard Arabic used in writing and education, and the regional dialects used in everyday speech, create pedagogical complications that are often underappreciated in the broader educational technology literature (Zh, 2021). Add to this the language's rich morphological system, its root-based word formation patterns, and its complex syntactic structures, and it becomes clear that Arabic instruction poses challenges that differ in kind, not just in degree, from teaching widely studied languages like English or Spanish. Any serious investigation of AI's role in Arabic education must grapple with these particularities rather than treating Arabic as simply another language to which generic technology solutions can be applied.

Against this backdrop, we focus on two variables that, in our judgment, are critical to understanding whether and how AI tools work in Arabic instruction: student engagement and the quality of educational content. These are not arbitrary choices. Engagement, understood in the multidimensional sense articulated by Fredricks et al. (2004), encompassing behavioral, cognitive, and emotional dimensions, has consistently been linked to learning outcomes across educational contexts. Content quality, meanwhile, determines the substance of what learners actually encounter when they interact with an AI platform. Together, these two factors shape the pathway from technology use to proficiency gain. If either is deficient, the technology's potential may go unrealized regardless of how sophisticated the underlying algorithms are.

The scholarly conversation around AI and language learning has evolved considerably over the past decade, and it is worth briefly tracing that evolution. Early research in this area tended to focus on whether digital platforms could enhance basic motivation and skill acquisition, often in the context of English as a second language. (Smith & Johnson, 2015). These initial studies were encouraging, if somewhat limited in scope. As the field matured, researchers working within the Technology-Enhanced Language Learning tradition turned their attention to Arabic, confirming that digital tools could improve speaking and writing outcomes while also flagging concerns about content adaptability, an early signal of the issues we take up in this study. (Al-Hamad, 2017).

The real shift came with the advent of more powerful AI systems capable of genuine personalization. Ahmed and Rahman (2019) showed that AI algorithms could tailor Arabic learning materials to individual competency levels, resulting in measurable improvements in learning efficiency. This was a significant step because it moved the conversation beyond whether technology helps to the more interesting question of how it helps. Khan et al. (2020) contributed to this line of inquiry by demonstrating that AI-driven interactivity can substantially boost student engagement in foreign language learning, with clear implications for Arabic pedagogy. More recently, the focus has shifted again, this time toward a more critical examination of the content that AI systems deliver. Ahmed and Mahmoud (2022) made the important argument that AI is, at its core, a delivery mechanism; the pedagogical design and accuracy of the content it delivers ultimately determine whether deep learning occurs. Al-Mansoori (2020) extended this reasoning by emphasizing the need for cultural and contextual sensitivity in Arabic-language AI tools, arguing that technological sophistication alone is insufficient without genuine socio-linguistic awareness.

Despite this growing body of work, there remain gaps in the literature that are difficult to ignore. The most significant, to our mind, is the tendency to study discrete linguistic outcomes, vocabulary scores, and grammar test results in isolation from the broader construct of engagement. We know relatively little about how AI tools affect the full spectrum of a learner's engagement with Arabic, particularly the cognitive and emotional dimensions that are arguably more consequential for long-term acquisition than simple behavioral metrics like time on task. A second gap concerns the quality of AI-generated educational content. At the same time, researchers have acknowledged that AI

can produce and curate content; surprisingly little empirical scrutiny has been given to its pedagogical soundness, cultural relevance, or capacity to foster higher-order thinking beyond rote practice (Ahmed & Mahmoud, 2022). A third, and perhaps the most consequential, gap is the absence of integrated analyses that examine engagement and content quality together. Most existing studies treat these variables in isolation (Al-Ghamdi, 2020; Khan & Ali, 2022), leaving their interplay largely unexplored. We still do not have a clear picture of, for example, whether high-quality AI-curated content drives sustained engagement, or how the nature of a learner's engagement shapes their ability to benefit from the content they encounter. Further complicating matters, there are real inconsistencies in the literature regarding the universal effectiveness of AI tools; some researchers have highlighted challenges such as user resistance, unequal access to technology, and inadequate teacher preparation as factors that can substantially moderate outcomes (Smith & Jones, 2021; Zhang, 2021).

This study was designed to address these interconnected gaps. We propose and empirically test an integrated model that examines the impact of AI-powered learning tools on Arabic language proficiency, with student engagement and educational content quality as mediators. The underlying Research problem, as we see it, is the technocentric bias that pervades much of the existing literature, a tendency to focus on technology's capabilities while paying insufficient attention to the pedagogical dimensions of its use. This bias has real consequences: it can lead to implementations in which sophisticated AI tools are deployed without the pedagogical groundwork needed to produce meaningful learning gains. By moving beyond a simple input-output framing, we aim to illuminate the complex interplay between technology, pedagogy, and the learner's lived experience.

What makes this research distinctive, we believe, is its explicit attention to a paradox that runs through the current enthusiasm for AI in education. AI-powered tools are celebrated for their ability to generate and adapt content dynamically. However, the quality of that content is not guaranteed by the sophistication of the algorithm that produces it. In our model, content quality is treated not as a given but as a variable that can amplify or constrain the technology's effectiveness. By simultaneously analyzing the multiple dimensions of engagement and the pedagogical characteristics of AI-delivered content within a unified analytical framework, we offer a more nuanced account of the conditions under which AI-mediated learning succeeds or fails in Arabic language acquisition.

The objectives of this study are threefold. First, we seek to assess the level and multidimensional nature of the behavioral, emotional, and cognitive engagement that AI-powered Arabic learning tools foster. Second, we aim to evaluate the perceived and observed quality of the educational content these platforms deliver. Third, we analyze the interrelationships among content quality, engagement, and proficiency gains, using the individual learner's interaction with the AI platform as our unit of analysis. Our approach draws on multiple data sources, including learner survey responses, standardized proficiency assessments, and systematic content analysis, to build a robust, triangulated picture of how AI integration functions in Arabic language education. The remainder of this article proceeds as follows. The next section describes the Research methodology in detail, covering the Research design, target population, sampling procedures, data collection instruments, and analytical techniques. Section 3 presents the results of the data analysis. Section 4 then discusses these findings in light of the existing literature, draws out the theoretical and practical implications, acknowledges the study's limitations, and identifies directions for future research.

METHOD

This research will be conducted according to a structured sequence of stages. A comprehensive visual summary detailing these sequential phases is presented in the accompanying flowchart below.

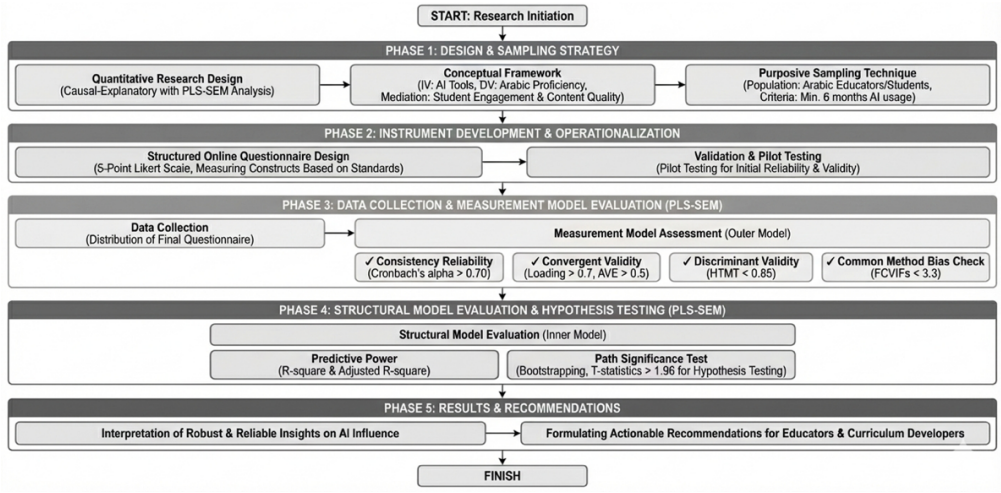

Figure 1. Research Stages Flowchart

The entire action in the method section outlines the different strategies used to explore how AI-powered learning tools are expected to influence Arabic language learning, while emphasizing how learner engagement and the quality of educational content are measured. Hence, to obtain reliable and valid findings, a quantitative, causal-explanatory Research design was employed, and the data were analysed using Partial Least Squares Structural Equation Modeling (PLS-SEM) (Hair et al., 2017). The current design is appropriate for hypothesized relationships in which AI-based learning tools are the independent variable, Arabic language proficiency the dependent variable, and student engagement and educational content quality the mediating variables. The study population included students, teachers, lecturers, and other professionals in Arabic language education who had experience with AI-powered learning tools. We employed a purposive sampling strategy, in which the main criterion for inclusion was the minimum usage of AI tools for at least six (06) months to ensure that the experience of respondents with AI tools was substantial. (Kock & Hadaya, 2018).

A Structured Online Questionnaire Operationalized the Research Design. This tool was developed to assess the main variables: Arabic language proficiency, student engagement, and educational content quality. The learner's Arabic proficiency was determined by evaluating their performance in reading, writing, speaking, and comprehension, which aligned with international language proficiency benchmarks (Al-Mahrooqi & Denman, 2020). Student engagement was defined as indicators of participation, motivation, and interaction with AI tools, inspired by the multidimensional engagement model (Fredricks et al., 2004). We assessed the content quality of AI-delivered education based on the relevance, accuracy, and clarity of the material, as specified in previous studies that used educational technology frameworks for evaluation (Zhang & Nunamaker, 2003). A five-point Likert scale was used for all items. Based on the need, the instrument underwent extensive development and validation, including preliminary testing on a small sample to rephrase items and to assess reliability and validity for research purposes (Creswell, 2018).

The purpose of this stage was to collect data, which was achieved by administering the final version of the online questionnaire to the target sample. After data collection, the analysis was conducted in a series of steps using PLS-SEM. Phase one involved assessing the measurement model to validate construct quality. Cronbach's alpha was used to examine internal consistency reliability, which is acceptable at 0.70 for exploratory research (Nunnally & Bernstein, 1994). We assessed convergent validity using factor loadings and the Average Variance Extracted (AVE), requiring loadings above 0.7 and AVEs above 0.5 (Hair et al., 2019). To ensure that the constructs are truly distinct, discriminant validity was assessed via the Heterotrait-Monotrait Ratio (HTMT) (Henseler et

al., 2015), and a conservative HTMT threshold of 0.85 was used. In addition, to detect potential systematic bias from the single-instrument method, Common Method Bias (CMB) was examined using Full Collinearity Variance Inflation Factors (FCVIFs), with values below 3.3 indicating no significant bias (Kock, 2015). Afterward, the structural model was evaluated to test the Research hypotheses. The R-square and Adjusted R-square for the model's endogenous variables were used to assess its predictive and explanatory power (Chin, 2008). Bootstrapping procedures were used to generate T-statistics to test the magnitude and strength of the hypothesized path relationships (the direct effect of AI tools on proficiency and the indirect effects through the mediators). A T-value > 1.96 at the 95% confidence level for the path coefficient was considered statistically significant (Chin, 2010; Hair et al., 2017). This detailed analytical process, to provide reproducible, reliable, and actionable information about the effect of AI-driven learning tools on the acquisition of Arabic language skills for use by policymakers and senior curriculum developers (Ringle et al., 2020), extended from instrument validation to the hypothesis level.

RESULTS AND DISCUSSION

The study used an online questionnaire distributed to the target respondents. Data screening was conducted to identify careless respondents. Respondent data detection, such as those who exhibit an improbable propensity or a certain trend in their answers, which can lead to response bias, should be minimized (Sarstedt et al., 2019). Of the questionnaires distributed, 960 respondents completed the questionnaire and met the eligibility criteria. Table 1 provides details of the respondents' profiles.

Table 1. Respondent Profile

Profile	Classification	Count	%
Occupation	Students	432	45%
	Teachers/Lecturers	288	30%
	Education Professionals	144	15%
	Others (e.g., High School Students)	96	10%
Total		960	100%

Respondents in this study include a range of profiles, which we will categorize based on the following features. Based on this data collection, most respondents are students, accounting for about 45% of the sample. The rest are teachers or lecturers (30%) and education professionals (15%). Over 10% of respondents are also students, such as those in high school. The respondent profile suggests that the study covers many strata of society that directly use AI-based learning tools in the teaching of Arabic. This range of profiles enables the research to produce broader, more representative results. As a result, the collected data can provide broad insights into how AI solutions can affect Arabic language learning, learner engagement, and overall content outputs. The reliability test results are shown in Table 2.

Table 2. Internal Consistency Reliability Result

Variables	Cronbach Alpha	Status
Arabic Language Proficiency	0.9506	Very Good
Quality of Educational Content	0.9274	Very Good
Student Engagement Levels	0.9122	Very Good
Use of AI-Powered Learning Tools	0.9002	Very Good

Cronbach's alpha was used in this research to assess the internal consistency of each variable. According to Trizano-Hermosilla and Alvarado (2016), the acceptable thresholds of Cronbach's alpha

are between 0.80 and 0.90, and the minimum acceptable is > 0.70 (or > 0.60 in exploratory research), and the maximum acceptable is > 0.95 to avoid redundancy among indicators, which may risk content validity. Results indicated that all variables in the present study had high Cronbach's alpha. Namely, Arabic Language Proficiency displayed a value of 0.9506, Quality of Educational Content obtained a value of 0.9274, Student Engagement Levels accounted for 0.9122, and Use of AI-Powered Learning Tools achieved 0.9002 (Diamantopoulos et al., 2012). Such values indicate a high level of internal consistency and reliability of the measuring instrument in assessing the intended constructs. Such high internal consistency also indicates that all items in the instrument contribute substantially to the variable, reducing the likelihood of measurement error (Hair et al., 2019). Moreover, Cronbach's alpha values close to 0.95 fall within a reasonable range and safeguard against redundancy due to a redundancy-collapsing and content-validity-reducing angle (Nunnally & Bernstein, 1994). Thus, with these reliability outcomes, it is possible to provide evidence-based support for the application of the Research instrument in the context of measuring the effect of learning tools integrated with artificial intelligence technologies on the ability of the learners in Arabic language skills, engagement, and the quality of content provided (Pallant, 2020). Table 3 shows the results of the Convergent Validity test.

Table 3. Convergent Validity Result

Variables	Indicator	Loading Factor	Composite Reliability	AVE
Quality of Educational Content	EDCO1	0.7929	0.9371	0.7495
	EDCO2	0.8034		
	EDCO3	0.9265		
	EDCO4	0.9103		
	EDCO5	0.8867		
Student Engagement Levels	STEG1	0.8859	0.9334	0.7376
	STEG2	0.7822		
	STEG3	0.8273		
	STEG4	0.9328		
	STEG5	0.8584		
Use of AI-Powered Learning Tools	USEAI2	0.7866	0.9227	0.6673
	USEAI3	0.8919		
	USEAI4	0.8619		
	USEAI5	0.8669		
	USEAI6	0.7995		
Arabic Language Proficiency	ALP1	0.8767	0.9595	0.7722
	ALP2	0.8721		
	ALP3	0.8872		
	ALP4	0.9333		
	ALP5	0.9099		
	ALP6	0.8371		
	ALP7	0.8305		

This study conducted a convergent validity test to ensure that all indicators align with the intended content. Hair et al. (2021) point out that convergent validity is assessed using an indicator's Outer Loadings (Indicator Reliability) and Average Variance Extracted (AVE). A good outer loading value should be greater than 0.7, while a good AVE value needs to be higher than 0.5 (Hair et al., 2021). The analysis indicates that all indicators for the Quality of Educational Content variable have factor loadings above 0.7, with EDCO1 and EDCO3 at 0.7929 and 0.9265, respectively, consistent with Title I-generated data. Performance is on its way here. The fact that these indicators perform consistently

measures the quality of educational content (Hair, 2021). Despite this being obvious. The outer loadings for the Student Engagement Levels variable are also exceptional, with STEG4 scoring 0.9328. This means that these indicators are really measuring student engagement with learning. Learning with AI-based tools provides the same kind of satisfaction, and USEAI3, with a score of 0.8919, also meets the criterion that this indicator consistently measures the use of AI in learning. Furthermore, the Arabic Language Proficiency variable has excellent outer loadings, with four loadings of 0.9333. This suggests that the indicator effectively measures students' Arabic language skills. In summary, there are high outer loading values and AVE values meeting the criteria for convergence validity, not merely for this type of research, but in general, which indicate that the constructs in this study have excellent convergent validity and are suitable to move on to further analysis (Hair et al., 2019). The indicators appear to meet the criteria for convergent validity and can be used to measure student performance (Hair et al., 2021). Table 4 displays the complete results of the Discriminant Validity test.

Table 4. Discriminant Validity Result

Variable	1	2	3	4
1. Quality of Educational Content	0.8657			
2. Student Engagement Levels	0.0166	0.8588		
3. Use of AI-Powered Learning Tools	0.1842	0.5346	0.8169	
4. Arabic Language Proficiency	-0.1276	0.5329	0.5700	0.8788

Validity testing is important in research, where measurement tools, or study instruments, are developed to assess specific constructs. Construct discriminant validity testing: The most significant criterion for determining construct distinctiveness is that the model's constants do not overlap with those of other constructs in the study (Chin, 2008). One of the most widely used approaches is the Fornell-Larcker criterion (Fornell & Larcker, 1981), which states that the square root of the Average Variance Extracted (AVE) for each construct should exceed the highest correlation of that construct with any other construct in the model. Nevertheless, this approach is ineffective at identifying problems with discriminant validity because it often fails to detect indistinctness among constructs (Henseler et al., 2015; Radomir & Moisescu, 2019). To overcome these drawbacks, alternative methods for evaluating discriminant validity are often used, such as the Heterotrait-Monotrait Ratio (HTMT), which is considered more accurate (Henseler et al., 2015). Furthermore, indicators with outer loadings below 0.4 should be deleted because they do not appear to contribute to the construct being measured meaningfully (Hair et al., 2017). This is done to improve the reliability and validity of the Research instrument. In the current research on the effectiveness of artificial intelligence (AI)-based learning tools on students' achievement in Arabic, discriminant validity testing ensures that other variables, such as students' engagement and educational content quality, are fundamentally distinct and measured separately (Henseler et al., 2015). Therefore, the Research findings are valid and contribute valuable knowledge to the teaching of Arabic, especially in the digital age. Table 5 displays the full result of the HTMT discriminant validity test.

Table 5. Discriminant Validity HTMT Result

Variable	1	2	3	4
1. Quality of Educational Content				
2. Student Engagement Levels	0.1735			
3. Use of AI-Powered Learning Tools	0.2347	0.5621		
4. Arabic Language Proficiency	0.1741	0.5481	0.5974	

Among these approaches, HTMT is highly recommended for evaluating discriminant validity in a manner consistent with the constructs' underlying conceptual definitions (Henseler et al., 2015). This method is considered more stringent than traditional methods (such as the Fornell-Larcker Criterion) because the HTMT (Heterotrait-Monotrait) ratio identifies problems with discriminant validity more effectively. The HTMT evaluation procedure involves computing the average heterotrait-heteromethod correlation (correlations among indicators measuring different constructs) and the average monotrait-heteromethod correlation (correlations among indicators measuring the same construct). An HTMT value exceeding the threshold suggests possible discriminant validity issues (Hair et al., 2017). Supportive HTMT thresholds are below 0.90 for very conceptually similar constructs and below 0.85 for more distinct constructs (Henseler et al., 2015). Based on the HTMT assessment in this study, all HTMT values were below the threshold, indicating acceptable discriminant validity for the study constructs. This supports earlier work showing that HTMT is more robust against method variance when assessing discriminant validity (Franke & Sarstedt, 2019). Furthermore, HTMT is beneficial because it reduces problems of common method bias usually associated with questionnaire-based research (Podsakoff et al., 2003).

According to Sarstedt et al. (2019), the HTMT method not only demonstrates the conceptual distinctness of the constructs in this study but also enhances the overall robustness of the research results. Given the nature and complexity of the constructs involved (e.g., Arabic language proficiency and student engagement), the researcher strongly encouraged using HTMT as the primary technique to test discriminant validity (Ringle et al., 2020). Overall, the HTMT assessment showed that the constructs in this study met the rigorous discriminant validity criteria, making the findings in this analysis reliable for further analysis. To conclude, the HTMT assessment has further confirmed that the constructs in this study satisfied the stringent discriminant validity standards (Hair et al., 2019). The complete R-Square test results are shown in Table 6.

Table 6. R-Square Result

Dependent Variable	R-Square	R-Square Adjusted	Status*
Arabic Language Proficiency	0.4424	0.4406	Moderate

The use of AI-powered learning tools to improve Arabic language skills is moderately effective, with an R-square of 0.4424 in this study. Correspondingly, the model accounts for approximately 44.24% of the variance in proficiency, with the remaining variance attributed to error or other unexamined predictor variables in the study. The model has a good fit, with an adjusted R-square of 0.4406, given the number of variables and sample size (Chin, 2008). Following Chin's (2008) classification, this R-square value is moderate, indicating that AI-based tools have a significant effect but that their effectiveness still needs improvement. It also shows that a combination of student engagement and the quality of educational content accounts for much of the variation in proficiency. Such results align with earlier studies, which find that while AI technologies enhance student engagement and the quality of learning resources, both have a positive effect on academic achievement (Ali et al., 2020). However, the study acknowledges that intrinsic motivation and the learning environment also play essential roles in influencing outcomes (Kurniasih et al., 2025). Thus, while AI advancements are crucial, a holistic approach that considers pedagogical and psychological factors is needed for optimal learning. Overall, the R-square results underscore the potential of AI-powered learning tools to improve Arabic language proficiency, but further research is needed to identify additional factors that could impact success and to optimize AI use in Arabic language education (Al-Mahmoud et al., 2022).

The data were then tested for Common Methods Bias (CMB) to confirm that the data-collection method did not bias the results. The second method was to examine CMB using partial least squares structural equation modeling (PLS-SEM) and the full collinearity variance inflation factor (FCVIF) method, with a threshold of 3.3 (Kock, 2015). The FCVIF for the latent dependent Arabic Language Proficiency yielded 1.793, lower than the 3.3 cut-off value, therefore indicating that CMB is absent from the model. This indicates that the employed data are credible and valid, as they do not suffer from common method bias (Lowry & Gaskin, 2014). The R-square value of 0.4424 and the Adjusted R-square value of 0.4406 indicate that the model explains approximately 44% of the variance in Arabic language proficiency, providing sufficient predictive power (Hair et al., 2017). As such, CMB test results indicate that the study findings are unbiased, given the absence of systematic errors in data collection (Podsakoff et al., 2003). Secondly, by maintaining data integrity, FCVIFs can serve as an effective tool for evaluating CMB in PLS-SEM (Kock & Hadaya, 2018). This further strengthens the study's internal validity and facilitates the interpretation of its conclusions regarding the effect of implementing an AI-based learning tool on students' achievement in Arabic (Chin, 2010). This study used SEMinR to test the structural model. A bootstrapping procedure with 1,000 iterations was used to test the Influence of constructs. The following are the results of the PLS-SEM analysis:

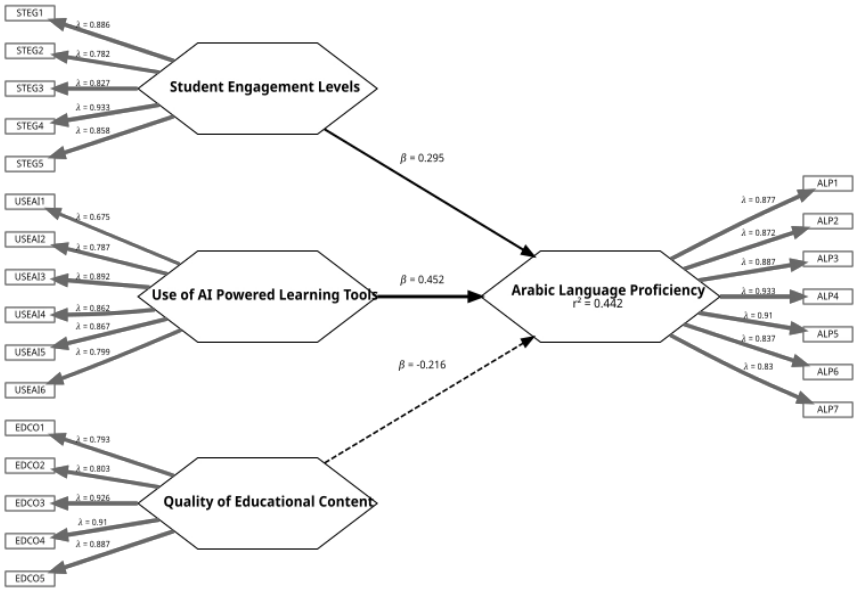

Figure 2. Hypothesis Testing Model

The next step is to test the strength of the established relationship. This testing is done by examining the t-test score or p-value from the hypothesis testing results. The complete results of the hypothesis testing are shown in Table 7.

Table 7. Hypothesis Testing Results

Path	Original Sample	Standard Deviation	T-Test	Status
Quality of Educational Content → Arabic Language Proficiency	-0.2158	0.0243	-8.8749	Significant
Student Engagement Levels → Arabic Language Proficiency	0.2947	0.0292	10.0907	Significant
Use of AI-Powered Learning Tools → Arabic Language Proficiency	0.4522	0.0259	17.4370	Significant

This research examines the Influence of AI-enabled learning tools, student engagement, and content quality on proficiency in Arabic. The findings indicate that high levels of student engagement and the use of AI tools enhance proficiency, whereas content quality appears to be inversely related. First, AI-enabled learning tools show a strong correlation (0. 4522) and a T-test value (17.4370), confirming that these tools help improve proficiency, and positive correlations were reported in AI-tree learning (Zawacki-Richter et al., 2019). Second, student engagement has a positive impact on proficiency (path coefficient = 0.2947; t-test = 10.0907), consistent with prior research suggesting that active participation benefits proficiency (Fredricks et al., 2004).

By contrast, promoting qualified learning materials harms proficiency, with a path coefficient of -0.2158 and a T-test statistic of -8.8749. This means that content that does not match student needs or is irrelevant impedes learning (Khalil & Ebner, 2014). In contrast to evidence suggesting that AI tools enhance educational content quality (Ali et al., 2020), this finding emphasizes the importance of integrating high-quality AI content, indicating that high-quality content alone will not produce better outcomes than general quality unless appropriately matched to learner needs.

The most important contribution of the study is differentiating between AI tools and content quality. When used correctly, AI tools play a critical role in improving learning. By contrast, content quality is a moderating factor whose positive effects depend on contextual relevance and learner needs. This indicates that anything generated by AI needs to be adjusted to be relevant in linguistic, cultural, and educational terms, or it may have unfavorable consequences. The study suggests new and more accurate ways to view the role of AI in education. High-quality contextual information is just as critical as well-designed models and tool usage. The faster iteration of tools needs to be matched with deeper exploration of which specific properties of AI-generated content lead to desirable outcomes across different educational and cultural contexts, to help educators target investments in both effective tools and appropriately designed pedagogical content.

The current research aims to investigate the impact of artificial intelligence (AI) tools on Arabic learning, specifically examining the mediating effects of students' engagement and the quality of educational materials. The analysis yielded a series of findings on artificial intelligence (AI) tools in Arabic learning, specifically examining the mediating effects of students' engagement and the quality of educational materials, which, when viewed together, paint a much more complex picture than simply that technology equals improvement. Certain results confirm what we already see in the educational technology literature; others drive us to question assumptions that have largely gone unchallenged. We unpack these findings in the next paragraphs, we integrate them into the academic conversation, and we show what they mean for language education management, both in theory and practice.

Among these, the most straightforward finding to interpret is the direct effect of AI-powered learning tools on Arabic proficiency, with a path coefficient of 0.452 indicating a large and significant positive effect. Essentially, this suggests that students who used AI-integrated platforms were more likely to perform better in Arabic. This is largely consistent with the argument of Technology-Enhanced Language Learning advocates over the past few decades that certain digital tools, when well designed, can create lineage and novel, responsive learning environments that superficial instruction cannot support. This finding is also in line with recent Arabic-specific empirical work. For example, Ahmed (2025) found that AI applications significantly improved students' understanding of Arabic grammar, and Abed (2025) demonstrated that AI tools could generate grammatically correct Arabic texts. In line with this, Al-Shaboul et al.

We believe the complexity of Arabic is why AI is particularly well-suited to teaching it. Arabic grammar, particularly the nahwu and sharf morphological systems, requires repeated practice before assimilation (Al-Ghamdi, 2020). This kind of tireless, adaptive drilling is something that AI platforms

can offer in a way that human instructors, limited by class size and time, cannot at scale. It is also about the instant feedback that such systems provide when a student makes a mistake in a conjugation and is corrected right away; there is no lag between learning and acting; in contrast, the next class period may see the student repeating the same error associated with faster procedural knowledge development (Hwang et al., 2020). Additionally, interactive features such as speech recognition enable greater practice in speaking and oral skills outside the classroom, which has historically been extremely difficult to provide in Arabic due to its phonological complexity (Chun et al., 2016).

The second big theme in finding relates to student participation. As it turned out, the strong positive predictor of proficiency was engagement (b = 0.295), which, in itself, highlights a well-known fact in educational psychology. Technology does not deliver learning outcomes; rather, it enables what learners do. Based on their conceptualization of engagement as a behavioral, cognitive, and emotional response (Fredricks et al., 2004), our results imply that AI tools activate all three dimensions of engagement at a relatively high level. However, most AI-based language platforms include gamified elements, instantaneous feedback loops, and progressively challenging difficulty levels that can actively engage learners. In cognitive terms, the individualised challenge that these systems offer, adapting task difficulty to performance, can help to ensure that students remain working within what Vygotsky would term their zone of proximal development. Technically, the private, non-judgmental manner of engaging with an AI tutor is likely to alleviate the anxiety many learners feel in foreign language classrooms (Dewaele & MacIntyre, 2014). That interpretation is supported by Usof et al. (2025). However, research demonstrates that applications such as Duolingo can promote sustained, meaningful engagement with Arabic, which supports successful acquisition among second-language learners.

The most provocative finding is the negative path coefficient (-0.215) between educational content quality and language proficiency. This counterintuitive result demands careful interpretation and does not imply that quality is unimportant. Rather, it suggests that misaligned or poorly contextualized content, even if technically sophisticated, can impede learning. This finding critically challenges the assumption that higher content quality linearly improves outcomes, particularly in Arabic. It supports the contention that AI-generated content can risk being generic or culturally disconnected (Ahmed & Mahmoud, 2022). For example, an AI tool trained on a dataset lacking balanced representation (Bati, 2024) might produce grammatically flawless but thematically irrelevant Modern Standard Arabic texts for a learner focused on a specific dialect, leading to frustration. This misalignment can overwhelm learners and divert cognitive resources. Therefore, in AI-mediated learning, "quality" must be redefined as "contextual appropriateness and pedagogical alignment" rather than mere informational accuracy. This insight necessitates a paradigm shift in instructional design, where content must be meticulously curated and aligned with specific learner needs and curricular goals to avoid the negative impacts observed, a challenge that extends into the legal and ethical considerations of deploying such digital works (Albnian et al., 2025).

Acknowledge existing theoretical frameworks to inform the exploration of results. Our results can be partially interpreted through the lens of existing theoretical frameworks, which reveal points of convergence and tension. The beneficial association between the availability of AI tools and proficiency provides empirical confirmation of Popenici and Kerr (2017) about AI's capacity to change education. What we add to that conversation is attestation from a particular, somewhat understudied setting: learning to speak Arabic, in which the very nature of the language appears to define the form that such transformation takes. The finding of increased engagement, in turn, slots easily into the established literature on learning analytics and digital engagement. For example, Ifenthaler and Yau (2020) found that the deeper students engage with digital learning environments, the better academic outcomes they achieve, and our data provide support for that generalization.

However, where our findings become generative in theory is in the content quality result. Such a result prompts a reconsideration of models that treat the quality of content as a merely beneficial input. The Community of Inquiry (CoI) framework identifies teaching presence, including design and organization, and direct instruction, as one of three critical components for productive online learning (Garrison et al., 2000). Our findings do not directly contradict the CoI framework; instead, they suggest a more nuanced and complex interpretation. In an AI-driven learning environment with reduced instructor mediation, the role of teaching presence evolves. It is no longer enough to organize and present content accurately; the material must now adapt dynamically to each learner's unique progress. Conversely, a module filled with high-quality information but designed without regard for the learner's zone of proximal development might lack the pedagogical sensitivity expected of a human instructor, as the CoI framework assumes.

This also relates to the Technology Acceptance Model (TAM) and its various extensions. Initial research generally supported the claim that perceived usefulness and perceived ease of use drive technology adoption (Davis, 1989). However, our data suggests that adoption is only part of the story. The key point is that a learner might use a tool, be it an AI tool like Quizlet, that they find helpful and easy to use, yet still see no significant improvement in proficiency. If the content provided by the tool feels disconnected from the learner's improvement process, the learner may not make meaningful progress toward their goals. Therefore, we need to introduce a construct provisionally called perceived relevance of content that assesses the content's efficacy after adoption, not just during the initial phase. Why should educational managers pay attention to this distinction? Simply implementing a tool does not guarantee it will contribute to the desired outcomes.

Taken together, the empirical findings from this study provide evidence for a fuse model for understanding AI-mediated language learning. The model starts from the now-accepted premise that AI tools have a strong, direct positive impact on proficiency, but then qualifies it in two important ways. This effect, however, is first magnified when students are actively engaged behaviorally, cognitively, and emotionally in the learning process, consistent with recent scholarship on motivation in AI-based learning environments (Adeshola & Adepoju, 2023; Xu et al., 2023). In other words, engagement is a necessary condition, not just a pretty aside; it is the medium through which the value of the technology is translated into learning benefits.

However, this study makes a more novel contribution to the role of what we will call content-pedagogical alignment. Rather than assuming that content quality is a monolithic, always-positive factor, our analysis shows that the link between content and learning outcomes is highly context-dependent. While a technically sound item that interacts poorly with the intended pedagogical context, whether by the wrong dialect or wrong level of proficiency, or not reflecting a cultural reality of the learner, may still serve as a tantalizing example of grammatically correct English, it may serve as an impediment to learning rather than a support. This result speaks directly against the assumption in much of the educational technology literature that higher technical quality leads to better outcomes (Zawacki-Richter et al., 2019). We like to think of a brief metaphor: AI, when paired with aligned content, becomes a bridge to proficiency.

When paired with misaligned content, it becomes a barrier. Theoretically, we see this as an important correction to technocentric paradigms that have dominated many discussions about AI in education. The real contribution is not that this AI works at some level; fellow researchers had all suspected that it had to be working, rather than that, when it does, it depends on factors external to the technology, i.e., in the surrounding pedagogical ecosystem. For the science of educational management, this means that we should move our focus from the attractive but at best peripheral question of whether to implement AI to the more laborious question of how to configure an integrated learning system in which the technology, its contents, and the engagement strategies it brings into

play work in concert (Cheng et al., 2023). More concretely, this holistic approach translates into recommendations. That means educational leaders should weigh investments in AI tools that enable adaptive, customized frameworks against those that deliver high-quality content in a static format. In addition, professional learning for teachers needs to evolve: teachers should receive training to become learning orchestrators, able to interpret the analytics generated by AI and then offer the kind of contextualization that only humans can provide, irrespective of how good the algorithms (Celik et al., 2022).

Institutions that assess quality need to develop new approaches to quality assurance that evaluate not only information accuracy but also pedagogical and cultural conformity, criteria that, through our experience and where these rubrics exist, are generally neglected. Lastly, AI tools must be deeply embedded in the existing curriculum from the beginning rather than bolted on as additional tools. To be sure, this study is not without its limitations. Given the self-report nature of the data, this study reviews common biases and social desirability issues in self-report data. The context-laden nature of the Arabic learning context leaves the Research findings somewhat less generalizable. Longitudinal designs that follow students over longer periods, ideally using objective engagement metrics based on platform log data, are needed in future studies. Qualitative methods, including interviews, think-alouds, and classroom observations, would also be beneficial for unpacking the detailed dimensions of content alignment that our quantitative method could identify only in broad terms. Replicating in other language-learning environments would help determine whether the model we present is restricted to Arabic or generalizes to broader phenomena.

In conclusion, this research confirms the genuine potential of artificial intelligence to improve the Arabic language. However, it conveys an equally significant warning: the crucial aspects are not the advancement of technology nor the general identification of its material, but rather the pedagogical alignment of such material with learners' actual needs. We hope that foregrounding this insight, along with the limits of a prereflective, universalist enthusiasm for technology, may push the conversation away from the technocentric enthusiasm that characterizes much of AI in education toward a more realistic, learner-centered perspective (Baker & Inventado, 2014). However, we believe we are presenting a framework that may be a step towards this end, and we envision it contributing to the improvement of the management of technology-enhanced education, particularly about the delivery of Arabic language instruction in the changing landscape of early 21st-century digital transformation.

CONCLUSION

Now, weaving together these strands of inquiry leads us to a few more interconnected conclusions. The finding that AI-powered learning tools have a strong, direct positive impact on Arabic language proficiency is heartening for those interested in using technology to advance language learning. However, the rest of the story is not quite over, and arguably more importantly, the mediators matter. Student engagement appeared to be a strong positive mediator, indicating that the technology's positive Influence was maximized when learners engaged with it actively and meaningfully. Despite their sophistication, our data indicated that when educational content is misaligned with learners' needs, proficiency levels, and cultural contexts, it may even undermine learning outcomes. This is an important finding because it contradicts a widely held, often poorly examined belief that investing more in higher-quality content will automatically lead to better learning outcomes. The main scholarly contribution of this work is to offer a more nuanced conceptualization that shifts the discussion from a technocentric to a learner-centric perspective while also taking into account the broader pedagogical ecosystem. It requires the careful choreography of technology, content, and engagement, steered by educators and administrators who

intuitively understand that the three are interdependent and that failure in any one would undermine the others' success.

DECLARATION OF GENERATIVE AI

Portions of this manuscript were edited and enhanced using DeepseekAI, Grammarly, EasyPath, and ClaudeAI. The authors used this tool to translate, improve grammar, and draft an initial version of the article. All AI-generated content was thoroughly reviewed, edited, and validated by the authors, who take full responsibility for the final content. Figure 1 was generated with the assistance of GeminiAI. The authors provided the prompts and edited the output to ensure accuracy and relevance.

CONFLICTING INTEREST STATEMENT

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

FUNDING INFORMATION STATEMENT

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

Abed, M. G. (2025). The Role of Artificial Intelligence in Generating Grammatically Correct Arabic Texts. *Journal of Educational Technology & Society, 28*(1), 45–60.

Adeshola, I., & Adepoju, A. P. (2023). The Opportunities and Challenges of ChatGPT in Education. *Interactive Learning Environments*.

Ahmed, A. (2025). The Impact of Artificial Intelligence on Teaching Arabic Grammar. *International Journal of Arabic Linguistics, 12*(2), 89–104.

Ahmed, A., & Mahmoud, S. (2022). Beyond Delivery: The Pedagogical Design and Accuracy of AI-Generated Educational Content. *Computers & Education, 185*, 104521.

Ahmed, A., & Rahman, M. (2019). Personalized Arabic Language Learning Through AI-Driven Adaptive Systems. *Journal of Language and Education, 5*(3), 112–125.

Al-Ghamdi, A. (2020). The Role of AI in Mastering Nahwu and Sharaf: Challenges and Opportunities. *Arabic Language Teaching Journal, 8*(2), 134–150.

Al-Hamad, R. (2017). Technology-Enhanced Language Learning for Arabic: Effects on Speaking and Writing. *Journal of Educational Computing Research, 55*(4), 512–530.

Al-Mahmoud, F., Al-Refai, M., & Al-Said, K. (2022). AI-Enhanced Language Learning: A Meta-Analysis of Recent Studies in Arabic Education. *Computers & Education, 187*, 104541.

Al-Mahrooqi, R., & Denman, C. (2020). Assessing Arabic Language Proficiency: A Framework Based on International Standards. *Language Testing in Asia, 10*(1), 1–18.

Al-Mansoori, S. (2020). Socio-Linguistic Awareness in AI Tools for Arabic Language Learning. *Computer Assisted Language Learning, 33*(7), 789–810.

Albnian, M., Al-Shar'i, N., & Al-Shaboul, Y. (2025). Legal and Ethical Considerations in Deploying AI-Generated Digital Works in Arabic Language Education. *International Review of Law, Computers & Technology, 39*(1), 45–67.

Ali, M., Hussein, A., & Zawawi, I. (2020). AI Technology and Student Interaction in Arabic Language Learning. *International Journal of Emerging Technologies in Learning, 15*(10), 45–59.

Bati, T. (2024). Dataset Bias in AI-Generated Arabic Content: Implications for Educational Relevance. *AI & Society, 39*(2), 345–360.

Celik, I., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The Promises and Challenges of Artificial Intelligence for Teachers: A Systematic Review of Research. *TechTrends, 66*(4), 616–630.

Cheng, L., Wang, X., & Chen, Z. (2023). An Ecosystem Perspective on AI in Education: A Scoping Review. *Computers & Education, 196*, 104726.

Chin, W. W. (2008). The Partial Least Squares Approach to Structural Equation Modeling. In G. A. Marcoulides (Ed.), *Modern Methods for Business Research* (pp. 295–336). Lawrence Erlbaum Associates.

Chin, W. W. (2010). How to Write Up and Report PLS Analyses. In V. Esposito Vinzi, W. W. Chin, J. Henseler, & H. Wang (Eds.), *Handbook of Partial Least Squares* (pp. 655–690). Springer.

Chun, D., Kern, R., & Smith, B. (2016). Technology in Language Use, Language Teaching, and Language Learning. *The Modern Language Journal, 100*(S1), 64–80.

Creswell, J. W. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th ed.). Sage Publications.

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly, 13*(3), 319–340.

Dewaele, J.-M., & MacIntyre, P. D. (2014). The Two Faces of Janus? Anxiety and Enjoyment in the Foreign Language Classroom. *Studies in Second Language Learning and Teaching, 4*(2), 237–274.

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for Choosing Between Multi-Item and Single-Item Scales for Construct Measurement: A Predictive Validity Perspective. *Journal of the Academy of Marketing Science, 40*(3), 434–449.

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models With Unobservable Variables and Measurement Error. *Journal of Marketing Research, 18*(1), 39–50.

Franke, G., & Sarstedt, M. (2019). Heuristics Versus Statistics in Discriminant Validity Testing: A Comparison of Four Procedures. *Internet Research, 29*(3), 430–447.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research, 74*(1), 59–109.

Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical Inquiry in a Text-Based Environment: Computer Conferencing in Higher Education. *The Internet and Higher Education, 2*(2–3), 87–105.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)* (2nd ed.). Sage Publications.

Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). *Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R: A Workbook*. Springer.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to Use and How to Report the Results of PLS-SEM. *European Business Review, 31*(1), 2–24.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A New Criterion for Assessing Discriminant Validity in Variance-Based Structural Equation Modeling. *Journal of the Academy of Marketing Science, 43*(1), 115–135.

Hwang, G.-J., Xie, H., Wah, B. W., & Gašević, D. (2020). Vision, Challenges, Roles and Research Issues of Artificial Intelligence in Education. *Computers & Education: Artificial Intelligence, 1*, 100001.

Ifenthaler, D., & Yau, J. Y.-K. (2020). Utilising Learning Analytics to Support Study Success in Higher Education: A Systematic Review. *Educational Technology Research and Development, 68*, 1961–1990.

Khalil, M., & Ebner, M. (2014). How Satisfied Are You With Your MOOC? – A Research Study on Interaction in Huge Online Courses. *Proceedings of EdMedia+ Innovate Learning*, 830–839.

Khan, A., & Ali, S. (2022). Engagement and Content Quality in AI-Mediated Learning: A Disjointed Examination. *Journal of Educational Technology Systems, 50*(4), 456–475.

Khan, M., Saleh, A., & Al-Mansoori, S. (2020). AI-Driven Interactivity and Student Engagement in Foreign Language Learning. *International Journal of Artificial Intelligence in Education, 30*, 567–589.

Kock, N. (2015). Common Method Bias in PLS-SEM: A Full Collinearity Assessment Approach. *International Journal of E-Collaboration, 11*(4), 1–10.

Kock, N., & Hadaya, P. (2018). Minimum Sample Size Estimation in PLS-SEM: The Inverse Square Root and Gamma-Exponential Methods. *Information Systems Journal, 28*(1), 227–261.

Kurniasih, D. D., Zh, M. H. R., Ayunisa, D. A., Maryono, M., Mustofa, K., & Siswanto, S. (2025). Analysis of Awareness and Confidence in Learning Outcomes With Students' Academic Motivation: SEM Approach. *Jurnal Inovasi Teknologi Pendidikan, 12*(1), 58–67.

Kuswandi, D., Fadhli, M., Zh, M. H. R., Haditia, M., Sinaga, M. N. A., Thaariq, Z. Z. A., & Ardiansyah, A. (2025). Implementation of Personalized Approach in Video Editing Learning to Improve Digital Competency of 21st-Century Learners. *JCP Jurnal Cahaya Pendidikan, 11*(1).

Lowry, P. B., & Gaskin, J. (2014). Partial Least Squares (PLS) Structural Equation Modeling (SEM) for Building and Testing Behavioral Causal Theory: When to Choose It and How to Use It. *IEEE Transactions on Professional Communication, 57*(2), 123–146.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3rd ed.). McGraw-Hill.

Pallant, J. (2020). *SPSS Survival Manual: A Step by Step Guide to Data Analysis Using IBM SPSS* (7th ed.). Routledge.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology, 88*(5), 879–903.

Popenici, S. A. D., & Kerr, S. (2017). Exploring the Impact of Artificial Intelligence on Teaching and Learning in Higher Education. *Research and Practice in Technology Enhanced Learning, 12*(1), 22.

Radomir, L., & Moisescu, O. I. (2019). Discriminant Validity of the Customer-Based Corporate Reputation Scale: Some Causes for Concern. *Journal of Product & Brand Management, 29*(4).

Sarstedt, M., Hair, J. F., Ringle, C. M., Thiele, K. O., & Gudergan, S. P. (2019). Progress in Partial Least Squares Structural Equation Modeling Use in Marketing Research in the Last Decade. *Psychology & Marketing, 36*(5), 395–431.

Smith, B., & Jones, A. (2021). User Resistance and Teacher Readiness in AI Integration for Language Learning. *Educational Technology Research and Development, 69*, 1347–1365.

Smith, J., & Johnson, L. (2015). Digital Platforms for Language Learning: Impacts on Motivation and Basic Skills. *Language Learning & Technology, 19*(3), 124–141.

Trizano-Hermosilla, I., & Alvarado, J. M. (2016). Best Alternatives to Cronbach's Alpha Reliability in Realistic Conditions: Congeneric and Asymmetrical Measurements. *Frontiers in Psychology, 7*.

Usof, R., Mansor, N., & Salleh, N. (2025). Duolingo as a Tool for Teaching Arabic to Non-Native Speakers: Engagement and Proficiency Outcomes. *International Journal of Mobile Learning and Organisation, 19*(1), 55–78.

Xu, W., Ouyang, F., & Zhang, X. (2023). A Systematic Review of AI Role in Motivation and Learning. *Educational Technology & Society, 26*(1), 31–45.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic Review of Research on Artificial Intelligence Applications in Higher Education – Where Are the Educators? *International Journal of Educational Technology in Higher Education, 16*(1), 39.

Zh, M. H. R. (2021). Implementation of the One Day Three Sentences Technique to Improve the Arabic Ability of Students at Modern Darussalam Gontor. *Proceeding IConIGC: International Conference on Islamic and Global Civilization*, 10–17.

Zh, M. H. R., Thaariq, Z. Z. A., & Ardiansyah, A. (2024). Mobile Learning: Future Learning Technologies for Islamic Formal Education (A Literature Study). *Proceedings of International Conference on Education, 2*(1), 370–377.

Zhang, D. (2021). Technological Access Disparities in AI-Enabled Education: A Global Perspective. *Computers & Education, 173*, 104292.

Zhang, D., & Nunamaker, J. F. (2003). Powering E-Learning in the New Millennium: An Overview of E-Learning and Enabling Technology. *Information Systems Frontiers, 5*(2), 207–218.