

ACTFL-Based Computerised Adaptive Testing for Quality Management in Arabic Instruction at an Islamic University

Ihwan Mahmudi^{*1}, Moh. Ismail¹, Taufik Rizki Sista¹, Ahmad Saifulloh¹, Syarifah¹

¹ Universitas Darussalam Gontor, Indonesia

*Corresponding Author: ihwanm@unida.gontor.ac.id,

Submitted: 05-01-2026

Revised : 07-05-2026

Accepted: 20-07-2026

ABSTRACT. This research examines the implementation of Computerized Adaptive Testing based on ACTFL standards at Universitas Darussalam Gontor (UNIDA). In the midst of the increasing adoption of global language assessment standards. This research employs a qualitative approach, utilizing thematic analysis. Data were obtained through in-depth interviews with university leaders, quality assurance units, lecturers, system developers, and students. The results show that CAT-ACTFL not only functions as an instrument for measuring language proficiency, but also becomes an important part of academic governance, curriculum management, and the strengthening of data-based quality assurance systems. Its implementation reflects a strategic alignment between global assessment standards and UNIDA's institutional vision framework, positioning Arabic as a pillar of academic policy and learning planning. This study introduces a new model of Pesantren-Aligned Language Assessment Governance as a key novelty, demonstrating that the global CAT-ACTFL standards can integrate harmoniously with the local values of pesantren. This research makes a significant contribution to educational management by transforming static quality assurance into a dynamic, data-driven system for differentiating Arabic language instruction in higher education. Nevertheless, this research also indicates the importance of further evaluation of the quality of the ACTFL instruments used. Therefore, further research is recommended to conduct validity and reliability testing and to analyze the item characteristics of the CAT-ACTFL instruments using psychometric approaches, such as Item Response Theory (IRT) and the Rasch Model, to ensure more precise and scientific measurement of language proficiency.

Keywords: Adaptive Testing, ACTFL, Arabic Language Assessment, Quality Assurance, Education Management

 <https://doi.org/10.31538/munaddhomah.v7i3.2908>

How to Cite Mahmudi, I., Ismail, M., Rizki Sista, T., Saifulloh, A., & Syarifah. (2026). ACTFL-Based Computerised Adaptive Testing for Quality Management in Arabic Instruction at an Islamic University. *Munaddhomah: Jurnal Manajemen Pendidikan Islam*, 7(3), 604–620.

INTRODUCTION

Arabic language proficiency assessment in Islamic higher education has experienced a significant shift in meaning, from merely a pedagogical instrument to a strategic system that influences academic governance, institutional credibility, and the direction of learning development (Mahmudi et al., 2025). Language evaluation is no longer understood only as a final measure of student achievement, but as a structural mechanism that has a role in curriculum planning, formulation of academic policy, and the strengthening of educational quality legitimacy within the global academic landscape (Almelhes, 2024; Musliadi, Triyono, & Jamilah, 2024; Sahayu, Triyono, & Ungu, 2026; Sartini, Triyono, & Triastuti, 2025). This change shows that

assessment is not only a tool for competency certification but also part of the infrastructure for decision-making and institutional quality development (Alsubaie, 2022).

As a response to this dynamic, many higher education institutions have begun to adopt international standard-based assessment frameworks to improve consistency, objectivity, and comparability of language evaluation results (Haerullah, Jundi, Hasibuan, Sari, & Asmarita, 2024). One of the frameworks that is widely used is ACTFL (American Council on the Teaching of Foreign Languages), which offers a classification of language ability based on structured and measurable performance levels (Harsch & Malone, 2020; Tayyara, 2022). The integration of ACTFL with Computerised Adaptive Testing (CAT) further expands the function of assessment by allowing the system to adjust the level of question difficulty dynamically based on participant responses, thus producing measurement that is more accurate, efficient, and diagnostic (Demir & Gelbal, 2025). This approach opens space for assessment to not only measure, but also to inform learning strategies and curriculum development in a more precise manner (Amalia, 2026; Hussain, 2025; Isnaeni, Budiman, Nurjaya, & Mukhlisin, 2025; Junaidi, Afrinaldi, Artika, Al-Momani, & Hulwana, 2025; Mahyudin Ritonga et al., 2024).

However, most of the existing studies are still focused on technical aspects such as psychometric validity, system efficiency, and algorithmic performance, so that they tend to separate assessment from the broader institutional context. The research that was conducted by El-Msayer (Msayer, Aoula, & Bouihi, 2024) and Ahmad Safeer (Safeer, 2025) has highlighted the strengths of psychometric and the advantages of time-saving that CAT provides. In contrast, Oller Jill has been investigating the application of it in Arabic language education by using IRT, CTT, and digital platforms (Jill, Gilkerson, Jane, Hannon, & Steven, 2023). Nonetheless, most of the initiatives have focused mainly on the development of technical tools, failing to incorporate adaptive testing into broader quality assurance frameworks. Although Joseph de Curto acknowledged the significance of the quality of curriculum and the training of teachers, they failed to integrate adaptive evaluation as a systematic and data-informed component within the institutional frameworks (De Curtò, De Zarzà, Fervier, Sanagustín-Fons, & Calafate, 2025).

Although the research conducted by Pransiska utilised the ACTFL framework for developing a comprehensive assessment instrument for Arabic language textbooks only (Pransiska, Sugiyono, Widodo, & Sayyid, 2024). Adaptive assessment is often treated as a standalone technological innovation, without being deeply examined regarding how it interacts with institutional governance, quality assurance systems, and more strategic education management practices (Wang, Medaglia, & Zheng, 2018). This condition creates a conceptual gap related to how CAT-ACTFL can be positioned not only as a measurement tool, but as an integral part of the institutional architecture that influences policy making and the direction of academic development (Oladele & Ndlovu, 2021; Ozan, Ozarslan, & Calisir Zenci, 2025).

This gap becomes more visible in the context of Islamic higher education, where the Arabic language has a very central epistemological position as the language of the Qur'an and the main medium of Islamic knowledge transmission (Safeer, 2025). At Universitas Darussalam Gontor, Arabic language is not merely taught as a subject, but becomes an intellectual foundation that shapes the academic orientation of the institution as reflected in the Gontor Waqf Charter (Pransiska et al., 2024). Therefore, the implementation of the ACTFL-based assessment system in this environment raises an important question about how global standards are implemented without shifting the strategic meaning of the Arabic language within the *pesantren* education system (Al-Assaf, 2025; Arif, Zekkari, & Addawiyah, 2025; Baihaqi, Amaliyah, Awaliyah, Khoerunnisa, & Laksono, 2023; Fatah, Khoiruddin, & Fakhruddin, 2025; Hasan, Mujahidin, Azizah, & Solechan, 2024).

This research intends to answer those questions by analysing the integration of the computer adaptive test based on the ACTFL standard framework in the academic governance system at the University of Darussalam Gontor (Mahmudi et al., 2024). Instead of focusing on the technical aspect only, this study discussed how CAT-ACTFL runs as a strategic mechanism that

supports curriculum management, encourages academic monitoring, and supports quality improvement of Arabic teaching. The main focus is on how the data assessment utilised in the decision-making process influences the practical teaching, also encourage institution accountability through evidence-based evaluation.

METHOD

This research uses an interpretive qualitative case study design. (Yin, 2018) to examine how the Computerised Adaptive Testing system based on ACTFL standards (CAT–ACTFL) is integrated into academic governance and quality assurance at Universitas Darussalam Gontor. This approach is selected because the research does not only focus on the technical aspect of assessment, but also on the process of meaning construction, institutional dynamics, and the interaction between the global assessment system with institutional structure and *pesantren* educational culture (Miles & Huberman, 1994). The data collection techniques employed in this study comprise three primary methods: in-depth interviews, participant observation, and documentation reviews (Creswell, 2018). In-depth interviews are conducted with purposively selected informants from various levels of academic governance to comprehensively represent the dynamics of CAT–ACTFL integration, starting from strategic policy to direct user experiences (Anderson, Leahy, DelValle, Sherman, & Tansey, 2014).

Data analysis in this study follows Creswell's thematic analysis approach, which proceeds through a linear, hierarchical hierarchy of data transformation (Braun & Clarke, 2006). The process begins with organising and preparing the raw data, including transcribing interviews and field notes from UNIDA Gontor (Bogdan & Biklen, 1998; Braun & Clarke, 2006). Next, the data is read thoroughly to obtain a general sense of the information and reflect on its overall meaning. Coding is then conducted by labelling the text to generate specific descriptions and themes (Creswell, 2018). In accordance with the object of this research, the coding process specifically filters themes related to how the CAT–ACTFL system aligns with the Arabic language as an epistemological foundation, and how global assessment standards adapt to the institutional values of the *pesantren* educational culture (Miller & Salkind, 2002).

Table 1. Typology of Participants and Spectrum of Governance
Perspectives in CAT–ACTFL Integration

Informant	Governance Level	Main Interview Focus
VR1	Macro Governance	Arabic language quality vision, rationale of ACTFL–CAT adoption, integration in SPMI & accreditation, policy challenges
DPB	Language Governance	ACTFL adaptation into the <i>pesantren</i> context, system design, and challenges
KPR	Meso Governance	Impact of ACTFL data on curriculum, placement, remedial, and class policy
GPM	QA Governance	Integration of assessment data in the quality evaluation system
DLC	Micro Academic	Change of teaching strategy and implementation of assessment results
DEV	Technical Layer	Adaptive mechanism, data flow, system validity & security
CST	User Experience	Perception of fairness and impact of test results on learning strategy

Source: Data Collection, 2026

RESULT AND DISCUSSION

Result

Implementation And Development System

This thematic domain describes how CAT-ACTFL is used in practice through a mix of technical design, administrative stages, and continuing development at the institution. The results reveal that CAT-ACTFL is not just a digital system; it is also an organised instructional framework where the system's structure, localisation, validation processes, and technical protections all work together to support its main purpose. The network code mapping provided here illustrates the connections between these components as revealed by the Atlas. ti study. To visually represent the internal dynamics of this domain, a network code mapping is presented below, illustrating the relational structure among key implementation-related codes generated through Atlas.ti analysis.

The illustration emphasises the importance of different codes, including localisation, architectural systems, validation systems, technical safeguards, system professionalism, and other supporting procedures, and how they relate to each other. The dependability of technology isn't derived from a single component alone. The source lies in the design of the system, the processes of quality control, and the interconnected changes that adapt to the institution's needs. The close grouping of codes related to architecture and validation highlights the need for a strong and stable structure. At the same time, the presence of localisation and technical safeguard elements shows a continuous effort to adapt global testing standards to local institutional needs. Overall, this network suggests that the success of CAT-ACTFL implementation depends on a coordinated process combining technological accuracy, procedural consistency, and contextual sensitivity, rather than relying only on separate technical solutions.

Researchers selected texts specifically tailored to the context and background of UNIDA students. In this reduction process, the themes raised focused on socio-religious phenomena commonly found in Islamic boarding schools. This was done to ensure the material presented remained relevant to the daily lives of students in these environments. (Code-DEV)

The text selection is adapted to the characteristics of UNIDA students by selecting socio-religious themes that are familiar with the Islamic boarding school ecosystem. (Code-VR1).

These findings also show that although ACTFL is considered a strong framework and quite flexible in classifying language skills, its effectiveness in *pesantren*-based higher education is highly dependent on the level of adjustment toward local needs. For the assessment system to remain technically accurate, pedagogically relevant, and aligned with institutional direction, changes are required in the formulation of indicators, reporting format, and curriculum development direction. Several supporting visuals display how this process is designed systematically, starting from the formulation of Arabic language proficiency level parameters based on ACTFL standards, the structured construction of the evaluation item bank, examples of test item forms in the database, and the interface display of the test page and system login page.

These visuals emphasise that the system development is not purely technical in nature, but also reflects a mature conceptual process. Furthermore, this system development process does not only stop at the digital infrastructure aspect, but also includes a comprehensive approach that integrates global assessment principles, adaptation based on *pesantren* tradition, and strong technological support for data-based continuous evaluation. This foundation has a strategic role in improving the quality of Arabic language learning in Islamic higher education, as well as strengthening the position of ACTFL as an adaptive and relevant assessment framework in the local context.

Rationality and Orientation of ACTFL Usage

This theme describes basic logical and strategic reasoning to decide the adoption of ACTFL as the main frame in Arabic assessment. This finding shows that the ACTFL is not only accepted as an external benchmark, but also as a normative guiding compass that rebuilds the proficiency that can be measured understandable in a university context. The informant's consistency framed this adoption as a response to the increasing need for a need about standard assessment which transparent, accountable, and comparable globally.

A strong emphasis on this theme pointed to an effort to create fairness and adaptive assessment. ACTFL decided to provide a measurable standard to assess multiple and various students' proficiency in more detail and ensure the equity of the assessment process. This orientation reflects the massive institutional support to change from a subjective assessment to an accountable and measurable assessment system, as well as strategic mapping of learning outcomes. As mentioned in various informant's quotation below:

CAT-ACTFL is completely different. Its adaptive logic is built on ACT, so I feel this test is clearer and fairer than the usual ones. It follows, but today the expectations are much more complex. We're dealing with digital-native students who require dynamic evaluation, where language proficiency must integrate seamlessly with global communication skills and complex problem-solving." (VR1-DEV-CST).

ACTFL is seen as a symbol of the institution's seriousness in adopting globally recognised standards, while at the same time strengthening trust among academic actors and reinforcing the university's position within the international academic landscape. This orientation also reflects a broader paradigm shift, where assessment is no longer viewed as a standalone academic procedure, but rather as an integral component of strategic educational planning. To describe the internal conceptual structure of this theme, a network code mapping is presented below, which shows the interrelation among the main codes that define the rationality and orientation of ACTFL usage.

This visual network represents the interconnectedness among dominant codes such as fair and adaptive assessment, measurable standards, system accountability, global standards, and strategic mapping. This visual shows how these codes meet each other and form a complete logic, which positions ACTFL not only as a technical framework, but also as a strategic reference point in institutional policy formulation, assessment philosophy, and long-term direction of education. The pattern of linkage in this network shows that the orientation toward ACTFL is built on multi-layered rationality, where technical consideration intersects with managerial vision and pedagogical objectives. This thematic configuration demonstrates that the adoption of ACTFL reflects both operational needs and strategic aspirations, therefore strengthening its position as the main pillar in continuously developing assessment culture within the institutional environment.

Impact On Integration With Quality Assurance Governance

This theme shows how CAT-ACTFL begins to be integrated in a real manner with the internal quality assurance system and academic governance of the university. As mentioned by one of the informants. The findings reveal that the assessment data not only stops in the measurement function, but also is utilised as the basis for the decision-making process, program evaluation, and continuous improvement through the PPEPP cycle. Several informants emphasise that the CAT-ACTFL results become an important reference in quality evaluation meetings, curriculum adjustment, and follow-up planning, as mentioned below.

This integration shows the shifting of the assessment function from merely evaluation of learning outcomes to a strategic instrument in quality management. The produced data is used to strengthen quality documentation, support accreditation evidence, and become a foundation in strengthening academic policy that is more based on data. This integration also demonstrates the shifting of the assessment function from merely evaluation of learning outcomes to a strategic instrument in quality management. The produced data is used to strengthen quality

documentation, support accreditation evidence, and establish a foothold in strengthening academic policy that is more based on data.

The ACTFL result is now seen not merely as a score, but as critical data to evaluate overall student performance. This data-driven approach is supported by the university's internal system, as another informant noted, At UNIDA, we already have a well-established quality cycle. I always ensure that every assessment tool aligns perfectly with our institutional standards and continuous improvement goals. (Code VR1-GPM)

In practice, this comprehensive evaluation directly impacts classroom instruction because we finally know the exact proficiency level of the students, so the teaching methods are no longer generalised or treated the same for everyone. Ultimately, this system encourages individual growth, as shared by a student who remarked, "The result makes me realise my specific language gaps, especially in my listening and speaking skills" (Code CST-DLC).

This network visual displays the interconnection among codes such as QA integration, follow-up action, PPEPP cycle, QA evidence, and support-based QA. The network shows that the integration is not linear in nature, but takes place through a reciprocal relationship between the quality evaluation mechanism, the reporting system, and the policy follow-up action. This condition emphasises that CAT-ACTFL plays a role as an important driver in strengthening data-based quality culture. This theme describes how the implementation of CAT-ACTFL gives direct influence on learning practice and curriculum arrangement at the classroom level. The findings show the existence of changes in teaching strategy, class grouping based on ability level, and adjustment of learning targets that are more realistic and measurable.

Another impact that appears is the development of a differentiated learning approach, adjustment of teaching materials, and the arrangement of a learning pathway that is more adaptive. ACTFL not only guides the final evaluation, but also shapes the direction of the learning process since the beginning of the semester. This figure shows the relationship among codes such as differentiated instruction, learning pathway, diagnostic insight, and learning target setting. The network displays how the CAT-ACTFL assessment data influence teaching strategy, determination of learning targets, and development of learning approaches that are more responsive to students' needs. Therefore, ACTFL plays a role in creating learning that is more directed, contextual, and aligned with the expected competency achievement.

The Challenges of ACTFL Implementation

This theme describes the various obstacles and tensions that appear during the process of CAT-ACTFL implementation, both at the technical, pedagogical, and institutional levels. The challenges faced are not only related to system readiness, but also touch on the aspect of human resource readiness, work culture adaptation, and the change of mindset toward assessment based on global standards. Several informants reveal that the early phase of CAT-ACTFL implementation is hit by confusion and resistance, especially in understanding the adaptive logic of the system and the assessment mechanism, which is considered different from conventional evaluation practice. One informant quoted below:

Other challenges also appear in the technical aspect, such as limitations of network infrastructure, server load, and access stability when the test takes place simultaneously. This also influences user comfort and creates anxiety toward the reliability of the system. On the institutional side, the issue of data literacy and managerial readiness has also become a concern. Not all academic units have the same understanding of utilising CAT-ACTFL result data as the basis for decision-making. In addition, the shifting of assessment culture from a subjective approach toward a standards-based system also creates psychological tension, especially for lecturers who have long been accustomed to traditional evaluation methods.

In the beginning, many students were confused because the adaptive testing format was entirely new to them. This confusion was often exacerbated by infrastructure limitations, which directly impacted the students' emotional state during the examination (Code DLC).

Sometimes during the test, the signal is not stable, so we become quite anxious (about our progress). Despite these field-level constraints, the institution emphasises that the ultimate goal of this transition goes beyond technicalities, requiring lecturers and staff to "really understand how to read and use this ACTFL data, not only seeing the final score" (Code GPM-CST).

This transition process demands adjustment that is not only technical in nature, but also mental and cultural. This network visual shows the relationship among codes such as system anxiety, adaptive confusion, technical constraints, server load, data literacy, and system transition. The network shows that the emerging challenges are interconnected and not standalone. Technical barriers intersect with users' psychological pressure, while low data literacy strengthens resistance toward system change. These findings affirm that the success of CAT-ACTFL implementation is not only determined by technological sophistication, but also by the readiness of human and institutional culture to accept change. Therefore, mentoring strategy, continuous training, and strengthening of internal communication become important factors in reducing challenges and smoothing the assessment transformation process in *pesantren*-based educational environments.

Overall, this research result section show that the implementation of CAT-ACTFL in *pesantren*-based higher education environment cannot be understood only as technical innovation in language assessment, but as transformation process that touch structural, pedagogical, cultural, and quality governance dimension simultaneously, where each theme that has been presented display close interconnection between the rationality of ACTFL usage, system development mechanism, integration with quality assurance, its impact on learning practice, the dynamic of challenges, until the process of *pesantren* cultural adaptation, and the findings indicate that CAT-ACTFL function as meeting point between global standard and local reality.

Table 2. Thematic Structure of Findings.

Thematic domain	Analytical Focus	Interpretative Description
Rationality & Orientation of ACTFL Usage	Conceptual Legitimacy	Reflects the institutional justification for adopting ACTFL as a credible, standardised, and globally recognised framework to ensure objective, accountable, and measurable Arabic proficiency assessment.
Implementation and Development System	Structural Execution	Describes the technical and procedural mechanisms supporting CAT-ACTFL deployment, including system design, localisation, infrastructure reliability, and procedural standardisation.
Integration with QA	Governance Interface	Illustrates how CAT-ACTFL data is embedded within institutional quality assurance processes and functions as strategic input for curriculum monitoring and decision-making.
Impact on Teaching and Curriculum	Pedagogical Transformation	Highlights the shifts in instructional strategy, curriculum adjustment, and adaptive learning design based on proficiency-level insights.
Challenges of ACTFL Implementation	Transitional Barriers	Captures technical, cognitive, and organisational tensions encountered during early implementation stages.
Implication in the <i>Pesantren</i> Culture Correlation	Contextual Negotiation, Strategic Outlook	1. Represents the alignment between ACTFL standards and <i>pesantren</i> epistemology, reflecting a hybridisation of global assessment models within Islamic educational traditions. 2. Indicates long-term institutional direction toward data-informed governance, international benchmarking, and sustainable quality transformation.

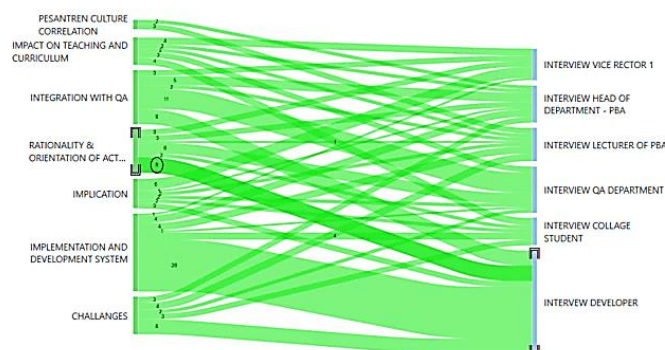
Source: Data Collection, 2026

The table presented above serves a purpose beyond merely summarising the thematic outcomes of the analysis. This serves as a framework for interpretation, influencing the organisation, understanding, and discussion of the findings presented in the Results section. Each

thematic domain is then examined through thoughtfully crafted narrative explanations, bolstered by direct empirical evidence obtained from the field data. (Virginia Braun & Victoria Clarke, 2022).

When viewed collectively, these seven domains create a unified thematic constellation that demonstrates how CAT-ACTFL transcends its traditional role as merely a technical assessment tool. Rather, it appears to operate as a cohesive system that shapes teaching methods, guides institutional decisions, and helps navigate cultural values within the educational environment. This thematic structure serves as the main framework for the analysis in the upcoming sections, offering a solid foundation for a thorough presentation and interpretation of the study's findings.

Figure 1. Co-Occurrence Flows Between Stakeholder Interviews Documents and Major Thematic Domains



Source: Data Collection, 2026

The co-occurrence analysis shows that each thematic area is represented with different levels of intensity among various institutional actors, highlighting a clear pattern of emphasis. The domain of implementation and development system stands out in the developer's interview data, showing the highest number of mentions. This suggests that concerns about system architecture, technical functionality, and operational reliability are mainly viewed as responsibilities tied to technology. In contrast, the theme of integration with QA is more prominently reflected in the views of quality assurance staff and senior leadership. This suggests that ACTFL is primarily viewed at this level as a strategic instrument for managing institutions and maintaining quality governance.

DISCUSSION

CAT-ACTFL as an Instrument of Governance in Global Education Management

The findings of this research show that CAT-ACTFL has expanded beyond its function as merely a language assessment tool and has started to play a role as a governance instrument in institutional education management, where assessment is no longer positioned only as a mechanism to measure student learning achievement (Azizah & Mardiana, 2024; Daulay, Haidir, & Firmansyah, 2024; Deesongkram, 2025; Rifkin, 2019), but has transformed into a strategic system that also directs decision-making processes, quality control, and the strengthening of institutional accountability (Rice, Pêgo, Collares, Kisielewska, & Gale, 2022).

In the global discourse of education management, the assessment system is increasingly understood as part of a governance mechanism that influences institutional behaviour, policy formulation, and academic performance management (Chung, Jeon, Lee, Lee, & Yoo, 2018), because data-based assessment not only functions as an evaluation tool, but also as a source of legitimacy and control that shapes accountability culture within higher education, and the implementation of CAT-ACTFL in this research demonstrates a similar pattern, particularly through its linkage with the quality assurance system and its role in curricular decision-making (Imawan, Retnawati, Haryanto, & Ismail, 2025).

The use of ACTFL as a normative reference framework also places the institution within the global education governance landscape that emphasises standardisation, benchmarking, and continuous performance measurement, where assessment becomes part of an accountability regime that encourages higher education institutions to demonstrate performance and credibility before the international academic community.

However, what differentiates CAT-ACTFL in this research is its adaptive character, because this system does not operate as rigid external control, but rather as a mechanism that can be negotiated and adjusted to pedagogical needs. (Kissling & O'Donnell, 2015), cultural values, and institutional strategic priorities, reflecting a shift from compliance-based governance toward adaptive governance. In practice, CAT-ACTFL functions as an institutional intelligence system that generates continuous data flow to support strategic planning, curriculum restructuring, and strengthening of academic policy, showing that assessment no longer stands as a separate administrative procedure, but becomes an important node within institutional governance architecture that still respects local context and institutional identity.

Reconceptualising Quality Assurance through Adaptive Assessment Systems

This section emphasises that the implementation of CAT-ACTFL does not merely strengthen the mechanism of academic evaluation but also encourages the occurrence of a paradigm shift in higher education quality assurance practices (Nagy, 2025). In global discourse, quality assurance is increasingly understood not only as a compliance process toward formal standards (León, Díaz, Rojas, Giraldo, & Gallegos, 2025), but as an institutional learning system that relies on a continuous feedback cycle (Kaddouri, Jmad, Azzimani, Mhamdi, & Abbadi, 2025). Models such as European Standards and Guidelines (ESG), Outcomes-Based Education (OBE), and Total Quality Management (TQM) in education emphasise the importance of authentic data and real-time monitoring as the foundation of quality improvement (Andriyani & Leksono, 2024; Halim, Mohamad, Hj Ali, Bakar, & Ujir, 2025; Rohmadi, Rahmat, & Ardianto, 2024).

An adaptive assessment system enables institutions to obtain a more granular and contextual picture of student performance. (Ihichr, Oustous, El Bouzekri, & Lahcen, 2024). The data no longer stops at the aggregate of final scores, but presents patterns of competence development that can be strategically followed up on. In this context, quality assurance shifts from merely administrative documentation into a living diagnostic process, where every assessment result becomes a source of information for learning improvement and policy decision-making. (Yasin, Utina, & Asrifan, 2025). Furthermore, CAT-ACTFL expands the scope of quality assurance from compliance orientation toward learning organisation orientation.

Assessment not only tests students, but also reflects the effectiveness of the curriculum, the relevance of teaching strategies, and consistency of standard implementation, so that quality is no longer interpreted as static achievement, but as a continuous negotiation process between targets, reality, and improvement strategies.

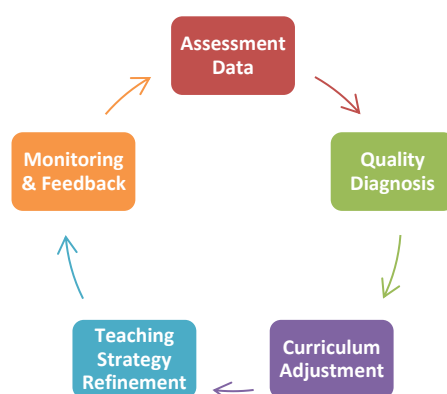


Figure 10. Adaptive Quality Assurance Loop Reinforced by CAT-ACTFL

From the perspective of global education management, this phenomenon strengthens the idea that adaptive assessment has the potential to become the core of a modern quality assurance system. It functions as a bridge between policy, teaching practice, and learning outcomes, while accelerating the decision-making cycle based on evidence. Within this framework, CAT-ACTFL not only supports internal quality reform, but also reinforces institutional position within the global education ecosystem that is increasingly competitive and audit-oriented (Halim et al., 2025). Therefore, this research proposes a new understanding of quality assurance as a practice that is no longer top-down in nature, but as an adaptive system that develops through interaction between assessment technology, quality policy, and pedagogical practice, showing that quality is not merely a target to be achieved, but a dynamic process that continues to be learned, interpreted, and refined.

Cultural Negotiation between Global Standards and Local Educational Identity

The findings of this study show that the integration of CAT-ACTFL does not occur as a linear process of transplantation of global standards, but rather as a complex field of cultural negotiation between the international assessment regime and *pesantren*-based educational identity. This situation highlights how global education management interacts with glocalisation. (De Velasco, 2007). It shows that global standards are not simply accepted; instead, they are reinterpreted in light of local values, traditions, and ways of knowing (Zapp, 2019). The ACTFL standard, rooted in the modern Western academic tradition, embodies a logic of rationality, an emphasis on measurable achievements, and a strong focus on performance. In contrast, *pesantren* education is grounded in values that prioritise adab, the blessings of the learning process, and the development of character as essential components of language acquisition (Hashim & Jemali, 2017).

In global literature, this condition is often understandable through the framework of hybrid governance and cultural translation. (Marshall, 2018), where educational institutions in non-Western regions re-contextualise international standards to remain aligned with their own value structures and institutional vision. CAT-ACTFL in this context functions not merely as a technical instrument, but as a dialogical medium between two knowledge systems: a global standard that focuses on the unification of competence and a *pesantren* tradition that is stagnant in uniqueness of value and spirituality (Badriyah & Wardi, 2026). The process of adapting indicators, selecting contextual test situations, and adjusting learning strategies shows that the institution does not fully submit to the logic of homogenization, but actively reshapes the meaning of assessment according to local needs (Sholehuddin, Hardini, Jaeni, Sofa, & Ho, 2025).

In a global perspective, this change is aligned with the concept of teacher agency and professional reflexivity, where educators are no longer placed as objects of policy but as subjects who have space to interpret data, shape strategy, and take pedagogical decisions based on evidence (Badriyah & Wardi, 2026). This condition indicates that assessment becomes not only an evaluation tool, but also a reflective instrument that strengthens the capacity of the lecturer in managing the learning process more consciously (Derakhshan, Molana, & Nazari, 2024). At the managerial level, the presence of an adaptive assessment system also strengthens the role of academic leadership in managing the learning process in a more structured and measurable manner (Derakhshan et al., 2024). Thus, learning management shifts from an administrative pattern toward a performance-based approach that is more strategic and responsive. (Estaji & Ghasvand, 2025).

The conceptual model presented in the table below shows that the lecturer is no longer positioned only as a routine executor of assessment procedures, but becomes an active actor who interprets assessment data, formulates learning strategies, and adjusts pedagogical interventions according to real classroom conditions (Soodmand Afshar & Movassagh, 2021). In this framework, the decision-making process is not based on habit or personal experience alone, but increasingly grounded on systematic reading of adaptive assessment results that reflect students' actual competence level (Fatkuroji, Wahyudi, Lateh, & Fadhilah, 2025; Katsaros, 2025). This

transformation also illustrates how pedagogical practice becomes more responsive and flexible, where the lecturer is required to continuously reflect on teaching effectiveness and learning progress (Karimli et al., 2025).

Toward a Conceptual Model of *Pesantren*-Aligned Language Assessment Governance

This section affirms that CAT-ACTFL in the context of UNIDA Gontor needs to be understood more precisely as a language assessment system that is integrated with the institutional vision of the *pesantren*, not as a moral-spiritual-based evaluation model or normative religious practice (Prasojo, Zulkarnain, Ferdiansyah, Puspasari, & Holidin, 2025). CAT-ACTFL functions purely as an instrument for measuring Arabic language proficiency; its implementation does not take place in a value-empty space. Arabic language at UNIDA is positioned as the language of the Qur'an and the main foundation for scientific development, as reflected in the Waqf Charter of Gontor, which emphasises the role of language as a pillar in the formation of Islamic intellectuality and civilisation. Therefore, the CAT-ACTFL assessment is not interpreted as a religious device, but as a strategic mechanism that supports the vision of *pesantren* education based on the language of revelation (Zarkasyi, 2020).

This research proposes the framework of *Pesantren*-Aligned Language Assessment Governance, which places CAT-ACTFL as a language assessment governance system that is aligned with the academic tradition of the modern *pesantren*. (Kawakip, 2020). This framework does not shift the function of CAT-ACTFL, but rather emphasises that the management of language assessment must support the role of Arabic as the main medium of transmission of Islamic knowledge and thought (M. Ritonga, Widayanti, Alrasi, & Halim, 2020). In this context, assessment is positioned as a strategic instrument that ensures the quality of language mastery is maintained in accordance with global standards, while still being relevant to the scientific orientation of UNIDA.

This conceptual model operates through three main axes: first, ACTFL standards as objective parameters for measuring language competence; second, the UNIDA governance system, grounded in the Waqf Charter as the framework of institutional direction (Zarkasyi, Mas'ud, Agung Hidayatullah, & Khakim, 2024); and third, the position of Arabic as the language of knowledge and the language of the Qur'an, which becomes the core of the academic structure of the *pesantren* campus (Muflihah, Hanifah, Thoha, & Haji Abdullah, 2025). The interaction of these three dimensions forms an assessment configuration that is not culturally neutral, but strategically institutional, and academically maintains language quality.

This research makes a significant contribution to the development of theory in the field of education management, both at the global level and the local level, especially in the study of assessment governance, quality assurance, and integration of international standards within a value-based education context (Marshall, 2018). The findings expand the understanding that so far tends to position assessment as a technocratic device by emphasising that assessment can function as a strategic mechanism that shapes power relations, knowledge production, and direction of institutional policy. (Ithichr et al., 2024).

Table 3. Analytical Summary of the Research Discussion

Discussion Domain	Research Findings	Theoretical Analysis	Research Contribution
CAT-ACTFL as a Global Governance Instrument	CAT-ACTFL at UNIDA has evolved beyond a mere language assessment tool into a strategic governance instrument used for institutional decision-making, quality control,	Aligns with global education governance discourse, where data-driven assessment acts as a source of institutional legitimacy and shapes a culture of accountability in higher education	Demonstrates a paradigm shift from rigid, compliance-based external control to flexible and responsive <i>adaptive governance</i>

	and accountability		
Reconceptualising Quality Assurance Systems	CAT-ACTFL data is deeply embedded within the PPEPP cycle, transforming quality assurance from static administrative documentation into a live, real-time diagnostic process.	Supports modern quality frameworks such as Outcomes-Based Education (OBE) and Total Quality Management (TQM). Which advocate for granular monitoring of student performance	Expands the scope of quality assurance from a mere compliance orientation to a <i>learning organisation orientation</i> . Turning assessment into a dynamic evaluation of curriculum effectiveness.
Cultural Negotiation and Glocal Identity	The integration of Western-standard ACTFL guidelines undergoes a complex process of cultural negotiation, adapting to local pesantren values (such as <i>adab</i> , blessings, and character building) through customised text selection and indicators.	Grounded in the frameworks of <i>hybrid governance, cultural translation</i> , and glocalisation within international educational development literature	Rejects standard global homogenization; highlights <i>teacher agency</i> and professional reflexivity in reshaping global assessment parameters to match local institutional realities
Toward a Pesantren-Aligned Assessment Model	CAT-ACTFL functions as an objective proficiency measure that operates without eroding the Gontor Waqf Charter, which structurally positions Arabic as the pillar of Islamic intellectuality and revelation.	Formulates a newly proposed conceptual framework called <i>Pesantren-Aligned Language Assessment Governance</i>	Offers a unique theoretical contribution showing that international benchmarking can coexist with and reinforce value-driven, non-Western institutional epistemologies.

Source: Data Collection, 2026

CONCLUSION

This research demonstrates that the implementation of CAT-ACTFL at Universitas Darussalam Gontor has successfully evolved from a technical assessment instrument into a strategic driver of institutional governance. Addressing the core research focuses, the study reveals that the system successfully bridges global benchmarks with local institutional values by centring its question item bank on socio-religious themes familiar to the Islamic boarding school ecosystem. Furthermore, the empirical findings show that CAT-ACTFL data have been deeply integrated into the university's internal quality assurance cycle (PPEPP), shifting the educational evaluation paradigm from traditional, generalised classroom teaching toward data-driven, differentiated instruction based on students' actual proficiency levels. This integration highlights how standards-based adaptive testing effectively reshapes classroom instruction management, academic policy orientation, and long-term curriculum planning.

Theoretically, this study enriches the literature on education management by offering the framework of *Pesantren-Aligned Language Assessment Governance*, proving that international benchmarking can coexist with and reinforce non-Western educational identity. However, a notable limitation of this study is its primary focus on qualitative stakeholder perceptions and systemic implementation, without conducting a deep statistical analysis on the testing tool's psychometric properties. To address this constraint, future research is strongly recommended to conduct quantitative validity testing, reliability testing, and item characteristic analysis of the CAT-ACL database through sophisticated psychometric approaches, such as item response theory (IRT) and the Rasch model, to scientifically ensure and refine the measurement accuracy.

Acknowledgment

The research team extends sincere gratitude to the Ministry of Higher Education, Science, and Technology (KemendiktiSaintek) for providing funding for this applied research. Appreciation is also directed to the Language Development Directorate of Darussalam Gontor University for serving as a research partner.

REFERENCES

- Al-Assaf, D. M. (2025). Using the sense relations in the classrooms of teaching Arabic to non-native speakers. *Dirasat: Human and Social Sciences*, 52(3), 5522. <https://doi.org/10.35516/hum.v52i3.5522>
- Almelhes, S. (2024). Enhancing Arabic Language Acquisition: Effective Strategies for Addressing Non-Native Learners' Challenges. *Education Sciences*, 14(10), 1116. <https://doi.org/10.3390/educsci14101116>
- Alsubaie, F. B. S. M. (2022). Arabic language use and Islamic education in teaching value to graduates in Saudi universities. *Ijaz Arabi Journal of Arabic Learning*, 5(3), 903–912. <https://doi.org/10.18860/ijazarabi.v5i3.18354>
- Amalia, E. R. (2026). Religious Pluralism in Indonesia's National Curriculum: A Critical Review of Interfaith Religious Education. *Attaqwa: Jurnal Ilmu Pendidikan Islam*, 22(2), 162–178. <https://doi.org/10.54069/attaqwa.v22i2.1169>
- Anderson, C. A., Leahy, M. J., DeValle, R., Sherman, S., & Tansey, T. N. (2014). Methodological application of multiple case study design using modified consensual qualitative research (CQR) analysis to identify best practices and organizational factors in the public rehabilitation program. *Journal of Vocational Rehabilitation*, 41(2), 87–98. <https://doi.org/10.3233/JVR-140709>
- Andriyani, F., & Leksono, A. A. (2024). Effective Management in the Implementation of Link and Match Curriculum for Quality Graduates. *Journal of Education and Learning Innovation*, 1(2), 151–160. <https://doi.org/10.59373/jelin.v1i2.45>
- Arif, M., Zekkari, M., & Addawiyah, R. (2025). Enhancing Santriwati's Islamic Literacy Through the LIT-CON Model of Women's Fiqh Learning in Pesantren. *Tafkir: Interdisciplinary Journal of Islamic Education*, 6(4), 966–992. <https://doi.org/10.31538/tijie.v6i4.2216>
- Azizah, I., & Mardiana, D. (2024). Learning Transformation: Increasing Student Achievement through Discovery Learning. *Dirasah International Journal of Islamic Studies*, 2(2), 155–166. <https://doi.org/10.59373/drs.v2i2.42>
- Badriyah, L., & Wardi, M. (2026). Policy analysis of Islamic educational institutions in facing the challenges of society 5.0: Innovation, learning, and technology-based infrastructure. *Multidisciplinary Reviews*, 9(1). Scopus. <https://doi.org/10.31893/multirev.2026012>
- Baihaqi, M. R., Amaliyah, H., Awaliyah, Y. S., Khoerunnisa, S. P., & Laksono, B. A. (2023). Analisis Swot Kebijakan Pemerintah Dalam Upaya Mencegah Kekerasan Seksual Di Lingkungan Pesantren. *Attaqwa: Jurnal Ilmu Pendidikan Islam*, 19(2), 181–191. <https://doi.org/10.54069/attaqwa.v19i2.591>
- Bogdan, R., & Biklen, S. K. (1998). *Qualitative research for education: An introduction to theory and methods* (3rd ed). Boston: Allyn and Bacon.
- Braun, V., & and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Chung, B. G., Jeon, I. S., Lee, R. H., Lee, I., & Yoo, S. S. (2018). Global Governance of Education and Training: As reviewed from Jomtien via Incheon to New York. *Asia Pacific Education Review*, 19(3), 319–336. Scopus. <https://doi.org/10.1007/s12564-018-9544-7>
- Creswell, J. W. autor. (2018). Research design qualitative, quantitative, and mixed methods approaches.

- Daulay, B., Haidir, & Firmansyah. (2024). The influence of managerial competence and achievement motivation on the leadership effectiveness of high school principals. *Cakrawala Pendidikan*, 43(2), 411–421. <https://doi.org/10.21831/cp.v43i2.61912>
- De Curtò, J., De Zarzà, I., Fervier, L. S., Sanagustín-Fons, V., & Calafate, C. T. (2025). An institutional theory framework for leveraging large language models for policy analysis and intervention design. *Future Internet*, 17(3), 96. <https://doi.org/10.3390/fi17030096>
- De Velasco, F. D. (2007). Religion, identity and education for peace: Beyond the dichotomies: Confessional/non-confessional and global/local. *British Journal of Religious Education*, 29(1), 77–87. Scopus. <https://doi.org/10.1080/01416200601037544>
- Deesongkram, M. (2025). A Bibliometric Analysis of STEAM Integration in Chemistry Learning at the Secondary Education Level. *Online Learning In Educational Research (OLER)*, 5(2), 289–303. <https://doi.org/10.58524/oler.v5i2.555>
- Demir, H., & Gelbal, S. (2025). A systematic review on computerized adaptive testing. *Erzincan Üniversitesi Eğitim Fakültesi Dergisi*, 27(1), 137–150. <https://doi.org/10.17556/erziefd.1577880>
- Derakhshan, A., Molana, K., & Nazari, M. (2024). “I wonder who I am when my assessment ideas are not respected”: A case-study inquiry into the role of institutional policies in English language teachers’ assessment identity. *Asia Pacific Journal of Education*. Scopus. <https://doi.org/10.1080/02188791.2024.2328237>
- Estaji, M., & Ghiasvand, F. (2025). Beyond the gradebook: Embracing the potentials of teacher assessment identity (TAI) in (re)shaping English language professors’ professional development and success in higher education. *Language Testing in Asia*, 15(1). Scopus. <https://doi.org/10.1186/s40468-025-00391-8>
- Fatah, M., Khoiruddin, M. A., & Fakhruddin, F. M. (2025). Pesantren-Based Character Education: The Dzuriyah Strategy in Shaping the Character of Santri Ndalem. *Interdisciplinary Journal of Social Sciences*, 2(2), 65–75. <https://doi.org/10.59373/ijoss.v2i2.191>
- Fatkuroji, Wahyudi, Lateh, F., & Fadhillah, A. R. (2025). Evaluation of Academic Information Systems in Realizing Good University Governance. *Munaddhomah: Jurnal Manajemen Pendidikan Islam*, 6(1), 139–154. <https://doi.org/10.31538/munaddhomah.v6i1.1675>
- Haerullah, I. S., Jundi, M., Hasibuan, R., Sari, R., & Asmarita, D. (2024). Constructivism in Arabic language pedagogy: An exploration through Islamic higher education settings. *ELOQUENCE: Journal of Foreign Language*, 3(1), 10–25. <https://doi.org/10.58194/eloquence.v3i1.1495>
- Halim, R. A., Mohemad, R., Hj Ali, N., Bakar, A. A., & Ujir, H. (2025). Validation of an Adaptive Decision Support System Framework for Outcome-Based Blended Learning. *International Journal of Advanced Computer Science and Applications*, 16(2), 1210–1219. Scopus. <https://doi.org/10.14569/IJACSA.2025.01602119>
- Harsch, C., & Malone, M. E. (2020). Language Proficiency Frameworks and Scales. In P. Winke & T. Brunfaut (Eds), *The Routledge Handbook of Second Language Acquisition and Language Testing* (1st edn, pp. 33–44; By P. Winke & T. Brunfaut). New York: Routledge, 2020. | Series: The Routledge handbooks in second language acquisition: Routledge. <https://doi.org/10.4324/9781351034784-5>
- Hasan, M. S., Mujahidin, Azizah, M., & Solechan. (2024). Fostering A Moderate Attitude in Sufi-Based Pesantren Culture. *At-Tadzkir: Islamic Education Journal*, 3(2), 171–188. <https://doi.org/10.59373/attadzkir.v3i2.66>
- Hashim, I., & Jemali, M. (2017). Key aspects of current educational reforms in Islamic educational schools. *Global Journal Al-Thaqafah*, 7(1), 49–57. Scopus. <https://doi.org/10.7187/gjat12620170701>

- Hussain, S. (2025). Islamic Education in Contemporary Central Asia: Reviewing Ideas, Curriculum and Pedagogy. *Dirasah International Journal of Islamic Studies*, 3(1), 64–76. <https://doi.org/10.59373/drs.v3i1.47>
- Ihichr, A., Oustous, O., El Bouzekri, Y., & Lahcen, A. A. (2024). A Systematic Review on Assessment in Adaptive Learning: Theories, Algorithms and Techniques. *International Journal of Advanced Computer Science and Applications*, 15(7), 855–868. Scopus. <https://doi.org/10.14569/IJACSA.2024.0150785>
- Imawan, O. R., Retnawati, H., Haryanto, H., & Ismail, R. (2025). Innovations in Assessment Methods: Computerized Adaptive Testing (CAT) for Sustainable Energy Efficiency. In Mansour Y., Abdelhadi A., Subramaniam U., Mustaffa Z., Al-Atroush M., & Abowardah E. (Eds), *Lect. Notes Civ. Eng.: Vol. 557 LNCE* (pp. 161–168). Springer Science and Business Media Deutschland GmbH. Scopus. https://doi.org/10.1007/978-981-97-8348-9_20
- Isnaeni, F., Budiman, S. A., Nurjaya, N., & Mukhlisin, M. (2025). Analysis of the Readiness for Implementing Deep Learning Curriculum in Madrasah from the Perspective of Educators. *Attadrib: Jurnal Pendidikan Guru Madrasah Ibtidaiyah*, 8(1), 15–30. <https://doi.org/10.54069/attadrib.v8i1.841>
- Jill, O., Gilkerson, G., Jane, D., Hannon, B., & Steven, H. (2023). Adaptive learning to develop and apply to Arabic language learning in higher education. *Al-Hijr: Journal of Adulearn World*, 2(2), 136–149. <https://doi.org/10.55849/alhijr.v2i2.544>
- Junaidi, Afrinaldi, Artika, F. S., Al-Momani, M. O., & Hulwana, H. (2025). Islamic Moderatism in Curriculum Development Of Islamic Educational Institutions in Aceh. *Jurnal Pendidikan Agama Islam*, 22(1), 43–56. <https://doi.org/10.14421/jpai.v22i1.8313>
- Kaddouri, M., Jmad, S., Azzimani, T., Mhamdi, K., & Abbadi, Z. (2025). Transforming Assessment Practices: Effective Digital Tools for Meaningful Feedback. In *Digit. Tools and Platf. For Eff. And Pers. Learn.* (pp. 155–202). IGI Global. Scopus. <https://doi.org/10.4018/979-8-3373-6013-3.ch006>
- Karimli, V. M., Khudaverdiyeva, T. S., Huseynova, F., İsmayilli, F. Z., Aliyeva, K. M., Aliyarova, N. N., ... Narayana, V. S. K. (2025). The Role of Mobile Computing in Adaptive Testing for English Language Learners: Personalizing Assessment to Improve Outcomes. *Forum for Linguistic Studies*, 7(6), 149–160. Scopus. <https://doi.org/10.30564/fls.v7i6.9663>
- Katsaros, K. K. (2025). Gen Z Tourism Employees' Adaptive Performance During a Major Cultural Shift: The Impact of Leadership and Employee Voice Behavior. *Behavioral Sciences*, 15(2), 171. <https://doi.org/10.3390/bs15020171>
- Kawakip, A. N. (2020). Globalization and Islamic Educational Challenges: Views from East Javanese Pesantren. *Ulumuna*, 24(1), 105–131. <https://doi.org/10.20414/ujs.v24i1.385>
- Kissling, E. M., & O'Donnell, M. E. (2015). Increasing language awareness and self-efficacy of FL students using self-assessment and the ACTFL proficiency guidelines. *Language Awareness*, 24(4), 283–302. <https://doi.org/10.1080/09658416.2015.1099659>
- León, G. M. O., Díaz, A. R., Rojas, E. M., Giraldo, L. F. G., & Gallegos, A. (2025). Evaluating quality assurance in education: A systematic literature review. *TPM - Testing, Psychometrics, Methodology in Applied Psychology*, 32(S4), 542–557. Scopus. Retrieved from Scopus.
- Mahmudi, I., Bujanah, H., Ismail, M., Saifulloh, A., Fatawi, I., & Hasanah, M. (2025). Analysis Of Online Arabic Language Proficiency Test Questions Based On Bloom's Cognitive Taxonomy. *Ijaz Arabi Journal of Arabic Learning*, 8(2). <https://doi.org/10.18860%20/ijazarabi.%20V8i2.29595>
- Mahmudi, I., Ismail, M., Resmalasari Dewi, N. R., Istibsyaroh, N., Islamanda, H. S., & El-Fani, A. R. (2024). Designing an arabic vocabulary test based on the ACTFL (american council on the teaching of foreign languages) standards for the novice level. *Abjadia: International Journal of Education*, 9(3), 759–766. <https://doi.org/10.18860/abj.v9i3.29055>

- Marshall, K. (2018). Global education challenges: Exploring religious dimensions. *International Journal of Educational Development*, 62, 184–191. Scopus. <https://doi.org/10.1016/j.ijedudev.2018.04.005>
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. SAGE.
- Miller, D. C., & Salkind, N. J. (2002). *Handbook of research design & social measurement* (6th ed). Thousand Oaks (Calif.): Sage.
- Msayer, M. E., Aoula, E.-S., & Bouihi, B. (2024). Artificial intelligence in computerized adaptive testing to assess the cognitive performance of students: A Systematic Review. 2024 *International Conference on Intelligent Systems and Computer Vision (ISCV)*, 1–8. Fez, Morocco: IEEE. <https://doi.org/10.1109/ISCV60512.2024.10620092>
- Muflihah, Hanifah, U., Thoha, M., & Haji Abdullah, H. R. B. (2025). Management of developing interactive multimedia-based Arabic teaching materials: Enhancing learning for diverse students at Indonesian Islamic universities. *Munaddhomah: Jurnal Manajemen Pendidikan Islam*, 6(1), 82–98. <https://doi.org/10.31538/munaddhomah.v6i1.1469>
- Musliadi, M., Triyono, S., & Jamilah, J. (2024). Enhancing speaking agility: Unveiling Indonesian lecturers' hybrid teaching experiences in oral communication skills. *Language Learning in Higher Education*, 14(2), 401–433. <https://doi.org/10.1515/cercles-2024-0018>
- Nagy, G. (2025). Performance Measurement and Quality Assurance in Higher Education: Application of DEA, AHP, and Bayesian Models. *Trends in Higher Education*, 4(3). Scopus. <https://doi.org/10.3390/higheredu4030054>
- Oladele, J. I., & Ndlovu, M. (2021). A review of standardised assessment development procedure and algorithms for computer adaptive testing: Applications and relevance for fourth industrial revolution. *International Journal of Learning, Teaching and Educational Research*, 20(5), 1–17. <https://doi.org/10.26803/ijlter.20.5.1>
- Ozan, O., Ozarslan, Y., & Calisir Zenci, S. (2025). Towards an adaptive language MOOC: Examining differences of language error patterns across cultural domains. *Turkish Online Journal of Distance Education*, 26(2), 16–38. <https://doi.org/10.17718/tojde.1452717>
- Pransiska, T., Sugiyono, S., Widodo, S. A., & Sayyid, W. A. M. (2024). Al-Kutub Al-Madrasiyah Al-‘Arabiyah Fi Indūnisiyā Min Mazūri Waṭanīyin Wa ‘Ālamīyin. *Jurnal Ilmiah Islam Futura*, 24(2), 488. <https://doi.org/10.22373/jiif.v24i2.14964>
- Prasojo, E., Zulkarnain, Z. P., Ferdiansyah, J., Puspasari, D., & Holidin, D. (2025). Infusing indigenous Islamic values into Western-style public administration in Indonesia: The role of pesantren institutions. In W. Drechsler, S. Chafik, & R. Kattel (Eds), *Islamic Public Value* (pp. 308–334). Edward Elgar Publishing. <https://doi.org/10.4337/9781035333646.00026>
- Rice, N., Pêgo, J. M., Collares, C. F., Kisielewska, J., & Gale, T. (2022). The development and implementation of a computer adaptive progress test across European countries. *Computers and Education: Artificial Intelligence*, 3. Scopus. <https://doi.org/10.1016/j.caeai.2022.100083>
- Rifkin, B. (2019). The power of performance-based assessment: Languages as a model for the liberal arts enterprise. In *Educ. Ling.* (Vol. 37, pp. 15–21). Springer Science+Business Media B.V. Scopus. https://doi.org/10.1007/978-3-030-01006-5_2
- Ritonga, M., Widayanti, R., Alrasi, F., & Halim, S. (2020). Analysis of Arabic Language Learning at Higher Education Institutions with Multi-Religion Students. *Universal Journal of Educational Research*, 8(9), 4333–4339. Scopus. <https://doi.org/10.13189/ujer.2020.080960>
- Ritonga, Mahyudin, Mudinillah, A., Wasehudin, W., Julhadi, J., Amrina, A., & Shidqi, M. H. (2024). The effect of technology on Arabic language learning in higher education. *Journal of Education and Learning (EduLearn)*, 18(1), 121. <https://doi.org/10.11591/edulearn.v18i1.20867>

- Rohmadi, A., Rahmat, & Ardianto, A. (2024). Management of Facilities and Infrastructure in Improving the Quality of Learning. *Journal of Education and Learning Innovation*, 1(2), 161–173. <https://doi.org/10.59373/jelin.v1i2.55>
- Safeer, A. (2025). Evaluating an ai-driven computerized adaptive testing platform for psychological assessment: A randomized controlled trial. *International Journal of Innovative Research in Computer and Communication Engineering*, 13(05), 9439–9450. <https://doi.org/10.15680/IJIRCCE.2025.1305005>
- Sahayu, W., Triyono, S., & Ungu, N. K. (2026). Digital Technology in Differentiated English Language Teaching: A Systematic Review with Bibliometric Insights. *Indonesian Journal of Science and Technology*, 11(1), 167–188. <https://doi.org/10.17509/ijost.v11i2.89685>
- Sartini, S., Triyono, S., & Triastuti, A. (2025). Unveiling maritime English communication needs for seafarer: Strategic reformation for classroom instructional design. *Maritime Technology and Research*, 7(2). <https://doi.org/10.33175/mtr.2025.271921>
- Sholehuddin, M. S., Hardini, I., Jaeni, M., Sofa, E. M., & Ho, T. T. H. (2025). The Role of Local Wisdom in Shaping Internationalisation Strategies of Islamic Higher Education Institutions in Indonesia. *Intellectual Discourse*, 33(2), 607–634. Scopus. <https://doi.org/10.31436/id.v33i2.2315>
- Soodmand Afshar, H., & Movassagh, H. (2021). Towards a Critical Language Teacher Identity: Contributions of a Critical Teacher Education Course. *Applied Research on English Language*, 10(2). <https://doi.org/10.22108/are.2021.127277.1682>
- Tayyara, A. el-Rahman. (2022). The practicability of proverbs in teaching Arabic language and culture. *Language Teaching Research*, 26(4), 799–819. <https://doi.org/10.1177/1362168819895253>
- Virginia Braun & Victoria Clarke. (2022). *Thematic analysis: A practical guide*. London ; Thousand Oaks, California: SAGE.
- Wang, C., Medaglia, R., & Zheng, L. (2018). Towards a typology of adaptive governance in the digital government context: The role of decision-making and accountability. *Government Information Quarterly*, 35(2), 306–322. <https://doi.org/10.1016/j.giq.2017.08.003>
- Yasin, Z., Utina, S. S., & Asrifan, A. (2025). Personalized Learning Through Intelligent Assessment Systems: Enhancing Student Engagement. In *Adv. Soc. 5.0 Through AI-Driven Curric. Innov.* (pp. 163–194). IGI Global. Scopus. <https://doi.org/10.4018/979-8-3373-1062-6.ch006>
- Yin, R. K. (2018). *Case Study Research and Applications: Design and Methods* (6th edn). Singapore: SAGE Publications, Inc.
- Zapp, M. (2019). International organizations, educational development, and conflict: A (world) cultural perspective. *Prospects*, 47(4), 321–337. Scopus. <https://doi.org/10.1007/s11125-019-09442-4>
- Zarkasyi, H. F. (2020). Imam Zarkasyi's Modernization of Pesantren in Indonesia. *Qudus International Journal of Islamic Studies (QIJIS)*, 8(1), 161–200. <https://doi.org/10.21043/qijis.v8i1.5760>
- Zarkasyi, H. F., Mas'ud, F., Agung Hidayatullah, R., & Khakim, U. (2024). Strategy of Indonesian Pesantren University in Achieving Competence of Student: A Grounded Research at UNIDA Gontor. *KnE Social Sciences*, 2024, 119–138. <https://doi.org/10.18502/kss.v9i6.15260>