Cultivating Linguistic Competence in the Spirit of Islamic Higher Education: A CSE-UCLA-Based Evaluation of the Arabic Language Development Program

Ihwan Mahmudi[1], Taufik Rizki Sista[2], Ahmad Saifulloh[3], Siti Fadilah[4]

[1, 4]Postgraduate Program, Universitas Darussalam Gontor, Indonesia
[2, 3]Department of Islamic Education, Universitas Darussalam Gontor, Indonesia
*Email: ihwanm@unida.gontor.ac.id

ARTICLE INFO	ABSTRACT
Keywords: Higher Education Islamic Education Arabic Language-Education Program Evaluation	**Purpose -** This research aims to evaluate the effectiveness of the Arabic Language Development Program at the University of Darussalam Gontor by reviewing the alignment between institutional objectives, program planning, field implementation, and achieved outcomes within the context of *pesantren*-based Islamic higher education. **Methodology -** This research uses a Sequential Explanatory Mixed-Method design, beginning with the quantitative stage through the use of a structured evaluation checklist to assess four components of the CSE-UCLA model (Need Assessment, Planning, Implementation, and Outcomes), then continuing with the qualitative stage through semi-structured interviews, observation, and document analysis to explain and validate the quantitative results. The research participants consist of relevant stakeholders in the management and implementation of the Arabic language program, with qualitative data analysis conducted using a thematic approach. **Findings -** The findings show that the Need Assessment and Planning aspects are in the category of very good, with strong alignment between the institutional vision, program rationale, and curriculum design. However, there is a gap in the Implementation stage, which is influenced by variations in student abilities, the dynamics of academic–pesantren activities, and the limitations of productive assessment instruments. Nevertheless, the summative evaluation indicates that the program has a significant impact on improving students' academic competence and strengthening the institution's scientific identity. **Contribution -** This research contributes to the expansion of the application of the CSE-UCLA evaluation model in the context of Islamic higher education, and provides policy recommendations for strengthening performance-based assessment, harmonizing academic–*pesantren* environments, and improving program quality in a sustainable manner.

INTRODUCTION

The development of information and communication technology has led to a significant transformation in the world of education, necessitating continuous modernization in learning design, curriculum development, and academic governance by higher education institutions (Juliana et al., 2023). In the context of globalization, language competence becomes a strategic element that determines academic competitiveness, access to international knowledge, and the ability to participate in cross-cultural scientific dialogue (Quainoo et al., 2022; Surur, 2022). The Arabic language has become increasingly relevant, both as the language of Islamic scholars and as a medium for global academic interaction (Fitriani, 2021). Therefore, Islamic higher education institutions have epistemological and pedagogical responsibilities to ensure that the teaching of the Arabic language is not merely formalistic, but truly supports students' academic abilities in reading original texts, writing scientific works, and communicating within academic contexts (Asril & Hanafi, 2021; Salmas et al., 2024a).

The University of Darussalam Gontor (UNIDA Gontor) has become one of the institutions that consistently places the Arabic language as an academic identity, through the implementation of a bilingual system in learning, academic writing, and daily communication (Rohmah et al., 2024). However, initial observations and interviews indicate that not all faculties have the same needs or the same level of Arabic language use. Otherwise, non-Islamic faculty students show a relatively low utilization of Arabic, especially in daily academic activities. Many students view the learning of Arabic as merely an administrative requirement, rather than an academic need, particularly because some of their courses predominantly use Indonesian and English. Additionally, low intrinsic motivation was found among students, particularly female students in the dormitory environment, in actively using the Arabic language (Mahmudi et al., 2025). This imbalance indicates the presence of a gap between institutional objectives and learning practices at the student level.

Research on Arabic language development in Islamic higher education institutions has been widely conducted. Yet, most studies focus separately on learning evaluation, curriculum, or teaching strategies (Mujahid, 2021). No research has comprehensively evaluated a language development program using the CSE-UCLA evaluation model in a pesantren-based university. This is especially true for studies including all faculties, including non-Islamic ones. Thus, a significant research gap exists. No study has comprehensively assessed the effectiveness of the Arabic Language Development Program at UNIDA Gontor, covering the entire process from need assessment and planning to implementation and outcomes. This is critical for measuring the alignment between institutional vision and actual realization in the field (Zarkasyi et al., 2024).

This condition gives rise to the formulation of the main research problem of this study: To what extent has the Arabic Language Development Program at UNIDA Gontor fulfilled institutional needs, been systematically planned, effectively implemented, and had a significant academic impact on students across faculties? This question requires an evaluation based on the CSE-UCLA model because the model provides a comprehensive framework that connects context, program design, implementation processes, and final outcomes in a structured manner (Fathurrozi & Pamungkas, 2021).

Therefore, the purpose of this study is to comprehensively evaluate the effectiveness of the Arabic Language Development Program at UNIDA Gontor using the CSE-UCLA model, by assessing four main components: Need Assessment, Planning and Evaluation, Implementation, and Summative Outcomes. Theoretically, this study contributes to expanding the use of the CSE-UCLA evaluation model within the context of *pesantren*-based Islamic higher education. Practically, this study offers policy recommendations for UNIDA Gontor to enhance the system of Arabic language teaching and evaluation, particularly for non-Islamic faculties, ensuring the program serves not only as an administrative tool but also as an integral component of academic competence.

METHODOLOGY

Research Design

This study uses a Sequential Explanatory Mixed-Method Design. This staged mixed-methods approach begins with the quantitative phase, which aims to obtain objective evaluative scores through a structured evaluation checklist. This is followed by a qualitative phase, which consists of semi-structured interviews and structured observations. These qualitative methods are used to explain, confirm, deepen, and validate the quantitative findings (Sugiyono, 2020). The design was chosen because evaluating the Arabic Language Development Program at UNIDA Gontor requires both numerical measurements of performance indicators and interpretive understanding of the program's context. Important contextual factors include *pesantren* cultural dynamics, student motivation, instructor perceptions, and operational field conditions. The study was conducted in several strategic units of the University of Darussalam Gontor, each directly linked to program implementation. The Arabic Language Education Study Program served as the academic and pedagogical center.

Participant

Participants were chosen through purposive sampling and grouped into three categories. These were individuals directly involved in the policy, implementation, and use of the Arabic Language Development Program. Informants included general policy stakeholders, the Directorate of the Language Center (as the program planner and policymaker), Arabic Language Education lecturers (as technical executors of instructional activities), and students (as program users). All participants contributed to the quantitative phase by completing a structured evaluation checklist. They also participated in the qualitative phase through semi-structured interviews to help explain and confirm the quantitative findings.

Table 1. Research Participants

Informant Category	Role	Number	Phase
DPB	Program Implementer	4	Quantitative & Qualitative
LPM	Quality Assurance	1	Qualitative
LD	Program Implementer	5	Quantitative
HD	Program Implementer	1	Quantitative & Qualitative
ST	Program Users	9	Quantitative
VR	General Policy Holder	1	Qualitative

Data Collection

Data collection was conducted in two stages, in accordance with the Sequential Explanatory Mixed-Method design (Sugiyono, 2020). The first stage employed a quantitative approach through a structured evaluation checklist completed by all participants to obtain evaluative scores for each component of the CSE-UCLA model. The second stage employed a qualitative approach, utilizing semi-structured interviews, structured observations, and document analysis, to explain, confirm, and deepen the quantitative findings, thereby enhancing the validity and comprehensiveness of the program evaluation. These techniques were used to triangulate to ensure the reliability of the data (Creswell, 2011).

Table 2. Data Collection Techniques

Stage	Technique	Type of Data	Purpose
Quantitative	Structured Evaluation Checklist	Indicator scores (1–4)	Measuring program effectiveness objectively

Stage	Technique	Type of Data	Purpose
Qualitative	Semi-structured interviews	Explanations & clarifications	Explaining and confirming the quantitative results
Qualitative	Document analysis	SOPs, schedules, reports, modules	Triangulation of planning and implementation

Table 3. Rating Scale Category Description

Score	Category	Description
4	Excellent	The indicator is fully fulfilled, consistent, and supported by strong evidence.
3	Good	The indicator is fulfilled, yet there are still minor aspects that are not optimal.
2	Fair	The indicator is only partially fulfilled; evidence is weak; implementation is inconsistent.
1	Poor	The indicator is not fulfilled; no evidence is available; implementation does not proceed.

Instrument

The structure of the instrument in this study was arranged in the form of a structured evaluation checklist, developed based on the four main components of the CSE-UCLA evaluation model: need assessment, program planning, program implementation, and program outcomes/improvement (Mardapi, 2012). Each component is elaborated into a number of indicators that represent the key aspects of program success, so that the entire checklist provides a comprehensive and systematic evaluative overview. For the needs assessment component, there are six indicators that assess the clarity of the program rationale, its alignment with the institutional vision, the formulation of objectives, and the identification of student needs. The planning evaluation component also comprises six indicators, covering the readiness of facilities, availability of instructors, curriculum quality, and completeness of implementation guidelines (Mardapi, 2008). The program implementation component includes seven indicators that assess the execution of formal and non-formal learning, consistency of activities, student participation, and monitoring of implementation. Meanwhile, the program outcomes/improvement component consists of seven indicators that evaluate students' competence attainment, the program's impact on faculties, and the mechanisms of improvement and program sustainability. Each indicator is assessed using a 1–4 scale (Poor–Excellent), accompanied by columns for supporting evidence and evaluator notes, allowing the instrument to generate not only objective quantitative scores but also enabling the researcher to transparently trace the basis of judgment. With this standardized structure, the checklist serves as the primary measurement tool for the quantitative phase and the basis for interpretation in the qualitative phase within the Sequential Explanatory design.

Data Analysis

Data analysis was conducted in two stages in accordance with the Sequential Explanatory Mixed-Method design. The process began with quantitative analysis as the primary stage and was followed by qualitative analysis as the explanatory stage. In the quantitative stage, each indicator in the structured evaluation checklist was assigned a score of 1 to 4. These scores were then summed and converted into percentages to determine the level of effectiveness for each CSE-UCLA component. The quantitative results served as the initial basis for identifying areas that required further explanation. The qualitative stage was then carried out through the analysis of interviews, observations, and documents using thematic analysis. This process included familiarization with the data, initial coding, identifying patterns of meaning, grouping codes into themes, and reviewing and interpreting themes to explain phenomena from the quantitative findings (Abu Aser et al.,

2021). This qualitative stage aimed to provide contextual explanations, confirm quantitative results, and validate findings through the triangulation of sources and methods. Both stages were integrated systematically to ensure a comprehensive, objective, and accountable evaluation.

FINDINGS

To ensure that the results of the CSE-UCLA evaluation do not remain limited to quantitative analysis, this study formulated five main themes. These are the stages of Need Assessment, Planning, Evaluation, Implementation, Summative Evaluation, and Suggestions for Improvement. The themes represent the core components of the CSE-UCLA evaluation model and were intentionally selected as the framework for qualitative analysis to strengthen and validate the quantitative findings. Each theme was analyzed using the concept of Groundedness (GR), which indicates the frequency of quotation occurrences, and Density (GS), which indicates interconnections among themes in network analysis. In this way, the qualitative results can show whether the narrative patterns of the informant support, reinforce, or highlight aspects not fully captured by the numerical evaluation.

Table 5. Frequency of Quotation Occurrence by Theme and Informant

Theme (with total Gr; GS)	DPB (16)	HD (13)	LPM (5)	VR (24)	Total (56)
Need Assessment (Gr=17; GS=15)	4	4	3	6	17
Planning Evaluation (Gr=11; GS=9)	0	4	0	7	11
Implementation (Gr=10; GS=7)	4	0	1	5	10
Summative Evaluation (Gr=7; GS=6)	3	1	0	3	7
Suggestions for Improvement (Gr=11; GS=8)	2	4	0	5	11
Total per Informant	13	13	4	26	56

The results of the analysis show that the Need Assessment theme is the most dominant empirically (GR = 17) and at the same time the strongest conceptually (GS = 15). This indicates that all informants consistently emphasized the program's rationale, the urgency of improving students' abilities, and the institution's strategic direction in positioning the Arabic language as an academic identity. The Planning Evaluation theme (GR = 11; GS = 9) demonstrates the strength of planning at the managerial level, especially among the Vice Rector I and the Head of the Arabic Language Education Study Program, who frequently referred to competency standards, academic guidelines, and the adoption of international benchmarks. However, the Implementation theme (GR = 10; GS = 7) reveals the presence of a planning–implementation gap, marked by variations in students' abilities, limitations in productive assessment, the rhythm of *pesantren* activities, and conflicts with academic scheduling. The Summative Evaluation theme (GR = 7; GS = 6) and the Suggestions for Improvement theme (GR = 11; GS = 8) indicate that, although implementation challenges exist, the program's impact remains significant and generates a shared awareness to improve the quality of the evaluation system in the future.

The relationships among themes and each informant's contribution are visualized in the Sankey diagram. This diagram illustrates the flow of evaluation themes from their information sources—the interview documents on the right —to the evaluation themes on the left. Thicker flows indicate a higher Groundedness of a theme for an informant.
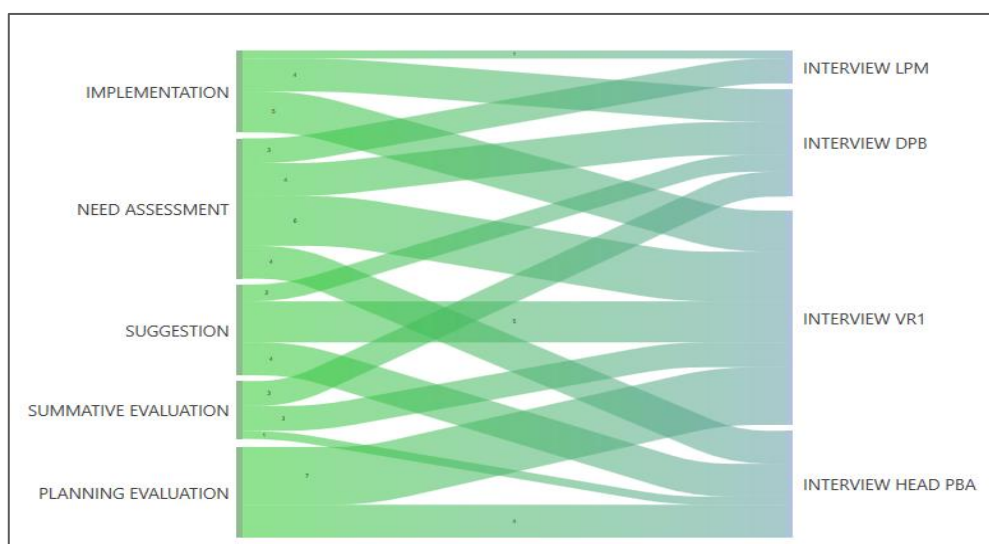
Figure 1. Qualitative Evidence Flow across CSE-UCLA Evaluation Components

Need Assessment

The following table presents the results of the Need Assessment aspect, which evaluates the relevance, rationale, and clarity of the Arabic Language Development Program's objectives. All components achieved a score of 100%. This confirms that the objectives, vision, and direction fully align with the institutional orientation of the University of Darussalam Gontor.

Table 6. Need Assessment Scores

Indicator	Score
Program rationale	100%
Alignment with vision and mission	100%
Clarity of program objectives	100%
Identification of implementation challenges	100%
Long-term strategic planning	100%
Average	100%

This alignment is consistent with the mandate of Law No. 12 of 2012. The law asserts that higher education institutions aim to develop learners who are knowledgeable, possess strong character, and have cultural and religious awareness. In the context of a *pesantren*-based university, mastery of the Arabic language becomes a principal foundation. This is because it serves as a key instrument for understanding Islamic literature, developing research based on primary sources, and supporting the Islamization of knowledge. Therefore, the Arabic language development program becomes a strategic element that cannot be separated from the institutional mission. These quantitative findings are further strengthened by empirical evidence from interviews. One of the informants emphasized the academic relevance of Arabic as a core institutional need:

"…the assessment for the Arabic language program is very clear, especially … its main scientific language." (HD)

Another informant explained the philosophical and institutional reasons for the Arabic language development program:

"The Arabic language program at UNIDA has a very strong rationale…philosophically and academically." (DPB)

An informant also stressed Arabic's foundational importance in all campus academic activity:

"Arabic language competence is not merely a technical skill, but it is the foundation …" (LPM)

A policy informant emphasized how the program fits the institutional vision:

"The Arabic language is an integral part of the Islamization of knowledge and the academic tradition of the pesantren." (VR)

Based on the findings above, it can be concluded that the Arabic Language Development Program is not only relevant but also strategically integrated with the University of Darussalam Gontor's mission. As a *pesantren*-based university, it is committed to the Islamization of knowledge and the strengthening of scholarly literacy rooted in the Arabic language. The figure below illustrates the flow of code relationships within the interview results related to the need assessment theme.
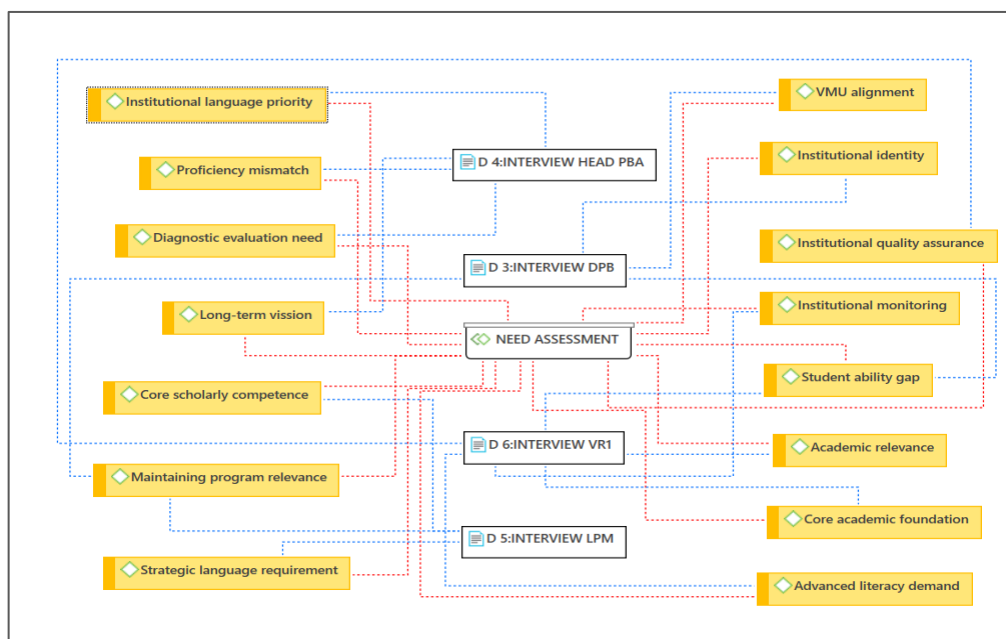
Figure 2. Network mapping of need assessment code co-occurrence analysis

Planning Evaluation

The evaluation of the planning aspect of the Arabic Language Development Program at the University of Darussalam Gontor shows that the overall quality of program planning falls into the good category, with an average score of 83.55%. The details of the evaluation results are presented in the following table:

Table 7. Planning Evaluation Score

Indicator	Score
Facilities and Infrastructure	83%
Student Engagement and Readiness	80%
Arabic Language Instructors	80%
Program Organizing Committee	75%
Curriculum Design and Content	100%
Academic Guidelines	83.3%
Average	83.55%

These results indicate that several planning components have performed very well, particularly curriculum design, which achieved a perfect score (100%). One informant provided a reinforcing statement:

"Our evaluation curriculum not only measures basic skills but also academic competencies …." (HD)

This strong curriculum demonstrates careful planning, focusing on outcomes that prepare students to write and defend a thesis in Arabic. This is further supported by another informant's statement:

"Evaluation planning must ensure that the instruments and competency standards are aligned with the demands of writing an Arabic-language thesis." (VR)

Nevertheless, the evaluation results indicate that the program organising committee obtained the lowest score (75%), which reflects the need to strengthen cross-unit coordination, particularly among the DPB, faculties, and the managers of course or matriculation activities. This finding is consistent with the interview results, which stated:

"… coordination across units is still not optimal, and this affects the consistency of … activities." (DPB)

Meanwhile, the aspects of facilities and infrastructure (83%) and academic guidelines (83.3%) are evaluated as good. Reinforcing this, one informant emphasised the critical role of formal documents as the foundation of planning:

"All standards and instruments must be included in the academic guidelines ..." (LPM)

Furthermore, the aspect of student readiness and engagement (80%) suggests that, despite some students' *pesantren* background, their academic readiness does not consistently meet university-level demands. One informant noted:

"… but they are not yet prepared to read … or write systematically in academic Arabic." **(HD)**

Overall, the strength of evaluation planning in curriculum, academic guidelines, and evaluation instruments is reinforced by observation and documentation. Field observations indicate that the evaluation blueprint guides the development of CAT examination items and matriculation activities. Additionally, official documents such as the 2025 Academic Guidelines, the DPB Evaluation Blueprint, and the Arabic Thesis Competency Standards underscore the consistency between planning and evaluation policy implementation.
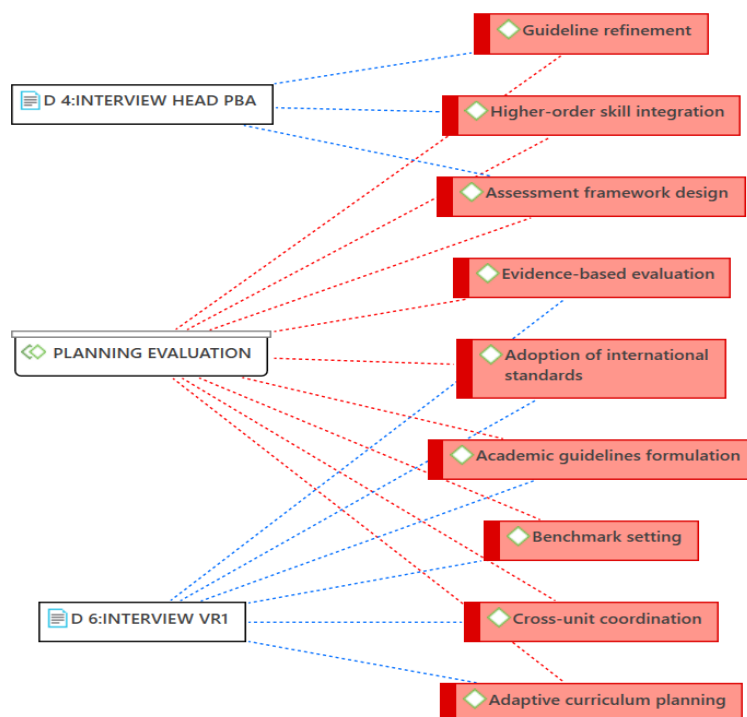
Figure 3. Network mapping of planning evaluation codes co-occurrence analysis

The network structure reveals that planning in this context is best understood as a major theme centered on establishing a high-quality, systematic, and evidence-based evaluation framework. Rather than focusing solely on schedules or administration, the theme reflects a comprehensive and adaptive approach, illustrated by interconnected clusters of codes.

Implementation

The evaluation results for the Implementation aspect of the Arabic Language Development Program at the University of Darussalam Gontor are presented in the following table:

Table 8. Implementation Evaluation Score

Indicator	Result
Instructional Planning of the Program	100%
Implementation of Learning Activities	54.25%
Program Evaluation	69.3%
Average	74.5%

Generally, the average score of 74.5% indicates that the program implementation falls into the moderately good category; however, a gap remains between the quality of planning (100%) and the execution of activities (54.25%). The highest score is found in the Instructional Planning indicator (100%), showing that the preparation of instructional plans and the program blueprint has been comprehensively carried out by the Directorate of the Language Centre (DPB). This is consistent with the statement of the informant, which stated:

"From the planning side, all schedules, learning flows, and materials have been prepared from the beginning…" (DPB)

However, the score of 54.25% on the Implementation of Learning Activities indicator shows that field implementation has not yet aligned with ideal planning. Several factors contribute to this issue. These are reflected in the informants' testimonies. For example, the Head of the Arabic Language Education Study Program (PBA) stated:

"… its implementation often differs because students' abilities vary …" (HD)

This finding is reinforced by LPM, which emphasised that implementation constraints constitute an indicator of the need for continuous evaluation:

"… it does not run optimally because … there is a need to divide focus with campus activities …" (LPM)

The Vice Rector, I highlighted further issues, especially faculty schedules and students' time management:

"At times, the implementation of this language program conflicts with the faculty schedule …" (VR)

The Program Evaluation indicator obtained a score of 69.3%. This indicates that program evaluation has occurred, yet it remains limited to the use of CAT, which primarily assesses receptive abilities. The Head of DPB affirmed this when stating:

"CAT can run well but has not yet been able to capture the ability …" (DPB)

This finding aligns with field observations. Non-formal activities such as language matriculation classes and language clubs or courses have not been consistently effective in improving students' productive competencies. Moreover, dense academic schedules and student activity workloads have led to several language-related activities being carried out less optimally. These overall findings indicate that the success of program implementation is strongly influenced by three main factors: discipline in carrying out learning activities, the motivation and readiness of students, and the alignment between the academic rhythm and the language-related activities. The Vice Rector, I emphasised this relationship clearly:

*"There is a gap between planning and realisation… it is not always aligned with … Arabic language.**" (VR)***

Thus, the implementation results show that although program planning is highly comprehensive and supported by strong academic guidelines, the effectiveness of its execution still needs improvement. This can be achieved through schedule harmonisation, enhancement of student motivation, consistency of non-formal activities, and refinement of evaluation instruments so they can capture productive competencies more comprehensively. To understand the dynamics of the Arabic Language Development Program implementation, this study did not rely solely on quantitative scores. It also conducted qualitative analysis using a code co-occurrence approach in Atlas.ti, as presented in the figure below.
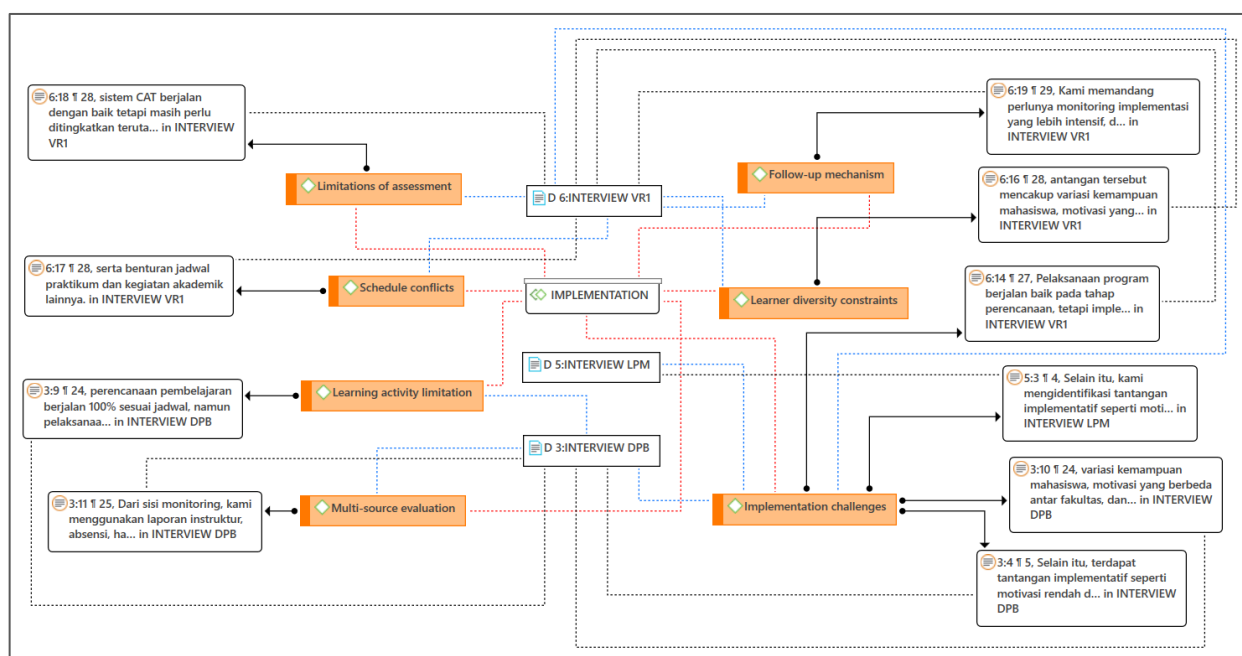
Figure 4. Network Mapping of Program Implementation Code Co-Occurrence Analysis

The network map illustrates that the challenges of implementation are not singular in nature, but rather the result of interactions among structural factors (scheduling, facilities, and monitoring), pedagogical factors (differences in students' abilities and the dynamics of learning activities), and managerial factors (follow-up mechanisms and cross-resource evaluation). Thus, this networking map serves as a visual representation of data triangulation, explaining the gap between the quality of planning (100%) and the implementation outcomes (54.25%). The network reveals several clusters of key challenges, including the limitations of assessment instruments, variations in student ability and motivation, obstacles to implementing learning activities, scheduling conflicts, and the need for more systematic monitoring. Informants consistently described these conditions, for example, when the Vice Rector I highlighted scheduling conflicts and the need for strengthened monitoring, or when the DPB emphasized that although the instructional plan had been well prepared, its realization was often hindered by classroom dynamics and students' readiness. LPM also added that non-formal language activities often did not run optimally due to the intensity of students' academic agendas.

The overall relationships among the codes suggest that the challenges of implementation primarily stem from operational conditions and learner characteristics, rather than weaknesses in program design. Thus, this networking map provides a more comprehensive understanding of the root causes of implementation issues and serves as a strategic foundation for designing more targeted improvement interventions in the future.

Summative Evaluation

The summative evaluation results of the Arabic Language Development Program at the University of Darussalam Gontor show high effectiveness. This is presented in the following table:

Table 9. Summative Evaluation Score

Indicator	Result
Human Resource Development	100%
Program Outputs	90%
Impact Generated by the Language Program	100%
Average	96.6%

Overall, the average achievement of 96.6% indicates that the program is not only successful at the implementation level but also produces strong and sustainable institutional impacts. The Human Resource Development component obtained a perfect score because the program is considered capable of improving the academic quality and language competence of both students and lecturers. This view is reflected in the statement of the informant, who emphasized the program's contribution to enhancing human resource quality:

"The greatest impact … is the improvement in human resource quality… of Arabic language proficiency." (VR)

This emphasis on human resource improvement is echoed by the Head of the Arabic Language Education Study Program, who has directly observed student progress:

"Students … show very noticeable improvement in their language abilities." (HD)

The Program Outputs indicator scored 90%, highlighting the program's effectiveness in producing academic results. The Head of DPB noted that most students reach sufficient language competence for institutional goals, including handling Arabic coursework and writing theses in Arabic:

"The student outputs are also improving; most of them reach the expected level of competence." (DPB)

In addition to this output, the DPB further emphasized that the program's substance contributes significantly to enhancing academic readiness across faculties:

"The program shows that … contributes greatly to the academic preparedness of students." (DPB)

The Impact Generated by the Program indicator received a perfect score, illustrating clear long-term impacts in strengthening the university's academic identity. This aligns with the interview results of the informant, who stated that the program plays an important role in reinforcing UNIDA's scholarly character as a *pesantren*-based institution:

"This program strengthens UNIDA's academic character … Arabic becomes an academic identity …" (VR)

Moreover, the informant also emphasized that enhanced language competence opens broader academic opportunities for students in advanced studies, research, and scientific publication:

"The development of language competence encourages students to be more prepared to read primary literature and to write scholarly works in Arabic." (DPB)

Interpretation of this summative score becomes even stronger when connected with the network mapping visualization, which reveals five major clusters shaping the program's success: (1) Learning Outcome Achievement, (2) High Student Output, (3) Sustainable Program Impact, (4) Skill Enhancement, and (5) Strengthening Academic Identity.

The network visualization shows that program success extends beyond grades, as it also enhances skills, shifts academic identity, and yields tangible contributions to graduate studies. For instance, one informant said student competencies gained from the program greatly increase academic readiness:

"The development of language competence makes students more prepared … thesis writing." (DPB)

Thus, the summative evaluation reveals that the Arabic Language Development Program at UNIDA Gontor yields not only short-term language gains but also broader institutional impacts, such as strengthening academic identity, enhancing human resources, and preserving the Arabic-based scholarly tradition. These findings confirm the program's success in aligning with the university's vision as a *pesantren*-based center for knowledge development.
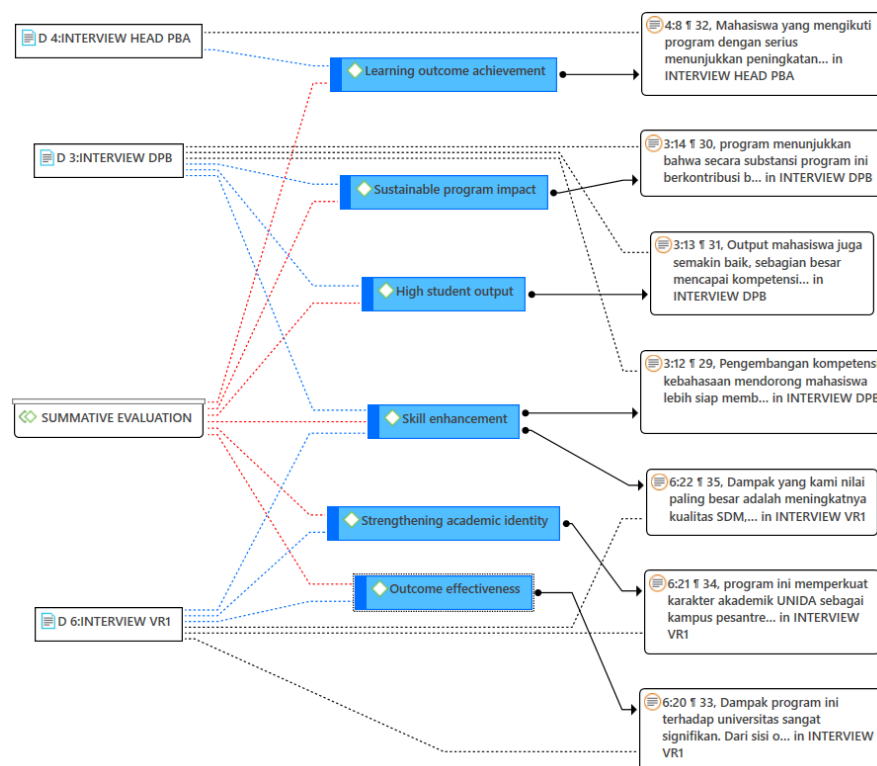
Figure 5. Network Mapping of Summative Evaluation Codes Co-Occurrence Analysis
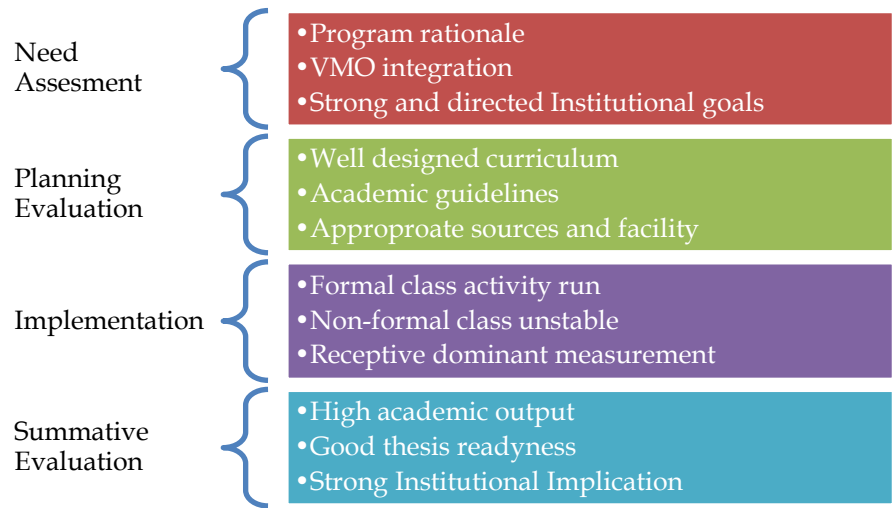
DISCUSSION

Integration of Findings with the CSE-UCLA Framework

The research findings demonstrate that the CSE-UCLA evaluation framework effectively provides a clear and comprehensive analytical structure for assessing the effectiveness of the Arabic Language Development Program at UNIDA Gontor. At the Context and Input stages, represented in this study through Need Assessment and Planning Evaluation, the program demonstrates strong alignment between objectives, institutional vision, and policy design (Aris, 2023). The high scores of these two aspects confirm that the program rationale has been formulated accurately, the curriculum has been systematically structured, and the academic guidelines have met competency-based evaluation standards. This pattern aligns with program evaluation theory, which states that the foundation of successful implementation is largely determined by the clarity of initial needs and the precision of planning (Booker, 2019). Within the context of Islamic higher education, such alignment reinforces the program's position as an institutional strategy for developing linguistic competence that supports mastery of Islamic sciences (Alwi et al., 2024). Therefore, the initial stages of the CSE-UCLA model affirm that the program possesses a strong foundation both structurally and epistemologically (Fatchurahman et al., 2022).

However, when moving to the process and product stages, the CSE-UCLA model helps reveal more complex dynamics. The moderate implementation score indicates an imbalance between strong planning and actual field realization, a phenomenon commonly observed in language education evaluations and noted in various previous studies (Fathurrozi & Pamungkas, 2021). The CSE-UCLA model succeeds in capturing that the gap originates from contextual factors such as the rhythm of academic–*pesantren* activities, heterogeneity of students' abilities, and limitations in productive assessment instruments. Nevertheless, the high product score indicates that the program continues to have a significant positive impact, both in enhancing student competence and in strengthening institutional academic identity. This pattern reflects the distinctive character of Islamic education evaluation, in which the strength of institutional vision and values is able to withstand the negative effects of implementation barriers (Mulyana & Sa, 2024). Thus, the integration of findings through

the CSE-UCLA framework not only assesses program effectiveness but also presents the dynamic relationship between context, which may serve as a strategic foundation for continuous improvement and decision-making (Samuel & Farrer, 2025).

Need Assesment
- Program rationale
- VMO integration
- Strong and directed Institutional goals

Planning Evaluation
- Well designed curriculum
- Academic guidelines
- Approproate sources and facility

Implementation
- Formal class activity run
- Non-formal class unstable
- Receptive dominant measurement

Summative Evaluation
- High academic output
- Good thesis readyness
- Strong Institutional Implication

Figure 6. Evaluation Landscape of the Arabic Language Development Program

Arabic Language and the Academic Institutional Identity

The research findings indicate that the success of the Arabic Language Development Program at UNIDA Gontor is closely tied to its institutional identity as an Islamic higher education institution rooted in the *pesantren* tradition (Mujahid et al., 2022). Within the framework of Islamic higher education, the Arabic language holds an epistemological role that surpasses its communicative function; it serves as the primary medium for understanding normative religious texts, accessing the intellectual heritage of *turats*, and constructing an Islamic epistemic framework that underlies the Islamization of knowledge. The strong compatibility between the university's vision and the program objectives, as reflected in the perfect Need Assessment score, indicates that Arabic is not merely a course, but an integral part of the institution's scholarly values and orientation. In line with the thoughts of Al-Attas and Zarkasyi, Arabic language competence is viewed as a key requirement for building an academic persona rooted in *adab*, deep literacy, and intellectual authority within the Islamic tradition (Hayden, 2024). This integration demonstrates that the Arabic language program at UNIDA serves as a bridge between the *pesantren* tradition and modern academic demands, ensuring that the educational process remains grounded in Islamic values while remaining relevant to contemporary scholarly requirements (Warnis et al., 2019).

Furthermore, this relevance is evident in the academic achievements of students and in the curriculum orientation that positions the Arabic language as the primary instrument for developing scholarly capacity (Umam et al., 2024). The program's approach, comprising matriculation, structured learning, and competency-based assessment, supports the creation of higher-order learning outcomes aligned with modern Islamic education literature, namely the ability to reason, read critically, write academically, and construct arguments in Arabic (Amirul Mukminin et al., 2025). The high summative results confirm that the program successfully forms strategic competencies needed by students to engage in Islamic academic discourse, such as writing an Arabic-language thesis and understanding primary texts. This is consistent with international research, which highlights that Arabic language programs in Islamic higher education become effective only when Arabic functions as a medium of knowledge rather than merely a technical subject (Alsaied et al., 2025). Thus, the Arabic language program at UNIDA can be viewed as an integrative model that connects linguistic competence with the epistemological identity of Islamic higher education, strengthening the institution's character (Samir et al., 2025) as a center for knowledge development rooted in *pesantren* tradition while simultaneously positioning it as an academic actor within the landscape of modern Islamic education.

The Gap Between Planning and Implementation of the Arabic Language Program

The gap between high-quality planning and less optimal implementation is a phenomenon commonly found in program evaluation, including in the context of Arabic language learning in Islamic higher education institutions (Al-Shaboul, 2025). Based on the CSE-UCLA framework, the input (planning) component is often easier to control and standardize because it consists of documents, designs, and policies. In contrast, the process (implementation) component is strongly influenced by human factors, institutional context, and operational field dynamics. The findings of this study clearly illustrate this pattern: program planning demonstrates strong alignment in curriculum, academic guidelines, and standardized instruments, yet classroom implementation and non-formal language activities encounter challenges related to student heterogeneity, the rhythm of academic–*pesantren* activities, and limitations of productive assessment (Mangadang et al., 2025).

This gap matches language program research by Hidayah (Hidayati & Muradi, 2025), who find that good curriculum design alone does not ensure success without a supportive learning environment, prepared students, and regular monitoring. We should view this gap as a natural consequence of ideal programs meeting complex academic realities, rather than a planning failure.

From the perspective of implementation gap theory in educational evaluation, this discrepancy can also be explained through the concept of ecological validity, which refers to the extent to which program design aligns with real conditions in the learning environment (Mangadang et al., 2025). The findings of this study indicate that the Arabic language curriculum design aligns with the vision of Islamic higher education; however, its implementation is influenced by the rhythm of *pesantren* activities, variations in student readiness across faculties, and the limitations of computer-based assessment instruments that predominantly measure receptive skills.

The consistency of these findings with other research, such as Salmas's study on Arabic language learning in Islamic universities, shows that implementation barriers are often structural and cultural rather than merely technical (Salmas et al., 2024b). They emphasize that Arabic language learning in Islamic higher education requires a stable linguistic ecology, synchronized scheduling, a conducive communicative space, well-managed student motivation, and evaluation instruments capable of capturing the complexity of language performance (Umam et al., 2024). This planning–implementation discrepancy does not diminish the overall effectiveness of the program; rather, it highlights the need to strengthen the *process* component by enhancing classroom management, adapting pedagogy to student diversity, integrating performance-based assessment, and harmonizing academic rhythms with *pesantren* routines. Therefore, this discussion demonstrates that improving implementation is not only an operational adjustment but also a strategy for reinforcing the identity of Islamic higher education, ensuring it remains adaptive and relevant within the modern academic landscape (Ritonga et al., 2021).

Enhancement of Student Competence and Academic Output

The research findings indicate that the Arabic Language Development Program has a significant contribution to the improvement of students' linguistic competence, aligning with the principle of linguistic cultivation in Islamic higher education (Alwi et al., 2024). From the perspective of language education studies, the improvement of student competence is not only viewed through receptive and productive abilities, but also through their capacity to access, understand, and process primary literature that forms the epistemological foundation of Islamic studies. This is consistent with the model, which emphasizes that successful language learning encompasses grammatical, discoursal, and sociocultural competence (Saif & Umar, 2021).

At a *pesantren*-based university like UNIDA Gontor, Arabic literacy means more than communication skills. It is necessary to engage with the Islamic intellectual tradition, read classical texts, and participate in source-based academic arguments (Al-Ghazali & Al-Zyoudi, 2025). The results of this study demonstrate that

the program supports student competence and fosters an Islamic scholarly identity in higher education (Harmeni, 2022).

In addition to competence improvement, the summative results also show that students' academic outputs have developed positively and are consistent with the institution's objective of forming graduates who are scientifically competent and rooted in *pesantren* character (Maimunah et al., 2021). Within higher-order learning outcomes theory, achievements such as the ability to write academic works in Arabic, construct scholarly arguments, and present ideas in academic forums indicate that students have reached a level of applied linguistic proficiency that meets the standards of higher education (Almelhes, 2024).

These results demonstrate that well-designed language programs can have a lasting impact, particularly in fostering students' participation in Islamic academic discussions (Rosowulan et al., 2025). The outputs show more than technical learning success. They also align with Islamic higher education's goal of developing scholars who combine language skills with a deep understanding of Islamic tradition. UNIDA's program stands as a best practice model for institutions seeking both language proficiency and holistic intellectual identity.

Research Contribution on Language Evaluation Systems

This study contributes to the development of Arabic language evaluation in Islamic higher education by linking academic goals with modern evaluation standards (Muflihah et al., 2025). Results show that the current computer-based system measures receptive skills well, but it does not yet capture the productive skills needed in higher education, such as speaking, argumentation, and academic text creation (Muflihah et al., 2025).

From the perspective of applied linguistics and language assessment theory, this condition confirms that multiple-choice-based testing alone is insufficient to represent communicative competence comprehensively (Wudthayagorn, 2025). International literature, such as the ACTFL Proficiency Guidelines, emphasizes the importance of performance-based assessment to obtain an authentic picture of language ability (Tayyara, 2022). Therefore, the findings of this study reinforce the urgency of implementing a more holistic and multimodal evaluation system, so that learning outcomes are not only technically measurable but also reflect the academic demands of Islamic scholarship, which require mastery of primary literacy and the ability to produce scientific discourse.

The research findings also described that the Arabic language evaluation system needs to be developed as part of a continuous academic quality management mechanism. The high summative results indicate that the final objectives of the program have been achieved, yet the implementation gap suggests the need for integrating formative evaluation throughout the learning process (Sista et al., 2022). Within the CSE-UCLA program evaluation approach, the quality of the *process* is strongly influenced by monitoring mechanisms, the suitability of instruments, and the sustainability of assessment. Therefore, the Arabic evaluation system should not merely function as a measurement tool, but must serve as a feedback system that guides academic decision-making, curriculum design, and remedial strategies (Fatchurahman et al., 2022).

The contribution of this study lies in affirming that Arabic language evaluation must run beyond the usual measurement of proficiency toward a quality control system that links learning outcomes, student readiness, and institutional strategic planning (Muflihah et al., 2025). This aligns with the needs of modern Islamic higher education to possess assessment instruments that not only reflect linguistic ability but also strengthen the academic culture and epistemological identity that form the core character of Islamic higher education (Mahmudi et al., 2025).

Implications for Policy in *Pesantren*-Based Higher Education

The findings of this research carry strategic implications for the development of policy in *pesantren*-based higher education, particularly in managing Arabic language programs as both epistemological instruments and institutional identity. Within the framework of Islamic higher education governance theory, the Arabic language serves not only as a linguistic skill but also as a prerequisite for forming an integrated knowledge

system, one that combines Islamic values with modern academic competence (Solahudeen et al., 2025). The evaluation results demonstrate that the program's success at the contextual, planning, and output stages reinforces the position of Arabic as an academic policy pillar inseparable from the institutional mission. This aligns with the concept of knowledge Islamization (Noraina et al., 2024), which posits that Arabic language mastery serves as the foundation for accessing Islamic epistemology authentically. Therefore, policy in *pesantren*-based universities must continue to prioritize the integration of Arabic across the entire academic spectrum, from compulsory cross-faculty curricula and capacity building for lecturers to the development of an academic environment that encourages the active use of Arabic in scientific and social activities on campus. This program demonstrates that strong language policy can function simultaneously as a guardian of *pesantren* scholarly tradition and as a driver of modern academic quality (Kamaruddin et al., 2025).

Furthermore, the findings related to implementation gaps provide important insights for formulating academic management policies in *pesantren*-based higher education institutions (Muhammad Iqbal et al., 2025). These gaps support the theory of educational ecology, which emphasizes that successful curriculum implementation relies not solely on documents and regulations, but also on the alignment of institutional rhythms, student motivation, and the readiness of evaluation tools (Mursal, 2024). In the context of UNIDA Gontor, this indicates that Arabic language policy must be developed not only from the top down through regulations, but also from the bottom up through the strengthening of academic culture among students and lecturers.

Research Limitations

This research has several limitations that should be noted as part of methodological reflection. First, the quantitative evaluation instrument used still relies on a structured evaluation checklist, meaning that the measurement of program effectiveness highlights standardized aspects that can be converted into numerical scores; however, it does not fully capture the qualitative and contextual dynamics of the learning process (Tatheer Aziz, 2025). Second, the language assessment that forms the core of program evaluation is still dominated by receptive-skill measurement through the Computerized Adaptive Test (CAT), resulting in productive abilities such as academic writing and scholarly speaking not being measured comprehensively, despite both being essential components of Arabic language learning outcomes in Islamic higher education (Sofa et al., 2023).

Another limitation is that the scope of informants remains limited to several key institutional actors, and therefore does not fully reflect the diversity of student experiences across faculties and levels of readiness (Yugo, 2025). Fourth, the characteristics of *pesantren*-based educational environments, with their dense and dynamic activity rhythm, may influence the consistency of program implementation, making the findings contextual and potentially different if applied in Islamic institutions without similar cultural settings. Therefore, although this research provides a strong and comprehensive evaluative overview, the results must still be interpreted with consideration of methodological limitations and the unique institutional context (Chotimah et al., 2022).

CONCLUSION

This research affirms that the Arabic Language Development Program at the University of Darussalam Gontor has a strong conceptual foundation aligned with the vision of *pesantren*-based Islamic higher education, as reflected in the excellent results of the Need Assessment and Planning Evaluation stages. Through the CSE-UCLA framework, this study reveals that although the curriculum design, academic instruments, and strategic direction of the program have been systematically arranged and oriented toward building linguistic competence as the epistemological basis of Islamic knowledge, program implementation still faces several challenges, particularly regarding student heterogeneity, the dynamics of *pesantren* activities that affect implementation consistency, and limitations in productive assessment which have not yet optimally measured academic speaking and writing skills. However, the summative achievements show that the

program continues to generate significant impact on the improvement of students' academic competence and the strengthening of the institution's scholarly identity. Overall, this study contributes to the development of evaluation frameworks for Arabic language education in Islamic higher education and emphasizes the importance of strengthening performance-based assessment, harmonizing the academic–*pesantren* ecology, and utilizing continuous evaluation as a foundation for strategic policy-making to improve program quality in the long term.

REFERENCES

Abu Aser, A. A., Tabieh, A. A. S., & Abu Awwad, F. M. (2021). Thematic and methodological approaches in curriculum and teaching methods researched in master's theses at middle East University in Jordan: A content analysis. In *Asian ESP Journal* (Vol. 17, Issue 3.2, pp. 145–173). Asian EFL Journal Press.

Alghazali, F., & Alzyoudi, M. (2025). Exploring Linguistic Paradigms in Language Applications to Augment Arabic Language Acquisition. *Novitas-ROYAL*, *19*(1), 249. https://doi.org/10.5281/ZENODO.15228458

Almelhes, S. A. (2024). Gamification for teaching the Arabic language to non-native speakers: A systematic literature review. *Frontiers in Education*, *9*(2), 02. https://doi.org/10.3389/feduc.2024.1371955

Alsaied, M. A., Aldarmaki, N. A. M. B., & Alghefari, Z. T. (2025). From Causes to Remedial Attempts: The Phenomenon of Aversion to Learning and Using the Arabic Language. *Journal of Posthumanism*, *5*(3). https://doi.org/10.63332/joph.v5i3.936

Al-Shaboul, I. A. (2025). Bridging the gap: The role of artificial intelligence in enhancing Arabic language learning, translation, and speech recognition. *Research Journal in Advanced Humanities*, *6*(2), 1–13.

Alwi, S., Syukran, S., & Sari, M. (2024). Internalization of Character Education Values in Arabic Language Learning at Islamic Boarding School in Indonesia. *Al-Ta'rib : Jurnal Ilmiah Program Studi Pendidikan Bahasa Arab IAIN Palangka Raya*, *12*(1), 89. https://doi.org/10.23971/altarib.v12i1.7882

Amirul Mukminin, Hanun, A., Zainuddin, Mushtofa, L., & Wassalwa, A. (2025). Integration of Bayani, Burhani and Irfani Epistemologies in Arabic Language Learning in Islamic Boarding School-Based Colleges. *Asalibuna*, *9*(01), 91–107. https://doi.org/10.30762/asalibuna.v9i01.5292

Aris, T. A. P. (2023). Evaluasi Program Planing PLP II (Pengenalan Lapangan Persekolahan): Menggunakan Model CSE-UCLA di PAUD. *Murhum : Jurnal Pendidikan Anak Usia Dini*, *4*(1), 11–20. https://doi.org/10.37985/murhum.v4i1.158

Asril, Z., & Hanafi, A. H. (2021). Reconstruction of Islamic Institution Education the 21st Hopes, Challenges, Paradigms, and Roles of Arabic. *Al-Ta Lim Journal*, *28*(1), 83–91. https://doi.org/10.15548/jt.v28i1.689

Booker, Z. Z. (2019). A Revised Logic Model for Educational Program Evaluation. *Journal of Modern Applied Statistical Methods*, *18*(2), 2–14. https://doi.org/10.22237/jmasm/1604190300

Chotimah, C., Khusna, N. I., Amin, S., & Palupi, C. (2022). Implementation of Islamic Boarding School Socio-Cultural Management-Based Curriculum in Developing of Islamic Educational Institutions Quality. *J-MPI (Jurnal Manajemen Pendidikan Islam)*, *7*(2), 63–75. https://doi.org/10.18860/jmpi.v7i2.15896

Creswell, J. W. (2011). *Research design: Quantitative, Qualitative research and mixed methods aproach*.

Fatchurahman, M., Setiawan, M. A., & Karyanti, K. (2022). Evaluation of the CSE-UCLA model on the performance of school counselor in Indonesia. *Perspectives of Science and Education*, *56*(2), 561–572. https://doi.org/10.32744/pse.2022.2.33

Fathurrozi, A., & Pamungkas, R. W. P. (2021). Pemberdayaan model evaluasi CSE-UCLA pada program sertifikasi kompetensi BNSP bidang network administrator madya dan pemrograman basis data di LSP-P1 Universitas Bhayangkara Jakarta raya. *Journal of Informatic and Information Security*, *2*(1), 67–76. https://doi.org/10.31599/jiforty.v2i1.629

Fitriani, A. Z. (2021). Evaluasi program e-learning pada prodi pendidikan bahasa arab Universitas Muhammadiyah Yogyakarta dengan model CIPP. *Maharaat: Jurnal Pendidikan Bahasa Arab*, *3*(2), 109–127. https://doi.org/10.18196/mht.v3i2.10481

Harmeni, H. A. (2022). Construct validity analysis: Assessment for learning primary mathematics questionnaire. *Journal for Educators, Teachers and Trainers*, *13*(2), 54–72. https://doi.org/10.47750/jett.2022.13.02.006

Hayden, W. (2024). The Epistemology of Ta'dib in Islamic Civilizational Discourse: Reviving and Reconstructing Contemporary Muslim Scholars' Views. *Journal of Al-Tamaddun*, *19*(1), 181–197. Scopus. https://doi.org/10.22452/JAT.vol19no1.14

Hidayati, H., & Muradi, A. (2025). Curriculum Integration in Semi-Modern Pesantren: A Holistic Approach in Arabic Language Education. *Jurnal Alfazuna : Jurnal Pembelajaran Bahasa Arab Dan Kebahasaaraban*, *9*(2), 154–176. https://doi.org/10.15642/alfazuna.v9i2.4505

Juliana, S. A., Liza, T., Fatimahtuzzahra, & Imel, M. A. H. (2023). Tantangan sosial di era digital pada interaksi manusia. *SIGNIFICANT: Journal Of Research and Multidisciplinary*, *2*(2), 245–261.

Kamaruddin, K., Abu Bakar, M., & Aziz, N. I. (2025). Challenges and strategies for mastering the Arabic language in the Islamic studies diploma program: A study at UNISHAMS. *International Journal of Education, Psychology and Counseling*, *10*(57), 801–823. https://doi.org/10.35631/IJEPC.1057052

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%2520/ijazarabi.%2520V8i2.29595

Maimunah, I., Huda, M., Haque, A., & Zubaidah, Z. (2021). Islamic Boarding School at University: A Strong Pathway for Integrating Religion and Science. *Jurnal Pendidikan Agama Islam*, *8*(1), 37–50. https://doi.org/10.18860/jpai.v8i1.15361

Mangadang, A. B., Ubayubay, R. M., & Comon, J. D. (2025). Implementation and Challenges of Arabic Language and Islamic Values Education Program: Basis for School Development Plan. *American Journal of Arts and Human Science*, *4*(1), 129–152. https://doi.org/10.54536/ajahs.v4i1.4139

Mardapi, D. (2008). *Teknik Penyusunan Instrumen Tes dan Nontes* (1st edn). Mitra Cendikia Press.

Mardapi, D. (2012). *Pengukuran Penilaian dan Evaluasi Pendidikan* (1st edn). Nuha Medika.

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

Muhammad Iqbal, Mohammad Ahsanuddin, & Muhammad Alfan. (2025). Enhancing Arabic Proficiency through Content and Language Integrated Learning: Implementation and Challenges in Pesantren. *International Journal of Arabic Language Teaching*, *7*(02), 207–222. https://doi.org/10.32332/ijalt.v7i02.10769

Mujahid, I. (2021). Islamic orthodoxy-based character education: Creating moderate Muslim in a modern pesantren in Indonesia. *Indonesian Journal of Islam and Muslim Societies*, *11*(2), 185–212. https://doi.org/10.18326/ijims.v11i2.185-212

Mujahid, I., Nasiruddin, M., & Hudayana, K. (2022). Evaluasi program "KEMBARA" sebagai upaya dasar peningkatan program pembelajaran bahasa arab di Universitas Darussalam Gontor. *VOX EDUKASI: Jurnal Ilmiah Ilmu Pendidikan*, *13*(1), 89–101. https://doi.org/10.31932/ve.v13i1.1420

Mulyana, R., & Sa, D. (2024). Exploratory data analysis to reveal learning loss condition in Islamic religious education. *International Journal of Evaluation and Research in Education (IJERE)*, *13*(1), 43–56. https://doi.org/10.11591/ijere.v13i1.26344

Mursal. (2024). Integrasi pendidikan tinggi dan pesantren dalam penguatan konsep ekonomi Islam. *Jurnal Al-Mizan: Jurnal Hukum Islam Dan Ekonomi Syariah*, *11*(1), 181–192.

Noraina, Robi'ah, & Darwis, M. (2024). Konsep Pendidikan Anak Dalam Keluarga (Tela'ah Buku Biografi K.H. Imam Zarkasyi Dari Gontor Merintis Pesantren Modern). *JIPKIS: Jurnal Ilmiah Pendidikan Dan Keislaman*, *4*(1).

Quainoo, E. A., Aggrey, R., Aggrey, D., Adams, F., Opoku, E., & Wanzam Abubakari, Z. (2022). The impact of globalization on education: A blessing or a curse. *Education Journal*, *11*(2), 70–70. https://doi.org/10.11648/j.edu.20221102.13

Ritonga, M., Widodo, H., Munirah, M., & Nurdianto, T. (2021). Arabic language learning reconstruction as a response to strengthen Al-Islam studies at higher education. *International Journal of Evaluation and Research in Education (IJERE)*, *10*(1), 355. https://doi.org/10.11591/ijere.v10i1.20747

Rohmah, Z., Sholihah, D. N., Basthomi, Y., & Milal, A. D. (2024). School-scape as a linguistic battlefield: An LL study at Darussalam education complex of Gontor, Indonesia. *Cogent Arts and Humanities*, *11*(1), 1–18. https://doi.org/10.1080/23311983.2024.2348301

Rosowulan, T., Hasyim, A. F., Sholikhun, M., Purwanto, P., Djamil, A., In'amuzzahidin, M., & Wijaya, R. (2025). Pesantren's Knowledge Identity Crisis in the Digital Era. *Jurnal Ilmiah Peuradeun*, *13*(1), 53. https://doi.org/10.26811/peuradeun.v13i1.1287

Saif, A., & Umar, I. (2021). Research on Online Problem-Based Learning Among Undergraduate Students: A Systematic Review. In *Lecture Notes on Data Engineering and Communications Technologies* (Vol. 72, p. 1191). Scopus. https://doi.org/10.1007/978-3-030-70713-2_105

Salmas, A. N. H., Bahruddin, U., & Syuhadak. (2024a). Exploring language proficiency: Dynamics of acquiring arabic speaking skills from Stephen Krashen's perspective at islamic universities. *AL-ISHLAH: Jurnal Pendidikan*, *16*(3). https://doi.org/10.35445/alishlah.v16i3.5100

Salmas, A. N. H., Bahruddin, U., & Syuhadak, S. (2024b). Exploring Language Proficiency: Dynamics of Acquiring Arabic Speaking Skills from Stephen Krashen's Perspective At Islamic Universities. *AL-ISHLAH: Jurnal Pendidikan*, *16*(3). https://doi.org/10.35445/alishlah.v16i3.5100

Samir, H. E.-D., Haidara, N. B., Alghazali, F., & Badran, M. A. (2025). The Impact of Language Immersion on Arabic Language Acquisition for Speakers of Other Languages: "Enhancing Arabic Language Skills for Speakers of Other Langugaes" Program as a Model. *Journal of Posthumanism*, *5*(3). https://doi.org/10.63332/joph.v5i3.919

Samuel, S., & Farrer, H. (2025). Integrating the PDCA cycle for continuous improvement and academic quality enhancement in higher education. *Journal of Comparative & International Higher Education*, *17*(2), 115–124.

Sista, T. R., Muttaqin, M. R., & Wijaksono, A. (2022). The Implementation of Human Resource Management in Development of Undergraduate Teacher at Modern Islamic Institution Darussalam Gontor Campus 2. *At-Ta'dib*, *17*(1), 113–141.

Sofa, F., Wahdah, Y. A., Baroroh, R. U., Amalia, N. M., & Afifah, S. N. (2023). Analysis of Assessment of Arabic Receptive and Productive Skills in Arabic Language Book. *An Nabighoh*, *25*(1), 1. https://doi.org/10.32332/an-nabighoh.v25i1.5764

Solahudeen, O.-A. I., Nurudeen, B., Shehu, A., & Oluwatosin, Y. I. (2025). Attitude of Islamic Studies Students towards Learning Arabic Language: Secondary Schools in Asa Local Government. *Jurnal Eduscience*, *12*(2), 356–365.

Sugiyono. (2020). *Penelitian Kombinasi (Mix Methode)* (3rd edn). Alfabeta.

Surur, M. (2022). Tantangan dan peluang Bahasa arab di Indonesia. *Risda: Jurnal Pemikiran Dan Pendidikan Islam*, *6*(2), 176–186. https://doi.org/10.59355/risda.v6i2.90

Tatheer Aziz. (2025). Language Learning Outcomes: A Critical Analysis of Assessment Methods and Their Effects on Student Achievement. *Journal of Information Systems Engineering and Management*, *10*(33s), 580–595. https://doi.org/10.52783/jisem.v10i33s.5567

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

Umam, M. K., Delimanugari, D., Dwi Wahyuni, A., & Hilatun Nisa, I. (2024). Arabic language learning at institutions for non-islamic education study programs. *Jurnal Asy-Syukriyyah*, *25*(1), 1–17. https://doi.org/10.36769/asy.v25i1.424

Warnis, W., Triana, H. W., Kustati, M., Remiswal, R., & Nelmawarni, N. (2019). Arabic Language as the Icon of Islamic Higher Education: A Study of the Implementation of Arabic Intensive Program. *TARBIYA: Journal of Education in Muslim Society*, *6*(1), 103–116. https://doi.org/10.15408/tjems.v6i1.10910

Wudthayagorn, J. (2025). Revisiting English-in-Education Policies in Thailand: Ambitious Goals, Contradictory Outcomes. *LEARN Journal: Language Education and Acquisition Research Network*, *18*(2), 1–13. Scopus. https://doi.org/10.70730/WCSZ5574

Yugo, T. (2025). Improving the Quality of Islamic Education through Pesantren-Based Management in Indonesia. *Journal of Educational Research and Practice*, *3*(2), 238–254. https://doi.org/10.70376/jerp.v3i2.357

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