



EVALUATION OF ARABIC LANGUAGE PROGRAM MANAGEMENT AT MA'HAD DARUL HIKAM MTSN IN KOTA BATU BASED ON THE CIPP MODEL



Aliza Nur Muhammad¹, Ari Sandi Pradana², Muhammad Hasbi Ashsidqi³

*Correspondence:

Email: alizanurmuhammad@gmail.com

Author Affiliation:

^{1,2,3} Universitas Islam Negeri
Maulana Malik Ibrahim
Malang, Indonesia

Article History:

Submission: December 26,
2025
Revision: April 27, 2026
Accepted: May 28, 2026
Published: June 30, 2026

Keywords:

Program Evaluation, CIPP
Model, Arabic Language
Management, Ma'had
Madrasah, Bi'ah Lughowiyyah

Kata Kunci:

Evaluasi Program, Model
CIPP, Manajemen Bahasa
Arab, Ma'had Madrasah, Bi'ah
Lughowiyyah

Abstract

The challenge of integrating the national curriculum with Arabic language proficiency targets in hybrid madrasah – pesantren educational institutions requires a comprehensive evaluation of program management. This study aims to evaluate the management of the Arabic language program at Ma'had Darul Hikam MTsN Kota Batu using the *Context, Input, Process, Product (CIPP)* model. The novelty of this study lies in the evaluation of an Arabic language program at a ma'had integrated with a public madrasah, which has characteristics distinct from both pesantren and regular madrasahs. The study employed a descriptive qualitative approach with subjects including the Ma'had Director, Language Coordinator, ustadz/instructors, musyrifah, and female students selected through purposive sampling. Data were collected through participant observation, in-depth interviews, and documentation. Data analysis utilized the interactive model developed by Miles, Huberman, and Saldana. The research findings indicate that: (1) the *context* aspect shows that the program's objectives are relevant in integrating Salaf and modern pesantren traditions; (2) the input aspect indicates that the textbook *'Durusul Lughah Al – 'Arabiyah'* aligns with the objectives, but the use of *Smart TVs* has not been optimized; (3) the *process* aspect indicates that the *"fun learning"* method enhances motivation, but the program is not yet sustainable and a *language-rich environment* has not yet been established; and (4) the *product* aspect indicates an improvement in cognitive understanding, but active oral communication skills are not yet optimal. The findings underscore the importance of strengthening language environment management to bridge cognitive competencies and communicative performance.

Abstrak

Tantangan integrasi kurikulum nasional dengan target kemahiran bahasa Arab di lembaga pendidikan hybrid madrasah–pesantren menuntut evaluasi manajemen program secara komprehensif. Penelitian ini bertujuan mengevaluasi manajemen program bahasa Arab di Ma'had Darul Hikam MTsN Kota Batu menggunakan model *Context, Input, Process, Product (CIPP)*. Kebaruan penelitian terletak pada evaluasi program bahasa Arab pada ma'had yang terintegrasi dengan madrasah negeri, yang memiliki karakteristik berbeda dari pesantren maupun madrasah reguler. Penelitian menggunakan pendekatan kualitatif deskriptif dengan subjek Mudir Ma'had, Koordinator Bahasa, ustadz/pengajar, musyrifah, dan santriwati yang dipilih secara purposif. Data dikumpulkan melalui observasi partisipan, wawancara mendalam, dan dokumentasi. Analisis data menggunakan model interaktif Miles, Huberman, dan Saldana. Hasil penelitian menunjukkan bahwa: (1) aspek *context* menunjukkan tujuan program relevan dalam mengintegrasikan tradisi pesantren salaf dan modern; (2) aspek input menunjukkan kitab *Durusul Lughah Al-'Arabiyah* sesuai tujuan, tetapi pemanfaatan *Smart TV* belum optimal; (3) aspek *process* menunjukkan metode *fun learning* meningkatkan motivasi, tetapi program belum berkelanjutan dan belum terbentuk bi'ah lughawiyyah; dan (4) aspek *product* menunjukkan peningkatan pemahaman kognitif, tetapi keterampilan komunikasi lisan aktif belum optimal. Temuan menegaskan pentingnya penguatan manajemen lingkungan bahasa untuk menjembatani kompetensi kognitif dan performa komunikatif.



INTRODUCTION

Arabic language programs within the context of Islamic education in Indonesia hold a strategic position as a means of developing language proficiency while strengthening students' understanding of Islam. However, their implementation in educational institutions that integrate the madrasah and pesantren (ma'had) systems still faces various challenges, particularly in aligning the requirements of the national curriculum with the distinctive characteristics of Arabic language instruction found in pesantren. The enactment of the Minister of Religious Affairs Decree (KMA) No. 450 of 2024 and the implementation of the Merdeka Curriculum increasingly require educational institutions to develop Arabic language programs that not only meet national learning outcomes but are also capable of building students' communicative and religious competencies (Jaafar et al., 2023). In practice, there remains a gap between program planning and its implementation in the field, limited human resources and learning support facilities, as well as a suboptimal program evaluation system that could serve as a basis for decision – making. These conditions result in many Arabic language programs operating routinely without clear measures of success or sustainable improvement mechanisms.

These issues indicate that the effectiveness of Arabic language programs is determined not only by the classroom learning process but also by the overall quality of program management. Several previous studies have examined the evaluation of Arabic language programs using the CIPP model in various educational contexts. Mahmudi et al. (2023) found that the success of Arabic language programs is determined by the interrelationship between the context, inputs, processes, and outcomes; thus, weaknesses in any one component will affect the quality of the final product. Research by Adam et al. (2024) indicates that the implementation of the Arabic language curriculum in madrasahs requires adjustments to align with the latest KMA policies to ensure that learning objectives remain relevant to students' needs. Meanwhile, Jaafar et al. (2023) revealed that Arabic language learning in Islamic boarding schools (pesantren) still faces various challenges, including limited teacher competence, insufficient learning materials, and low student motivation.

The findings of Fauzi and Inayati (2023) also confirm that systematic program evaluation can serve as a foundation for improving the effectiveness of learning management. The commonality among these studies indicates that the CIPP model is a comprehensive evaluation approach for assessing a program's effectiveness; however, the focus of these studies remains largely on evaluating instruction, curriculum implementation, or Arabic language programs in traditional Islamic boarding schools, regular madrasahs, and higher education institutions. On the other hand, research specifically evaluating the management of Arabic language programs in educational institutions that integrate the public madrasah system with ma'had remains very limited. In fact, this educational model has distinct characteristics because it must integrate the Ministry of Religious Affairs' formal education policies with the academic culture of pesantren through 24 – hour student guidance. This situation poses more complex challenges, ranging from aligning program objectives, managing resources, and implementing instruction to achieving Arabic language learning outcomes. Consequently, a research gap remains in the form of a lack of studies that comprehensively evaluate the interrelationships between *context*, *input*, *process*, and *product* in the management of Arabic language programs within the hybrid madrasah – pesantren educational model. This gap is important to examine because the evaluation results not only illustrate the program's effectiveness but can also serve as the basis for formulating recommendations

for the development of Arabic language program management that aligns with the characteristics of integrated educational institutions.

Based on the above conditions, this study offers a novel approach in the form of a comprehensive evaluation of Arabic language program management using the CIPP model at Ma'had Darul Hikam MTsN in Batu City as a representative of educational institutions that integrate the public madrasah and pesantren systems. Unlike previous studies, which have largely focused on evaluating the learning process or curriculum implementation, this study centers its analysis on the integration of policy context, resource readiness, program implementation, and outcomes as a unified Arabic language program management system. This approach is expected to provide a more comprehensive picture of the effectiveness of program management while offering evidence-based recommendations for the development of Arabic language programs in madrasahs that implement the ma'had system.

Therefore, the research question for this study is: How is the management of the Arabic language program at Ma'had Darul Hikam MTsN in Batu City assessed based on the four components of the CIPP model—*Context*, *Input*, *Process*, and *Product*? In line with this research question, this study aims to evaluate the alignment of the program's context with the institution's vision and KMA Regulation No. 450 of 2024; analyze the readiness of program inputs, including human resources, curriculum, and facilities and infrastructure; evaluate the implementation of the Arabic language learning process along with the challenges in its implementation; and analyze the outcomes of the program's implementation in enhancing students' Arabic language proficiency. The results of this study are expected to provide a theoretical contribution to the development of Arabic language program evaluation studies while also serving as practical recommendations for madrasah and ma'had administrators in improving the quality and sustainability of Arabic language programs at integrated Islamic educational institutions.

RESEARCH METHODS

This study employs a qualitative approach using *an evaluative research* design. The qualitative approach was chosen because this study aims to gain an in-depth understanding of the phenomenon of Arabic language program management occurring in its natural context, thereby providing a holistic picture of the program's planning, implementation, and outcomes. The evaluation model used is *CIPP (Context, Input, Process, Product)*, developed by Daniel L. Stufflebeam (1971). This model was chosen because it enables a comprehensive evaluation of the program through four main dimensions: *context* (alignment of program needs, objectives, and policies), *input* (readiness of human resources, curriculum, facilities and infrastructure, and funding), *process* (instructional delivery, program management, monitoring, and the implementation of differentiated instruction and *bi'ah lughawiyah*), and *product* (students' attainment of Arabic language competencies, changes in language behavior, and the program's success rate).

The study was conducted at Ma'had Darul Hikam MTsN in Batu City. The research location was purposely selected because it is one of the public madrasahs that integrates the formal education system with the pesantren (ma'had) training system, giving it characteristics that differ from both regular madrasahs and traditional pesantren. The integration of the madrasah curriculum with a pesantren-based Arabic language program makes this institution a representative location for to evaluate the effectiveness of Arabic language program management within a hybrid education model.

The research data sources consisted of primary and secondary data. Primary data were obtained from 10 key informants selected using *purposive sampling*—a technique for selecting informants based on the consideration that they possess knowledge, experience, authority, and direct involvement in the management of Arabic language programs. These informants consisted of one madrasah principal who served as the person in charge of institutional policy, one ma'had director as the primary manager of the ma'had program, one Arabic language program coordinator responsible for program planning and implementation, two Arabic language teachers directly involved in the learning process, two dormitory supervisors who supported the implementation of language immersion within the ma'had environment, and three students selected to represent high, intermediate, and low proficiency levels in Arabic language learning. The selection of these student characteristics was intended to ensure that the data collected could reflect learning experiences across various proficiency levels, thereby yielding more comprehensive information. Meanwhile, secondary data was obtained from official documents, including the Madrasah Operational Curriculum (KOM), teaching modules, Learning Objective Sequences (ATP), learning assessment tools, the ma'had activity schedule, the student handbook, internal policy documents, and reports on Arabic language program activities.

Data collection was conducted using three techniques: participatory observation, *in-depth* interviews, and documentary analysis. Observations were conducted directly on the implementation of Arabic language instruction in the classroom and language immersion activities within the boarding school environment, including the use of learning media such as *Smart TVs*, the application of *fun learning* methods, the implementation of *ta'lim dzu syu'bah*, teacher–student interactions, and the implementation of *bi'ah lughawiyah*. Semi–structured in–depth interviews were conducted with all informants to gather information regarding program planning, implementation, challenges faced, problem–solving strategies, and perceptions of the effectiveness of the Arabic language program. Documentation was used to obtain supporting data in the form of curriculum materials, the Durusul Lughah Al–'Arabiyyah textbook, teaching materials, evaluation documents, photos of activities, and administrative records related to the Arabic language program.

Data analysis utilized the interactive model developed by Miles, Huberman, and Saldana, which was implemented continuously from the data collection process through the drawing of conclusions. The first stage is data condensation, which involves selecting, *coding*, grouping, simplifying, and focusing the data based on the four components of the CIPP evaluation framework so that only data relevant to the research focus is retained. The second stage is *data display*, which involves organizing the results of the reduction into evaluation matrices, categorization tables, and descriptive narratives so that the relationships among findings regarding context, input, process, and product can be systematically analyzed. The third stage is drawing conclusions *and* verification, which involves interpreting patterns, themes, and relationships among the CIPP components, followed by continuous verification through rechecking the field data until consistent, logical conclusions supported by empirical evidence are reached.

Data validity was ensured through source triangulation, methodological triangulation, and temporal triangulation. Source triangulation was conducted by comparing information obtained from the madrasah principal, the ma'had director, the language coordinator, the Arabic language teacher, the the musyri'ah, and the students to assess the consistency of data regarding the management of the Arabic language program. Methodological triangulation is conducted by comparing the results of

observations, interviews, and documentation regarding the same phenomenon—for example, the alignment between the teaching practices observed in the classroom and the teachers' explanations as well as instructional materials. Temporal triangulation is carried out through observations and interviews at different times, both during formal learning activities and language immersion activities in the dormitory environment, ensuring that the data collected is not merely incidental. In addition, the researcher also conducted *member checking* with several key informants to ensure that the researcher's interpretations align with the experiences and information they provided.

The program evaluation analysis was conducted based on indicators for each component of the CIPP model. The *context* component was evaluated using indicators of the program's alignment with the madrasah's vision and mission, students' needs, KMA Regulation No. 450 of 2024, and the objectives of the Arabic language program. The "Input" component includes educators' qualifications and competencies, student characteristics, the integrated curriculum, instructional materials, learning media, facilities and infrastructure, funding, and institutional support. The *process* component covers instructional planning, the implementation of differentiated instruction (*ta'lim dzu syu'bah*), the use of teaching methods, the management of *the bi'ah lughawiyah*, monitoring, supervision, and periodic program evaluation. Meanwhile, the "product" component is evaluated based on the achievement of Phase D learning outcomes, improvements in the four Arabic language skills (*istima'*, *kalam*, *qira'ah*, and *kitabah*), changes in the use of Arabic in daily life at *the ma'had*, and the program's success rate in achieving its established objectives.

RESULTS AND DISCUSSION

RESULTS

Context Evaluation of the Arabic Language Program at Ma'had Darul Hikam

Table 1. Results of the *Context, Input, Process, and Product (CIPP)* Evaluation of the Arabic Language Program at Ma'had Darul Hikam

No	Component	Findings
1	<i>Context</i>	The program was created in response to the real needs of female students who were having difficulty understanding Arabic. The program integrates the modern and Salafi boarding school approaches; however, the program's objectives have not been systematically documented and have not been fully aligned with KMA No. 450 of 2024.
2	<i>Input</i>	The curriculum uses the *Kitab Durusul Lughah Al – Arabiyah*, which aligns with the program's objectives. Teachers are available and activities are conducted regularly; however, the use of technology – based learning tools is not yet optimal because the facilities are not located on the boarding school campus.
3	<i>Process</i>	Instruction employs the " <i>fun learning</i> " method through games, songs, and simple conversations. Implementation has not been sustained because the use of Arabic ceases once the learning activities are over, and a <i>language-rich environment</i> has not yet been established.
4	<i>Product</i>	The program improves students' basic understanding of Arabic and fosters a positive attitude toward learning Arabic, but has not yet resulted in active communication skills in everyday life.

Contextual Evaluation of the Arabic Language Program at Ma'had Darul Hikam

Research findings regarding the contextual evaluation reveal that the Arabic language program initiative at Ma'had Darul Hikam emerged from a strategic policy seeking to integrate two opposing educational models. This program is designed to combine the strengths of modern Islamic boarding schools—which prioritize speaking proficiency (*maharah kalam*)—with the strong tradition of Salafi boarding schools in mastering grammar (*qowa'id*). Based on interview data, the primary foundation of this program is a response to the actual situation of female students who feel alienated from the Arabic language. *"This program is a direct initiative of the Director, who recognized the need to integrate the two systems. Modern boarding schools excel in active language use, while traditional boarding schools excel in qowa'id. We want to combine the two."* (Interview with the Language Coordinator, December 3, 2025). This statement confirms that the program's development is not only based on normative institutional considerations but also stems from the female students' real need to build balanced Arabic language proficiency that integrates both communicative and structural aspects. Interview data indicate that the establishment of the Arabic language program at Ma'had Darul Hikam was an internal institutional policy based on the students' actual needs. This program was designed to integrate the strengths of modern boarding schools—which emphasize speaking skills (*maharah kalam*)—with the traditions of Salaf boarding schools, which focus on mastering grammatical rules (*qawa'id*). Furthermore, the language coordinator explained that *"This program is a direct initiative of the Director. He saw a need to integrate the two systems. Modern boarding schools excel in active language use, while traditional boarding schools excel in grammar. We want to combine the two."* (Interview, December 3, 2025).

The interview results also showed that before the program was implemented, most female students had difficulty understanding basic Arabic vocabulary, which impacted their learning of religious subjects. This situation served as the basis for developing the Arabic language program as a response to the students' needs. A review of program documentation indicates that the program has a development framework that integrates active language learning with grammar mastery. However, the program documents do not yet include objective formulations that explicitly refer to Ministerial Decree (KMA) No. 450 of 2024, nor do they contain systematically documented program success indicators. Based on the observation and documentation data, it can be concluded that the Arabic language program is grounded in the students' real needs; however, the formulation of program objectives remains implicit and has not been comprehensively documented. It is concluded that this initiative emerged as a direct response to the students' actual circumstances, rather than merely as a means to implement formal curriculum policies. This demonstrates that, within its context, the program has fulfilled the principle of a needs-based program—meaning it was designed based on the real needs of its target audience. However, field data also indicate that at the program's inception, the Arabic language program was not entirely aligned with the new national legislation—specifically Ministerial Decree No. 450 of 2024—or with the institution's vision to promote its identity as an academic institution. The absence of program objectives integrated into the regulatory framework resulted in uncertainty regarding the optimal long-term direction of the program.

Based on interview results, it was revealed that the Arabic language program initiative emerged as a direct response to the real-world conditions faced by students—rather than merely a formality in the implementation of national curriculum policies.

Previously, students struggled to understand religious studies due to a lack of mastery of basic Arabic vocabulary, creating an urgent need for more contextualized learning interventions. Thus, this program has fulfilled the principles of a *needs-based program*, in which its design and implementation focus on identifying real problems in the field, rather than merely following bureaucratic directives. The existence of this program reflects educators' awareness of the importance of a humanistic, learner – centered approach, and demonstrates that effective curriculum change should stem from real needs, not merely administrative compliance.

However, findings in the field also indicate that in the early stages of implementation, the Arabic language program has not yet fully aligned with the new national regulatory framework, particularly Ministerial Decree No. 450 of 2024, and has not been fully integrated with the institution's vision as an academic institution seeking to strengthen its scholarly identity. The absence of clearly defined program objectives linked to the regulatory framework creates uncertainty regarding the program's long – term development direction. Consequently, although the program stems from students' real needs, without a systematic connection to formal policies and the institutional vision, it risks losing institutional legitimacy, becomes difficult to evaluate objectively, and is vulnerable to future shifts in policy priorities. Therefore, efforts are needed to reformulate the program's objectives by integrating the principles of field – based needs and regulatory compliance, so that the program can be sustainable while remaining academically relevant.

However, contextual analysis also indicates that the program's implicit objectives align with the national curriculum and the development of 21st – century competencies, particularly in enhancing language skills and fostering a positive attitude toward the Arabic language. The findings suggest that, in this context, the program is aligned but not yet structured, meaning that clear and well – documented operational objectives are still needed, even though the program's vision and rationale are relevant. With more systematic and focused objectives from the outset, the Arabic language program at Ma'had Darul Hikam has the potential to have more focused areas of development and become a valuable resource for more effective and comprehensive learning materials, methods, and media.

Input Evaluation: Readiness of Materials, Human Resources, and Program Support Facilities

Regarding the input component, the data indicates that the learning materials are well – prepared but are hindered by imprecise facility management. The core learning material uses **Kitab Durusul Lughah Al – Arabiyah**, a standard reference used in modern Islamic boarding schools to practice conversation. However, field observations reveal the ineffective use of technology – based learning media. Learning tools such as *Smart TVs* are actually available, but the units are located in the madrasah classrooms, not in the dormitories or ma'had areas. As a result, these tools are very rarely used in afternoon or evening language activities because access is limited and it is impractical to gather the students (Field Observation Notes, November 26, 2025).

From the perspective of input components, the data indicates that this Arabic language program has a solid academic foundation, particularly regarding the readiness of well – developed instructional materials. The core learning materials utilize the textbook **Durusul Lughah Al – Arabiyah**, which serves as a standard reference and has

proven effective in various modern Islamic boarding schools for developing students' Arabic conversational skills. The availability of these systematic and tiered teaching materials is a key asset to the program's success, as teachers do not need to design materials from scratch and students can learn progressively according to their ability levels. However, the strengths of these materials have not been fully maximized due to constraints in another area: inadequate facility management.

Field observations indicate that the program's ineffectiveness actually stems from the use of technology-based learning media that is not suited to the context in which activities take place. Resources such as *Smart TVs* are actually available at the institution, but they are located in the madrasah classrooms, not in the dormitories or ma'had areas—which are the primary locations for afternoon and evening language activities. As a result, access to these resources is severely limited and impractical, since moving students from the dormitories to the madrasah classrooms during hours outside of formal instruction is considered cumbersome and time-consuming. This has led to *Smart TVs* being rarely—if ever—used during afternoon or evening language activities, even though these devices have great potential to support interactive learning, such as showing conversation videos, listening exercises, or vocabulary quizzes. Thus, the gap between the availability of resources and their location of use has become a real obstacle that reduces the quality of program implementation, even though the materials themselves have been deemed highly appropriate.

Limited access and the lack of widespread use of resources hinder the integration of audiovisual media into language learning routines, so that existing resources do not make a significant contribution to enhancing learning input (Qu & Miralpeix, 2025). The findings indicate that, in terms of input, this program has a solid foundation of instructional materials but still needs to reorganize its human resource management and facilities so that all components of the input can work together optimally. Observations reveal that the program uses the **Kitab Durusul Lughah Al-Arabiyah** as its primary learning material. This textbook is consistently used in language activities at the ma'had and serves as a reference for all instructors. In addition to the availability of materials, field observations found that the institution has provided learning media in the form of *Smart TVs*. However, these facilities are located in the madrasah classrooms, whereas language activities take place within the ma'had environment in the afternoons and evenings.

Furthermore, based on the observations, it was found that learning resources such as *Smart TVs* are actually available, but the units are located in the formal school area, not in the dormitories or the ma'had area. As a result, this media is very rarely used in afternoon or evening language activities due to limited access." (Field Observation, November 26, 2025). Interviews with instructors also revealed that the learning process relies primarily on textbooks, whiteboards, and direct communication because audiovisual media are difficult to utilize during ma'had activities. Overall, the data indicate that the input aspect is supported by adequate learning materials; however, the utilization of learning resources remains suboptimal due to limited access to supporting facilities.

Process Evaluation: Implementation of Learning and Program Implementation Challenges

Observations of the learning process indicate that the approach used focuses on *fun learning* methods. The instructor at employs games, singing, and casual everyday conversations to convey the material. However, a crucial finding regarding this aspect of

the process is the discontinuity in implementation. "The lessons use the *Kitab Durusul Lughah* textbook, sometimes supplemented with conversations, singing, or *fun games*. Usually, it's just simple conversations, followed by listening to announcements in Arabic, but these activities are rarely practiced outside the classroom and dormitory because there are no conversation partners" (Interview with an 8th – grade student, November 26, 2025).

Observations of the learning process indicate that, in terms of methodology, this Arabic language program has implemented a targeted approach—namely, *fun learning*. Instructors use a variety of active strategies, such as *games*, group singing, and casual everyday conversations, to convey material from the *Durusul Lughah Al – Arabiyah* textbook. Theoretically, this approach aligns with the principles of second – language learning in the early stages, where a relaxed and interactive atmosphere can reduce students' *language anxiety* and enhance their intrinsic motivation. These positive findings indicate that the program's methodological aspects are not problematic; teachers have made efforts to enliven the classroom with varied, student – centered activities.

However, a crucial finding regarding the process lies in the discontinuity of implementation—specifically, the disconnect between Arabic language learning experiences in the classroom and students' daily lives outside of class hours. Based on an interview with an 8th – grade student, it was found that classroom activities were indeed enjoyable, filled with simple conversations, songs, or *fun games*; the student even admitted to being able to understand Arabic announcements they heard. However, the use of Arabic barely continues when they are outside the classroom and the dormitory. The main cause is the lack of conversation partners, which indicates that *a collective linguistic environment* has not yet been established. As a result, Arabic language learning occurs only as episodic and isolated events within the classroom, rather than as *a habit* or culture that is part of students' daily lives. This situation threatens the sustainability of language acquisition, because without continuous practice, the skills acquired in class tend to fade quickly and never reach the level of *automaticity* required for active communication.

Observations within the ma'had environment reinforce these findings. Outside of class hours, interactions among students are still dominated by the use of Indonesian, so the practice of communicating in Arabic has not yet become a daily habit. Thus, the research results indicate that the classroom learning process has proceeded in accordance with the program design, but the implementation of Arabic language use has not carried over into daily life within the ma'had environment. These findings highlight interesting variations in teaching methods and also indicate that Arabic language learning at Ma'had Darul Hikam remains situational in nature and has not yet been systematically institutionalized. The engaging teaching methods used by teachers have proven to increase student participation during activities; however, their effectiveness is limited by a lack of language practice outside the formal classroom setting (Han & Li, 2025). This points to the need for a more comprehensive management approach, particularly in enhancing language activities in the dormitories in conjunction with classroom instruction. Consequently, students receiving instruction do not receive reinforcement through written language, so the language internalization process only reaches a level of passive understanding that does not develop into communicative habits. This indicates that the effectiveness of the learning process is not only determined by the effectiveness of the methods but is also influenced by the consistency of implementation and a supportive learning environment.

Product Evaluation: Students' Arabic Language Competency Achievement

The results of the product evaluation reveal mixed outcomes. On the one hand, the program succeeded in shifting students' perceptions from a state of ignorance to a basic level of understanding. However, this improvement in comprehension has not yet fully transformed into active communicative competence, particularly in the spontaneous use of Arabic outside the classroom context. Field findings indicate that most students remain proficient in speaking practice within the dormitory environment. Arabic is still used passively, meaning people may understand and respond to it in various ways, but it has not yet become a productive means of communication in daily life. This situation suggests that the program's outcomes focus more on the development of competencies (knowledge and understanding) than on practical skills (language proficiency).

The results of the program's product evaluation or learning outcomes present a mixed picture, meaning there are successes on one hand but also significant limitations on the other. On the one hand, the program succeeded in transforming students' initial perceptions—from a lack of knowledge regarding Arabic vocabulary and language structure—into a fairly solid foundational understanding. Students are able to recognize common words, understand simple instructions, and grasp the meaning of Arabic announcements made within the madrasah environment. This achievement is no small feat, considering that previously the students lacked adequate Arabic language skills. In other words, the program is effective in building a cognitive foundation and declarative knowledge of Arabic as a linguistic system.

However, this improvement in comprehension has not yet fully translated into active communicative competence, particularly regarding the spontaneous use of Arabic outside the classroom context. Field findings indicate that most students are still not proficient in speaking (*productive skills*) within the boarding school environment. Arabic is used more passively—meaning students may understand what they hear or read and are able to respond in limited ways, such as through gestures or short answers—but it has not yet become a means of productive communication in their daily lives. Spontaneous conversations in Arabic among students—for example, during meals, breaks, or play—are almost nonexistent. This situation clearly indicates that the program's outcomes focus more on developing *competence* (knowledge and understanding of the language) than on *performance* or practical *skills*. In fact, the ultimate goal of language learning should be active communicative ability, not merely passive comprehension. This gap between comprehension and oral production poses a serious challenge that must be addressed by redesigning learning strategies to better encourage active practice and the creation of an environment

The interview results show that the program has improved students' basic understanding of the Arabic language. Most students reported that they now find it easier to understand Arabic vocabulary, teachers' instructions, and announcements in Arabic than they did before participating in the program. One student remarked, "Now, when I hear announcements in Arabic, I'm starting to understand them." (Interview, November 26, 2025). Nevertheless, observation results indicate that students' active communication skills remain limited. The use of Arabic in daily interactions within the ma'had environment has not yet become a habit, so spontaneous conversations among students are rarely observed. Data from observations and interviews indicate that the program's achievements are more evident in improved comprehension (*receptive competence*) than in active speaking ability (*productive competence*). Overall, the product evaluation results show that the program has successfully improved students' basic Arabic language skills

and fostered a positive attitude toward learning Arabic. However, these results have not yet fully translated into active communication skills in daily life within the ma'had environment.

Overall, the findings of this evaluation indicate that the program has successfully built basic Arabic language skills; however, it has not yet succeeded in fostering sustainable behavioral change in language use. Successfully integrating this program is not easy and requires specialized research to ensure optimal program outcomes. The program's outcomes still focus on basic comprehension and a positive attitude toward the Arabic language, while mastery of active communication skills requires further support through practice, consistency, and an environment where Arabic is used more frequently. These findings underscore that the success of Arabic language learning is not only assessed based on comprehension of the material but also on students' ability to apply the language functionally in everyday social contexts.

DISCUSSION

Evaluation results based on the CIPP model indicate that the effectiveness of Arabic language program management at Ma'had Darul Hikam is determined not only by the quality of instructional delivery but also by the integration of context, input, process, and product as an interconnected system. These findings reinforce *Stufflebeam's* view that the success of an educational program cannot be measured solely through final outcomes but must be analyzed through the relationships among the program's constituent components. Thus, the various challenges identified are not isolated issues but reflect the suboptimal integration of the Arabic language program management system as a whole.

In the *context* dimension, the alignment of program objectives with students' needs indicates that program design was oriented toward *needs assessment*, rather than merely fulfilling the institution's administrative requirements. From a CIPP evaluation perspective, this suggests that program development decisions were based on the identification of real needs, thereby possessing strong academic legitimacy. This finding also demonstrates that the integration of Salaf pesantren traditions with modern learning approaches represents a form of curriculum adaptation that is responsive to changes in student characteristics. This phenomenon aligns with Stephen Krashen's *Affective Filter Hypothesis*, which explains that success in second–language acquisition is heavily influenced by students' affective states. When the learning process is able to increase motivation, build self–confidence, and reduce anxiety, *language input* is more easily processed into language proficiency. Therefore, the success of the “*fun learning*” approach is not solely due to the use of games or songs, but because this approach successfully creates a psychological climate conducive to language acquisition.

The above findings also expand upon the results of Aristya's (2026) study, which emphasized the importance of affective factors in foreign language learning. While that study focused on the effectiveness of classroom learning strategies, this study demonstrates that affective factors are also influenced by program management policies that provide teachers with the flexibility to develop more adaptable learning approaches. Thus, the contribution of this study lies in affirming that the success of Arabic language learning is determined not only by teachers' pedagogical competence but also by managerial decisions in designing program objectives and directions that align with learners' needs.

In the *input* dimension, the research findings indicate that a high–quality curriculum does not necessarily result in an effective program if it is not supported by

adequate resource management. From a CIPP perspective, input is not merely understood as the availability of resources but also involves the proper management of those resources to ensure all components are utilized optimally. The mismatch between the location of learning facilities and the center of students' activities indicates that the main issue lies not in limited facilities, but in poor management of facility utilization. This situation reveals a gap between facility planning and the operational needs of daily program activities.

These findings support the research by Puteri et al. (2025), which explains that the effectiveness of learning media is largely determined by the accessibility of its use, not merely by the availability of devices. However, this study offers a broader perspective by demonstrating that the management of facilities in a boarding school environment presents different complexities compared to regular schools. In addition to physical access, coordination among madrasah administrators, ma'had administrators, and musrifah plays a crucial role in ensuring the optimal utilization of facilities. Therefore, the input – related issues identified in this study are better understood as matters of resource management rather than merely limitations in facilities.

In addition to facilities, the readiness of human resources is also a factor influencing the program's success. Findings regarding the suboptimal technical guidance provided to musrifah indicate that the development of Arabic language programs requires a more structured mentoring system. Within the framework of educational management, human resources are a strategic component that determines the program's sustainability, as they serve as the primary implementers of the policies that have been designed. The contribution of this study is to demonstrate that evaluating inputs at integrated educational institutions should not be limited to an inventory of facilities and teaching staff but must also examine the quality of coordination, task distribution, and mentoring mechanisms among program administrators.

In terms of the *process* dimension, the research results indicate that the program's main issue does not lie in the teaching methods but rather in the lack of continuity in the language habituation process outside the classroom. From the perspective of language acquisition theory, communication skills do not develop solely through formal learning but require continuous interaction within a social environment. Therefore, the lack of a developed Arabic – speaking culture (*bi'ah lughawiyyah*) causes the language acquisition process to stall at the stage of understanding the material without progressing into a habit of communication. This finding reinforces the research by Mahmud (2025), which states that the language environment is a dominant factor in the success of Arabic language learning in Islamic boarding schools. However, this study indicates that the formation of a *bi'ah lughawiyyah* cannot be achieved solely through language use policies; rather, it requires a management system that integrates learning programs, dormitory guidance, supervision of musrifah, and the communication culture of the entire ma'had community. In other words, the existence of a language environment is the result of a systematically designed managerial process, not something that forms naturally.

The implications of these findings suggest that the success of Arabic language programs in integrated educational institutions requires a shift in management orientation—from a focus on learning activities to the management of a language ecosystem. In this context, the dormitory functions not only as a place of residence for students but also as a space for language acquisition that facilitates continuous communication practice. This perspective constitutes a conceptual contribution of the study because it broadens the meaning of Arabic language program management from merely managing academic activities to managing the learning environment as a whole.

In terms of the “*product*” dimension, program outcomes that remain dominated by receptive skills indicate that learning success indicators do not yet fully reflect the primary goal of language mastery as a tool for communication. In communicative competence theory, mastery of language rules does not automatically result in the ability to use the language in real – life situations. This situation highlights a gap between *competence* and *performance*: students have understood the material but have not had sufficient opportunities to apply those skills in daily interactions. The results of this study differ from some previous evaluations of Arabic language programs, which generally assessed success based on the achievement of learning targets or improvements in academic scores. This study demonstrates that a program's success should be measured by its ability to foster language – use habits in students' daily lives. Thus, product evaluation is insufficient when relying solely on cognitive indicators; it must also incorporate language behavior indicators as a measure of program success.

Overall, this study contributes to the development of Arabic language program management in educational institutions that integrate madrasahs and ma'had. Unlike previous studies, which primarily evaluated the learning process or curriculum implementation, this study demonstrates that program effectiveness is determined by the integration of policy, resource management, the creation of a language environment, and the mentoring system in the dormitories. These findings expand the application of the CIPP model within the context of integrated Islamic education and yield practical implications that the development of Arabic language programs should be directed toward building a sustainable language ecosystem. Thus, the success of a program is not merely measured by the implementation of learning activities but by the program management's ability to foster a culture of Arabic language use as an integral part of students' daily lives.

Research Limitations

This study has several limitations that must be considered when interpreting the findings. *First*, the study was conducted at a single location, namely Ma'had Darul Hikam MTsN in Batu City; therefore, the findings are best understood as a contextual snapshot and cannot yet be directly generalized to all educational institutions that integrate madrasahs and pesantren. *Second*, the study employed an evaluative qualitative approach; thus, the assessment of program outcomes was primarily based on interviews, observations, and documentation, and was not supplemented by longitudinal quantitative measurements of Arabic language proficiency. *Third*, the program evaluation was conducted during a specific research period, so the program's development after that period could not be fully observed. These limitations also open the door for future research to conduct comparative studies across several ma'had, using a mixed – methods design, and to conduct longitudinal evaluations to assess the sustainability of the program's impact on students' communicative skills.

CONCLUSION

Based on a comprehensive evaluation using the CIPP model, this study concludes that the management of the Arabic language program at Ma'had Darul Hikam has a strong conceptual foundation but faces challenges in technical and operational aspects. In the context dimension, the program proved highly relevant to the vision of integrating modern and Salafi curricula and aligns with KMA Regulation No. 450 of 2024, where the program's objectives are valid in addressing student alienation through an approach that reduces

affective filters. However, an evaluation of the input aspect revealed gaps in facility management; although the “Durusul Lughah” teaching materials were deemed appropriate, their effectiveness was hampered by the placement of supporting facilities—such as *Smart TVs*—which were difficult to access in the dormitory areas, as well as a lack of technical guidance for the female instructors.

Furthermore, the evaluation of the process and product dimensions revealed mixed results. The implementation of the “*fun learning*” method successfully created an enjoyable learning atmosphere, but the practice was interrupted due to the lack of a consistent language environment (*bi’ah lughowiyah*) outside of class hours. Consequently, the program’s final outcomes were ambivalent: the program succeeded in cognitively transforming students’ perceptions and foundational understanding, but failed to convert those competencies into active communication skills within the ma’had’s social environment, resulting in students tending to be passive in their daily speaking practices. These findings underscore a new insight: in an integrated education model, successful language mastery cannot rely solely on an adaptive curriculum or engaging classroom methods; rather, it absolutely requires the engineering of the social environment. The core idea emerging from this research is the urgency of transforming the dormitory into a language laboratory, where accessibility to facilities and environmental familiarization are key to bridging the gap between theoretical knowledge of language rules and the application of that language in daily activities.

REFERENSI

- Abdilah, A. J. (2025). Evaluasi Kurikulum Bahasa Arab dengan Model CIPP (Context, Input, Process, and Product). *Risenologi*, 10(1), 52 – 68.
- Aristya, S. (2026). The CIPP Model for Curriculum Evaluation in Islamic Education: A Comprehensive Literature Review. *Al-Hayat: Journal of Islamic Education*, 10(1), 52 – 66.
- El Zahraa, F. (2026). Teaching Arabic to Non – Native Speakers: Pedagogical Strategies, Challenges, and Prospects. 61–46, (2)6, *مجلة الشرق الأوسط للعلوم التربوية والنفسية*.
- Fauzi, A., & Inayati, N. L. (2023). Implementasi Evaluasi Pembelajaran Pendidikan Al Islam di Sekolah Menengah Atas Muhammadiyah. *Munaddhomah: Jurnal Manajemen Pendidikan Islam*, 4(2), 272 – 283. <https://doi.org/10.31538/munaddhomah.v4i2.438>
- Han, H., & Li, B. (2025). Perceptions of social interactions and second language willingness to communicate in different activities. *Asian-Pacific Journal of Second and Foreign Language Education*, 10(1), 4.
- Jaafar, A., Deni, E. P., Febriani, A., Lestari, R., Yelliza, M., & Sari, W. W. (2023). Problems of Learning Arabic in Islamic Boarding Schools. *International Journal of Multidisciplinary Research of Higher Education*, 6(3), 147 – 154. <https://doi.org/10.24036/ijmurhica.v6i3.141>
- Leak, A. E., Lenaburg, L. A., Sciaky, E., & Paden, M. O. A. (2025). Planning and proposing evaluations to support project success. *BMC Proceedings*, 18(Suppl 1), 31.
- Mahmud, I. S. (2025). *Evaluasi Program Pembelajaran Bahasa Arab Dan Inggris Pada Kelas XII Mapk Berdasarkan Model Cipp (Context, Input, Process, Product) Di Pondok Pesantren An-Nawawi Berjan Purworejo*. Universitas Islam Indonesia.
- Mahmudi, I., Fadilah, S., & Anwar, H. S. (2023). Evaluation of the Idhaf Lughawi Program at Islamic Boarding School Gontor For Gilrs 1 Using the CIPP Model. *Educan : Jurnal Pendidikan Islam*, 7(1), 14. <https://doi.org/10.21111/educan.v7i1.8956>
- Muhammad Zikran Adam, Syuhadak, Siti Aliyya Laubaha, & Nicky B. Lasanggo. (2024). Kurikulum Bahasa Arab KMA 183 pada Tingkatan Madrasah Aliyah. *AL-KILMAH*, 3(1), 63 – 76. <https://doi.org/10.58194/alkilmah.v3i1.1580>

- Nurhayati, N., & Usiono, U. (2025). Evaluation of Arabic Language Learning at Madrasah Tsanawiyah Ali Imron Bandar Selamat School Tembung – North Sumatra. *Ihya Al-Arabiyyah: Jurnal Pendidikan Bahasa Dan Sastra Arab*, 10(1), 68 – 77.
- Puteri, A. R., Nasution, W. N., & Nasution, M. I. P. (2025). Integrasi teknologi informasi dan komunikasi dalam pendidikan: konsep, perkembangan, dan inovasi media pembelajaran. *Jurnal Pendidikan Indonesia: Teori, Penelitian, Dan Inovasi*, 5(4).
- Qu, Y., & Miralpeix, C. M. I. (2025). *Audiovisual input and second language learning*. Routledge Journals, Taylor & Francis Ltd 2–4 Park Square, Milton Park
- Sumadi, E., & Huda, M. (2025). Curriculum Management and Implementation of Integrative Learning in Islamic –Based Madrasas in Central Java. *Quality*, 13(1), 75 – 92.

Wawancara

Inisial PTF (Koordinator Program Bahasa), *Wawancara*, Minggu, 10 Desember ruari 2025

Darissalamah, N., Anisah, L.N., (Musyrifah Ma'had). *Wawancara*. Rabu 12 November 2025

Tiga siswi inisial E, H, S, (Siswi kelas 8 MTsN Kota Batu). *Wawancara*. Rabu 26 November 2025