

Policy Challenges and Implementation Gaps in Arabic Proficiency Testing at Indonesian Islamic State Universities

Received 2025-12-16

Accepted 2026-02-30

Published 2026-07-20

Herdah¹, Raodhatul Jannah^{*2}, Darmawati³

¹²Pendidikan Bahasa Arab, IAIN Parepare, Indonesia

herdah@iainpare.ac.id¹, raodhatuljannah@iainpare.ac.id^{*2},

darmawati@iainpare.ac.id³

To cite this article: Herdah, Herdah., Jannah, Raodhatul., Darmawati, Darmawati. (2026). Policy Challenges and Implementation Gaps in Arabic Proficiency Testing at Indonesian Islamic State Universities. *Ijaz Arabi: Journal of Arabic Learning*, 9 (3), 1346-1360, DOI: <https://doi.org/10.18860/ijazarabi.V9i3.41203>

Abstract

The implementation of Arabic proficiency testing in Indonesian Islamic State Universities presents several challenges and opportunities for improving students' language proficiency. This study explores the experiences of students and educators in navigating the Arabic proficiency tests and examines the implications for institutional and pedagogical practices. The study focuses on three major challenges in the implementation of testing policy: (1) inconsistency in testing standards across institutions, (2) misalignment between the test content and the Arabic curriculum, and (3) inadequate institutional support for student test preparation. This study adopts a qualitative case study design of Arabic proficiency testing at the State Islamic Institute of Parepare (IAIN Parepare). Data were collected through policy analysis, test results, and interviews with students from various academic programs. The findings indicate that students often perceive the test as difficult due to unfamiliar test formats and content that exceeds the curricular scope. Moreover, educators struggle to integrate test-oriented skills into regular instruction due to the lack of clear institutional guidelines. The analysis reveals that, while the tests play a critical role in assessing students' language skills, these challenges hinder the policy's effectiveness. Despite these challenges, the study offers recommendations for improving the standardization of testing procedures, aligning the curriculum more closely with the tests, and enhancing institutional support to prepare students better. Addressing these issues will help create more equitable and accurate assessments of Arabic proficiency across all institutions.

Keywords: Arabic proficiency test; educational policy; language testing; standardization test; Islamic state university

INTRODUCTION

Arabic proficiency tests are a mandatory requirement in Indonesian Islamic State Universities, aimed at enhancing students' mastery of Arabic, which is critical for Islamic studies and religious practices. This policy is crucial, especially as a prerequisite for students who wish to continue their studies in the Middle East, making it highly sought after by students. According to the Ministry of Religious Affairs, the number of applicants for Middle Eastern scholarships has consistently increased each year, reflecting the growing interest in these tests. (*Seleksi Timur Tengah - Kemenag*, n.d.) (*Seleksi Timur Tengah - Kemenag*, n.d.) Data from the scholarship selection process between 2019 and 2022 show a significant increase in the number of applicants, indicating the enduring appeal and importance of Arabic proficiency tests as a key element in Islamic education policies in Indonesia (Hefner, 2009).

Previous research on Arabic proficiency tests has often focused on the analysis of test items and issues related to implementation. Studies show that the administration of these tests is inconsistent across institutions, resulting in disparities in the quality of assessment and students' proficiency in Arabic (Hasan, 2021). For example, a study found that the standards of Arabic proficiency tests vary between universities, leading to unequal assessment outcomes (Yusri, 2013). Other research indicates that variations in test implementation can affect students' motivation to learn (Cheng, 2009). Furthermore, reports show that the lack of standard assessment practices is a significant factor contributing to the differences in test results across institutions (Norris, 2018).

The policy of Arabic proficiency testing at Islamic religious universities in Indonesia aims to ensure that students master the Arabic language as an integral part of Islamic studies. This policy is designed to assess students' Arabic language skills in both academic and religious contexts, as well as serve as a prerequisite for further studies in the Middle East. The implementation of this policy varies in terms of assessment standards across institutions. The Indonesian Ministry of Religious Affairs (2022) reported that the Arabic proficiency testing policy is a strategic step to ensure that students meet the required Arabic language standards for Islamic studies. Data indicate an increase in the number of applicants for scholarships to study in the Middle East, reflecting the high demand and relevance of this test (Raish, 2021)(Muradi et al., 2022). Researchers have noted that this testing policy is designed to assess Arabic language competence in both academic and professional contexts, which is crucial for preparing students for further studies and career development. However, a study by Maulana found that despite the policy's goal of improving Arabic language mastery, its implementation varies among institutions, impacting the consistency and quality of assessments across universities (Indriana & Maksum, 2023). As such, the Arabic proficiency testing policy at Islamic religious universities serves to standardize students' Arabic proficiency, although the execution and assessment standards can differ.

The implementation of Arabic proficiency tests at Islamic religious universities often faces challenges related to standardization and effectiveness. Variations in test methods and assessment standards can affect test outcomes and students' language skills. These differences highlight the challenges of achieving consistency in assessments across institutions. Maulana identified that standardization issues in the implementation of Arabic proficiency tests at Indonesian universities result in uneven assessments and test outcomes. A study found that variations in test methods and assessment standards across higher education institutions affect the effectiveness of these tests, thus impacting students' mastery of the Arabic language. Rohfin emphasized that discrepancies in test implementation lead to significant differences in assessment outcomes between institutions, indicating a challenge in achieving consistency and standardization in the administration of these tests (Gestanti, Mufanti, & Nimasari, 2019). Therefore, the standardization and effectiveness of Arabic proficiency tests remain challenges that need to be addressed to improve the quality of assessments in higher education.

Government policies and the implementation of Arabic proficiency tests affect the quality of education and students' academic preparedness. Although these policies are designed to improve students' Arabic language mastery, ineffective implementation can hinder their readiness for further studies or professional careers. Evaluating these policies and test practices can provide solutions to these issues. Alfian (2018) studied the impact of Arabic proficiency test policies on educational quality in higher education and found

that these policies play a role in enhancing Arabic language mastery, although challenges remain in their implementation (Gay & Betts, 2020). Muhammad Abdel Latif observed that the policies and practices of test administration often do not align with expected standards, which can affect students' academic preparedness for further studies or professional careers (Abdel Latif & Alhamad, 2023). They noted that a comprehensive evaluation of test policies and practices is needed to identify and address issues affecting test outcomes and effectiveness, ensuring that these tests meet both academic and professional objectives. Evaluating Arabic proficiency test policies and practices is essential for identifying and addressing the issues that impact the quality of education and students' academic preparedness.

Standardization and Effectiveness of Arabic Proficiency Testing at Islamic Religious Universities are crucial to ensuring fair and reliable assessments of students' language skills. Standardization refers to the application of consistent criteria and assessment methods across institutions, which guarantees that the tests are accurate and equitable. Research by O'Loughlin highlights that inconsistent standard in language testing lead to unreliable assessments, particularly when institutions adopt different guidelines (May, 2011). Furthermore, Harsch point out that variations in test administration and scoring compromise the overall reliability of the tests, creating challenges in achieving consistent and accurate assessments across different universities (Mislevy, Carlson, Bridgeman, Golub-smith, & Stone, 2017). Thus, improving the standardization of Arabic proficiency tests is essential for ensuring a level playing field in student evaluations.

Variability in the implementation of Arabic proficiency tests affects the assessment outcomes and students' language mastery. This variability includes differences in how the tests are administered, the criteria used for evaluation, and how the results are interpreted across various institutions. As noted by Wafa, differences in test administration methods can lead to disparities in test results, resulting in an unequal assessment of students' language proficiency (Alowaydhi, 2016). Yonggang found that inconsistencies in the evaluation of language tests affect students' readiness for academic and professional challenges (Ren, 2014), highlighting the importance of addressing these discrepancies to ensure fair assessments.

The effectiveness of Arabic proficiency tests is linked to their ability to predict students' readiness for advanced studies or professional careers. An effective test should accurately assess students' Arabic language skills and predict their preparedness for academic or professional challenges. According to Pill and Harding, the effectiveness of language tests is measured by how well they predict students' future success in their studies or careers (Pill & Harding, 2013). A poorly designed test, as Liu suggest, can create a gap between theoretical knowledge and practical skills (Ren, 2014), potentially hindering students' professional and academic progress. Enhancing the design and methodology of Arabic proficiency tests is crucial to ensure they accurately measure students' preparedness for future endeavors.

The Arabic proficiency testing policy at Islamic religious universities in Indonesia impacts students' Arabic language skills and academic preparedness. This policy is designed to enhance students' Arabic proficiency while preparing them for academic and professional demands. However, the effectiveness of this policy can vary depending on its implementation across educational institutions. Thu Bui explains that consistent policy implementation is key to improving students' language proficiency, and any variations

can lead to differing academic outcomes across institutions (Bui, 2022). Rachael found that inconsistent application of the policy resulted in significant differences in academic preparedness (Ruegg, Hoang, & Petersen, 2024). Moreover, Nicola concluded that differences in policy application affected the overall effectiveness of Arabic proficiency tests, with some institutions showing better results than others (McLelland, 2021). To maximize the impact of Arabic proficiency testing, standardization across universities is essential.

Variability in the implementation of Arabic proficiency tests affects test results and academic readiness differently. Differences in test implementation between institutions can lead to varied impacts on students' language skills and preparedness. Greg noted that discrepancies in test implementation produce unequal outcomes in student preparedness, underscoring the importance of a standardized policy across institutions (Miller, 2018). Michael emphasizes that inconsistent application often reduces the effectiveness of proficiency tests (Posner, 2011), while Marja and Paivi pointed out that inconsistent policies hinder students' Arabic proficiency (Niitemaa & Pietilä, 2018). Therefore, reducing variability and improving standardization of the policy is necessary to ensure fair and effective assessments.

The impact of government policies on Arabic proficiency test outcomes has not been thoroughly explored. Variations in assessment standards can affect the quality of learning and students' preparedness for future studies or careers (Ritonga, 2020). This issue can create a gap between academic expectations and students' actual abilities. Research has identified that students' lack of understanding of the test format is one of the main challenges affecting the effectiveness of these tests (Musthofa, 2022). Additionally, technical and administrative difficulties have been noted as contributing factors that influence test outcomes (Mohamed, 2023). Another study highlights that differences in the implementation of testing policies across institutions can lead to disparities in learning outcomes (Schissel, 2019). These findings emphasize the need for further research to better understand how policy implementation affects Arabic proficiency tests and students' language abilities.

This article analyzes the challenges in achieving the goals of Arabic proficiency testing in Indonesia. By evaluating test results, government policies, and potential solutions, the study aims to fill gaps in previous research. The analysis considers how inconsistencies in assessment standards impact students' academic success. Moreover, inconsistent implementation of policies has been found to hinder the achievement of educational objectives (Spolsky, 2019)(Sánchez, 2018). Reports from relevant authorities also underline the necessity for improved assessment standards to enhance students' learning outcomes (Jong, 2018).

METHOD

This study employs a qualitative case study design to investigate the policy challenges and implementation gaps in Arabic proficiency testing at Indonesian Islamic State Universities, with a specific focus on the State Islamic Institute of Parepare (IAIN Parepare). The case study approach was chosen to enable an in-depth exploration of how Arabic proficiency testing policies are interpreted and applied at the institutional level, and to uncover the practical challenges that arise during their implementation. By examining IAIN Parepare as a representative case, this study aims to reveal systemic issues affecting the standardization and effectiveness of Arabic language assessment

across similar institutions in Indonesia. The findings are expected to highlight the broader implications of policy-practice gaps and their impact on students' language proficiency outcomes.

Data collection was conducted through multiple methods to provide a comprehensive view of the research problem. The primary data source was the TOAFL (Test of Arabic as a Foreign Language) scores from the IAIN Parepare Language Center, which provided the test results of 659 students who took the exam in 2024. This data set was essential for understanding student performance across the five stages of the TOAFL test. From this pool of students, 23 were selected for interviews, with four students chosen from each of the five test stages, representing different levels of proficiency. This selection ensured that the interview sample captured a range of experiences and provided diverse perspectives on the challenges faced during the test.

Observations of the TOAFL test administration at IAIN Parepare were conducted to gain insights into the procedures, logistical arrangements, and overall testing environment. These observations were crucial in understanding how the tests were practically implemented, how closely they aligned with institutional policy, and how they impacted students' test-taking experiences. In addition to the student interviews, semi-structured interviews were conducted with two Arabic language instructors at IAIN Parepare. These instructors were interviewed to gain insights into the Arabic language curriculum and its alignment with the proficiency test. The interviews explored the structure of the curriculum, the preparation of students for the TOAFL test, and the challenges faced in achieving the desired proficiency levels. The instructors' perspectives helped contextualize the student experiences and provided valuable information on how well the curriculum supported language learning and proficiency development.

A desk review of government policies and institutional regulations related to Arabic proficiency testing was also conducted. This included national policy documents, institutional guidelines, and reports from the IAIN Parepare Language Center. The review provided a broader policy context for the TOAFL test and allowed for the comparison of policy goals with practical outcomes. The data were analyzed following the Miles and Huberman framework, which includes three key stages: data reduction, data display, and conclusion drawing/verification (Greene, 1989).

1. **Data Reduction:** The data from TOAFL scores, student and instructor interviews, observations, and policy documents were organized by thematic areas such as "test implementation," "curriculum alignment," "policy impact," and "student experiences." Irrelevant or redundant data were eliminated, and qualitative coding techniques were used to group the data into categories that highlighted key patterns, such as inconsistencies in test administration, challenges in curriculum alignment, and gaps in institutional support.
2. **Data Display:** The reduced data were then presented in tables, matrices, and charts to facilitate comparison between different data sources. For example, student interview responses were compared with TOAFL scores, observations from the test process, and the curriculum insights provided by the instructors. This allowed for the identification of trends and discrepancies between policy intentions, curriculum design, and actual student experiences.
3. **Conclusion Drawing and Verification:** Conclusions were drawn through triangulation of the data from multiple sources, including student interviews, instructor feedback, TOAFL scores, and policy reviews. The analysis verified whether the curriculum

supported test preparation and identified areas where test standards and student preparation diverged. The conclusions were continually reassessed to ensure they were grounded in the data and supported by multiple sources of evidence.

This data collection and analysis process provided a robust framework for understanding the effectiveness of Arabic proficiency testing at IAIN Parepare, highlighting the critical role of both curriculum design and policy implementation in shaping student outcomes.

RESULTS AND DISCUSSION

Analysis of Government Policy on Arabic Proficiency Testing

The Indonesian government has established various policies to standardize Arabic proficiency and curriculum development within Islamic religious higher education institutions. These policies emphasize the integration of Arabic language skills into academic programs, ensuring that graduates are well-equipped for both religious and professional endeavors. Through comprehensive guidelines, these documents set the expectations for student competencies, outlining the essential skills needed to succeed in an increasingly globalized academic and work environment. The following table provides an overview of key documents that guide the implementation of these policies in higher education.

Table 1. Government Policy Implementation Documents for Arabic Proficiency and Curriculum Standards in Islamic Religious Higher Education Institutions

Document Name	Description	Link
Graduate Competency Standards (SKL) and Learning Outcomes (CPL) for bachelor's degree Programs at Islamic Religious Higher Education Institutions and Faculties of Islamic Studies	This document outlines the graduate competency standards for undergraduate programs at Islamic religious higher education institutions, including the expected Arabic language skills.	http://lpm.uinjambi.ac.id/media/files/2019/07/SKP_DAN_CPL.pdf
Decree of the Director General of Islamic Education Number 1521 of 2022 on the Implementation of the Independent Campus-Merdeka Learning Program at Islamic Religious Higher Education Institutions	The primary competencies for each study program: Ability to communicate orally and in writing using Arabic and English in academic and professional settings.	https://lp2m.uin-malang.ac.id/wp-content/uploads/2022/10/Juknis-Implementasi-MBKM-PTKI-2022.pdf
Curriculum Development Guide for PTKI Referring to KKNi and SN-Dikti	The learning outcomes for general skills in every undergraduate program include "Ability to communicate orally and in writing using Arabic and English in academic and professional development."	https://lpm.walisongo.ac.id/wp-content/uploads/2019/02/Panduan-Pengembangan-Kurikulum-PTKI-Kemenag.pdf
Decree of the Director General of Islamic Education Number 6994 of 2018 on the National Religious Research Agenda 2018-2028	In terms of skill qualifications, PTKI graduates should have general Islamic religious skills in the community, including a. Reading and writing the Qur'an; b. Proficiency in Arabic; c. Practicing worship; d. Islamic arts; e. Preaching; f. Adequate leadership	https://lp2m.uin-malang.ac.id/wp-content/uploads/2019/02/4.-sk-diktis-2019-4-sk.pdf

skills; g. Collaboration with various parties; h. Solving life problems.
--

The Indonesian government's policy emphasizes the importance of Arabic language proficiency for students at Islamic religious universities. The document titled "Graduate Competency Standards (SKL) and Learning Outcomes (CPL) for bachelor's degree Programs" stipulates that Arabic proficiency must be a key part of every student's learning outcomes, making it an essential skill in religious higher education. This policy guides universities to ensure that students can effectively communicate in Arabic as a primary graduation requirement.

In the "Decree of the Director General of Islamic Education Number 1521 of 2022," Arabic proficiency is listed as one of the core competencies under the Independent Campus-Merdeka Learning Program. This policy mandates that students at Islamic universities must be proficient in both Arabic and English to meet academic and professional demands. The inclusion of Arabic proficiency reflects the government's priority to equip students with language skills that are relevant for global careers. Additionally, the Curriculum Development Guide for PTKI underscores that Arabic language proficiency must be included in the learning outcomes for every study program. This ensures that Arabic language acquisition is integrated into the curriculum, aligning with the government's objective to prepare graduates who are not only competent in Islamic studies but also ready to contribute internationally. This policy highlights the government's commitment to enhancing the quality of religious higher education in Indonesia by stressing the importance of Arabic language proficiency. Subsequently, the implementation of these policies is explored and adapted by individual institutions.

Table 2. The implementation of Indonesian government policies at IAIN Parepare

Implementation	Description
UM-PTKIN and Independent Tests	Arabic literacy test as a requirement for university admission.
Arabic Language Course	A mandatory 2-credit course for all study programs.
TOAFL Test	A minimum score of 375 is required for non-Arabic language programs, and 400 for Arabic language programs, as a prerequisite for the final thesis examination.
Final Thesis Examination	Passing the TOEFL test is required to be eligible for the final thesis examination.

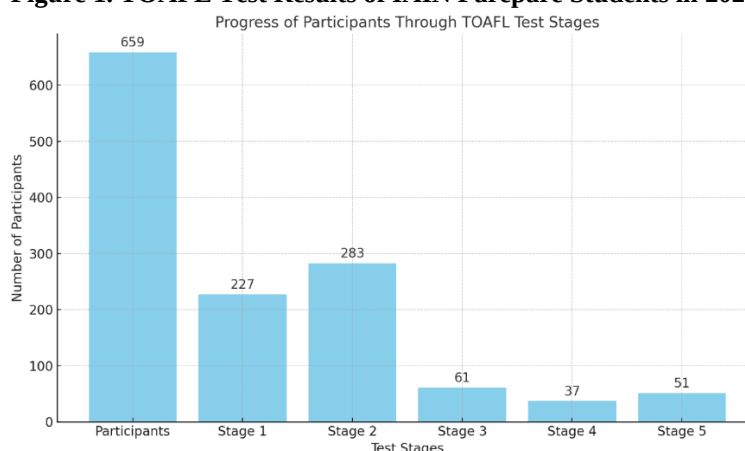
Table 2 outlines the implementation of government policies at IAIN Parepare, focusing on Arabic literacy and academic requirements. It highlights the structured progression from admission to graduation, with Arabic literacy evaluated during the entry process, and mandatory Arabic courses integrated into the curriculum. Additionally, proficiency in Arabic is emphasized as a key component of academic success, with specific TOAFL scores required for thesis completion. This reflects the institution's commitment to aligning with national standards while ensuring students' readiness for both academic and professional challenges.

Standardization and Effectiveness of Arabic Proficiency Testing: An Analysis of Challenges and Variability

To understand the effectiveness and standardization of Arabic proficiency testing at IAIN Parepare, an analysis of the TOAFL test results conducted by the UPT Bahasa is essential. This test is designed to assess the Arabic language proficiency of students, particularly in an academic context. However, the results reflect several challenges related

to the administration of the test and the preparation of students. Below is a summary of the TOAFL test results taken by IAIN Parepare students in 2024.

Figure 1. TOAFL Test Results of IAIN Parepare Students in 2024



Source: UPT Bahasa IAIN Parepare

The results of this test indicate that out of 659 students who participated, only 36% passed the first stage. This raises concerns regarding the effectiveness of the test and students' readiness to meet the established standards. Although the TOAFL test is designed to comprehensively assess Arabic proficiency, the outcomes reveal significant variability in student achievement.

One possible factor contributing to the low pass rate is the lack of standardization in both the preparation and administration of the test. Despite TOAFL being designed as a standardized measure, its implementation may vary depending on the resources and support available within each academic program. Variations in Arabic language teaching methods, the quality of materials, and the intensity of practice can significantly influence student performance on the test. This highlights the need for a thorough evaluation of the curriculum and teaching methods used to ensure that all students receive adequate preparation. Moreover, the low pass rate could indicate that the TOAFL test may be too difficult or not fully aligned with the skills taught in the academic programs. If the test does not correspond with the curriculum, students may struggle to meet the expected standards despite having completed relevant courses. This underscores the necessity of reviewing the alignment between the test and the curriculum and considering adjustments to the test's difficulty level to better reflect the skills students are expected to master. It was expressed in the interview as follows: *"We were never taught the material covered in the test because during our Arabic language classes, the lecturer only taught basic conversational Arabic for 16 sessions."* (Nurhidaya Putri, syaria banking program, interview translated, 7/5/2024). *"My Arabic skills are limited to speaking simple sentences, so when I took the TOAFL test, I didn't know how to answer the questions. So I had to take the test three times"*. (Muh. Amir, mathematics program, interview translated, 10/5/2024). *"The TOAFL test is quite difficult for me. Even though I can understand the reading passages provided, the answer choices are almost identical and very misleading."* (Muhammad Zainul Rusdi, Arabic Language Education Program, interview translated, 10/5/2024).

The analysis of TOAFL test results at IAIN Parepare reveals challenges in achieving standardization and effectiveness in Arabic proficiency testing. The low pass rate points to variability in both test preparation and administration, as well as possible

misalignment between the test and the curriculum. Therefore, improvements are needed to enhance the quality of teaching and to ensure that Arabic proficiency tests more accurately and fairly assess students' abilities.

The inability of students to achieve the expected results in the TOAFL test can be attributed to several factors affecting their capacity to comprehend and respond to the test questions. First, the lack of Arabic vocabulary proficiency presents a significant challenge for students, making it difficult for them to understand the questions. As noted by Asadi, limited vocabulary knowledge significantly hampers students' ability to interpret test questions accurately (Albirini, 2016). This deficiency in vocabulary knowledge prevents students from fully understanding the questions, leading to errors in their responses. Second, students face difficulties in translating and comprehending the reading passages (*qira'ah*) included in the test. Before answering the questions, students must first understand the given sentences or texts, but many struggle even to read them accurately, let alone translate them (Bohra, 2020). This highlights a gap in reading and translation skills, which directly affects test performance.

Additionally, weaknesses in mastering Arabic grammar, including *nahwu* (syntax) and *sarf* (morphology), create significant obstacles for students in constructing and comprehending sentences. As noted by Mofareh, a solid understanding of grammar is critical for language proficiency, and without adequate mastery, students are unable to respond correctly to the test items (ALQAHTANI, 2015). Even though Arabic is a compulsory subject in Islamic higher education, the limited credit hours and restricted learning time are insufficient to develop the level of proficiency required to pass the TOAFL test. Moreover, the multiple-choice format of the TOAFL test provides an opportunity for students to guess answers, meaning the scores do not always reflect actual proficiency. Finally, the lack of available preparatory materials and information about the test format often leaves students underprepared, as they are unaware of the types of questions and strategies required for success. This lack of preparation significantly affects their performance in the test.

The Impact of Arabic Proficiency Testing Policy on Language Skills and Academic Readiness of Students

The government policy regarding Arabic proficiency testing at IAIN Parepare has not significantly improved students' Arabic language skills. The low pass rate indicates that many students still struggle with the vocabulary and grammar necessary for proficiency. Data from IAIN Parepare reveal that only 36% of the 659 students passed the first stage of the TOAFL test, reflecting the ineffectiveness of this policy in enhancing Arabic competency. This suggests that the implemented policies have not achieved the expected outcomes in terms of improving language skills.

Furthermore, the TOAFL testing policy has not adequately enhanced students' academic readiness. Students who fail this test often face difficulties in attending classes conducted in Arabic, which diminishes their overall academic preparedness. Interviews with students who did not pass the exam indicate that they encounter challenges in understanding academic material due to their limited proficiency in reading (*qira'ah*) and grammar. Thus, this policy has yet to provide a significant positive impact on students' academic readiness. Moreover, the Arabic proficiency testing policy does not fully support the improvement of students' academic performance. Failing the TOAFL test serves as a significant barrier to completing their studies, forcing students to retake the

exam multiple times. Data show that students who do not pass must repeat the test, which consumes time and hinders their graduation, negatively affecting their academic achievements. This indicates that the policy is impeding academic progress rather than promoting improved performance.

Additionally, this policy appears to contribute to a decline in students' motivation to learn. The mismatch between the material taught in class and the TOAFL test generates frustration, which reduces their motivation to study Arabic. One student expressed, "What is the point of learning Arabic if the material taught in class differs from the test given? Furthermore, our Arabic skills are still very basic." This statement reflects the need for the policy to be adjusted to align more closely with classroom learning processes and to enhance students' motivation.

Ultimately, this policy serves as an obstacle for students preparing for further studies and careers. Students who do not pass the TOAFL test feel inadequately prepared to pursue higher education or enter the workforce, especially given the minimum TOAFL score requirement of 450 for the Ministry of Religious Affairs scholarships. Those who fail this test believe they lack the competencies necessary for advancing to higher educational levels or competing in the job market. This highlights that the policy has not been effective in helping students achieve their academic and professional goals, necessitating evaluation and adjustment of the policy to yield more positive results in the future.

The implementation of Arabic language policies in State Islamic Religious Universities (PTKIN) faces significant challenges, primarily due to inadequate Arabic language instruction at these institutions. Many PTKIN have yet to provide sufficient attention to the development of Arabic language curricula, resulting in students being less skilled in communicating effectively in the language. The lack of focus on both theoretical and practical aspects of Arabic language learning makes it difficult to achieve the government's policy objectives—that students should be able to communicate proficiently in Arabic. This situation is exacerbated by the implementation of policies at the university level, including a shortage of qualified instructors and supportive facilities, such as language laboratories or additional programs.

As a result of insufficient coursework, students often struggle to prepare for Arabic proficiency tests, such as the TOAFL. This test assesses practical skills that are not adequately taught in the classroom, leaving many students unprepared to face the challenges posed by the exam. Students who do not receive adequate Arabic language instruction often lack confidence and frequently fail to achieve the expected scores, ultimately hindering their academic and professional development. These challenges emphasize that without improvements in the approach to Arabic language learning at PTKIN, existing policies will not be able to yield the desired outcomes.

Furthermore, the challenges in Arabic language learning also affect student motivation. Students who feel inadequately prepared for proficiency tests tend to lose interest and enthusiasm for their studies. The disconnect between policy expectations and the reality on the ground creates dysfunction within the educational system at PTKIN, where students are unable to develop language skills that meet the expected standards. Therefore, significant efforts are required to enhance the quality of Arabic language education in PTKIN to ensure that policy objectives can be met and students are better equipped to communicate in Arabic in academic and professional environments. This

finding aligns with previous research, which indicates that inadequate preparation and support significantly impact students' language proficiency and motivation (Ngwa, 2020). The gap between government policies and their implementation at State Islamic Religious Universities (PTKIN) poses a serious issue that hinders the achievement of desired outcomes in Arabic proficiency testing. Policies aimed at improving students' Arabic language skills are often not effectively translated into the curriculum and teaching methods applied in the classroom. The curriculum tends to focus more on theoretical aspects, while proficiency tests such as the TOAFL assess practical skills that require deeper language mastery. This misalignment results in many students being unprepared for the tests, as reflected in the low pass rates.

The failure to align the curriculum with testing standards leads many students to feel that what they learn in class is not relevant to the material being tested. This creates confusion and diminishes students' confidence, ultimately negatively impacting their motivation to learn. Students who feel inadequately prepared for proficiency tests often struggle to achieve the expected results, affecting their overall academic performance. This gap reinforces the argument that without appropriate curriculum adjustments, government policies will continue to face challenges in achieving their intended objectives.

In addition to affecting test outcomes, this gap also impacts the quality of graduates from PTKIN. Students who do not receive relevant instruction related to Arabic proficiency tests tend to be less competent in using the language in academic and professional environments. This creates long-term issues, where PTKIN graduates may not be prepared to face challenges that require Arabic language proficiency, both in further studies and in their careers. Therefore, to address this gap, a comprehensive revision of the Arabic language curriculum and teaching methods at PTKIN is necessary to align more closely with the standards being assessed and the needs of students. This finding is consistent with previous research that highlights the critical need for curriculum reform to enhance student readiness and proficiency (Wiley, 2016).

Adjustments to policies are essential for improving Arabic language skills among students at State Islamic Religious Universities (PTKIN), given the numerous challenges and gaps that have been identified. One key strategy is to revise the curriculum to align more closely with the standards assessed in proficiency tests such as the TOAFL. A curriculum designed with a focus on mastering practical Arabic language skills will better prepare students for these tests, which in turn will enhance their pass rates and overall learning outcomes. Several PTKIN that have made curriculum changes have shown significant improvements in TOAFL test results, indicating that these adjustments have a positive impact.

In addition to curriculum adjustments, enhancing training for instructors is also a crucial strategy for improving students' Arabic language skills. Well-trained instructors who utilize practical teaching methods relevant to the test standards can greatly contribute to achieving policy goals. Continuous faculty training will ensure that instructors possess the necessary skills and knowledge to teach Arabic effectively, allowing students to develop competencies that meet their academic and professional needs. Studies have shown that instructors who receive specialized training can improve their teaching quality, ultimately enhancing students' Arabic language skills.

Another strategy to consider is the development of innovative and more practical teaching methods that can be implemented in Arabic language learning at PTKIN. The

use of learning technologies, conversation simulations, and other supportive programs can help engage students more actively in the learning process and improve their skills. Innovations in teaching will make Arabic language learning more engaging and relevant to the needs of proficiency tests, thereby reducing the gap between what is taught in the classroom and what is assessed in the tests. By implementing these strategies, PTKIN can enhance the effectiveness of Arabic language policies and ensure that students graduate with adequate language skills to succeed in academic and professional environments. This finding is consistent with previous research that emphasizes the significance of curriculum innovation and effective teacher training in enhancing student outcomes in language education (Rabe-Hemp, Woollen, & Humiston, 2009).

CONCLUSION

The findings from interviews, test results, and curriculum analysis highlight both the challenges and opportunities presented by Arabic proficiency testing at Indonesian Islamic State Universities. Data revealed that while the testing policy aims to enhance students' Arabic skills, significant gaps exist, including misalignment between curriculum content and testing standards, which leads to inadequate student preparation and low pass rates. Additionally, students expressed feelings of frustration and lack of confidence due to the disconnect between classroom instruction and the demands of the TOAFL test. Addressing these issues is crucial for improving academic readiness and language proficiency. Recommendations for future action include revising the Arabic curriculum to better align with proficiency standards, enhancing instructor training programs, and establishing a standardized Arabic proficiency testing system to ensure uniformity and fairness across institutions. This study is limited in its scope, focusing primarily on a single institution, which may affect the generalizability of the findings. Future research should explore multiple institutions to provide a broader understanding of how policy implementation impacts Arabic proficiency across diverse educational contexts. Furthermore, further investigation into effective pedagogical strategies for teaching Arabic in relation to proficiency testing is warranted to ensure that students are adequately prepared for both academic and professional challenges.

REFERENCES

- Abdel Latif, M. M. M., & Alhamad, M. M. (2023). Arabicization or Englishization of higher education in the Arab world? Controversies, policies and realities. *Frontiers in Psychology*, 14(February), 1–8. <https://doi.org/10.3389/fpsyg.2023.1093488>
- Albirini, A. (2016). *Modern Arabic sociolinguistics: Diglossia, variation, codeswitching, attitudes and identity*. Routledge. <https://doi.org/10.4324/9781315683737>
- Alowaydhi, W. H. (2016). Evaluation of Arabic Language Learning Program for Non-Native Speakers in Saudi Electronic University According to Total Quality Standards. *Journal of Education and Practice*, 7(32), 74–90.
- ALQAHTANI, M. (2015). The importance of vocabulary in language learning and how to be taught. *International Journal of Teaching and Education*, III(3), 21–34. <https://doi.org/10.20472/TE.2015.3.3.002>
- Bohra, M. (2020). *Non-Linguistic Factors Affecting Students' Acquisition of Arabic Speaking Skills in Al Jamea Tus Saifiyah University Nairobi* Mustafa Bohra. 25(11), 21–32. Retrieved from <https://www.iosrjournals.org/iosr->

- jhss/papers/Vol.25-Issue11/Series-6/D2511062132.pdf.
- Bui, T. (2022). A Review of Language Testing and Assessment in Online Teaching. *International Journal of English Linguistics*, 12, 54. <https://doi.org/10.5539/ijel.v12n4p54>
- Cheng, J. (2009). Automatic Assessment of Spoken Modern Standard Arabic. *Proceedings of the 4th Workshop on Innovative Use of NLP for Building Educational Applications, BEA 2009*, 1–9. Retrieved from <https://www.scopus.com/inward/record.uri?partnerID=HzOxMe3b%5C&scid=84922836646%5C&origin=inward>. <https://doi.org/10.3115/1609843.1609844>
- Gay, G. H. E., & Betts, K. (2020). From discussion forums to e-meetings: Integrating high touch strategies to increase student engagement, academic performance, and retention in large online courses. *Online Learning Journal*, 24(1), 92–117. <https://doi.org/10.24059/olj.v24i1.1984>
- Gestanti, R., Mufanti, R., & Nimasari, E. (2019). *Constructing Test Specification for Standardized Test of Academic English Proficiency (STAcEP)*. <https://doi.org/10.4108/eai.27-4-2019.2285347>
- Greene, J. C. (1989). Qualitative Data Analysis and Interpretation. *Creative Ideas For Teaching Evaluation*, 147–153. https://doi.org/10.1007/978-94-015-7829-5_17
- Hefner, R. W. (2009). Islamic schools , social movements, and democracy in Indonesia. *Making Modern Muslims*, 55–105. Retrieved from <https://www.scopus.com/inward/record.uri?partnerID=HzOxMe3b%5C&scid=84872713643%5C&origin=inward>. <https://doi.org/10.21313/hawaii/9780824832803.003.0002>
- Indriana, D., & Maksum, A. (2023). Implementation of The Special Arabic Language Learning Program (PKPBA) for New Student at The Tarbiyah and Teacher Training Faculty UIN Sultan Maulana Hasanuddin Banten| Implementasi Program Khusus Perkuliahan Bahasa Arab (PKPBA) Bagi Mahasiswa Baru Fakul. *Mantiqu Tayr: Journal of Arabic Language*, 3(1), 53–69. <https://doi.org/10.25217/mantiqu tayr.v3i1.3173>
- Jong, N. H. De. (2018). Fluency in Second Language Testing: Insights From Different Disciplines. *Language Assessment Quarterly*, 15(3), 237–254. <https://doi.org/10.1080/15434303.2018.1477780>
- Liddicoat, A. J. (2013). Language-in-education Policies: The Discursive Construction of Intercultural Relations. *Language-in-Education Policies: The Discursive Construction of Intercultural Relations*, 1–246. Retrieved from <https://www.scopus.com/inward/record.uri?partnerID=HzOxMe3b%5C&scid=84948983985%5C&origin=inward>. <https://doi.org/10.21832/9781847699152>
- May, L. (2011). Interactional Competence in a Paired Speaking Test: Features Salient to Raters. *Language Assessment Quarterly*, 8, 127–145. <https://doi.org/10.1080/15434303.2011.565845>
- McLelland, N. (2021). Language standards, standardisation and standard ideologies in multilingual contexts. *Journal of Multilingual and Multicultural Development*, 42(2), 109–124. <https://doi.org/10.1080/01434632.2019.1708918>

- Miller, W. G. (2018). The standardization journey and the path ahead. *Journal of Laboratory and Precision Medicine; Vol 3 (October 2018): Journal of Laboratory and Precision Medicine*. Retrieved from <https://jlpn.amegroups.org/article/view/4544>.
<https://doi.org/10.21037/jlpn.2018.09.14>
- Mislevy, R. J., Carlson, J., Bridgeman, B., Golub-smith, M., & Stone, E. (2017). *Psychometric Models for Reading*. (December).
- Mohamed, S. (2023). The development of an Arabic curriculum framework based on a compilation of salient features from CEFR level descriptors. *Language Learning Journal*, 51(1), 33–47.
<https://doi.org/10.1080/09571736.2021.1923781>
- Muradi, A., Amaluddin, L. O., Hilmi, D., Hery, B., Noor, F., Noor, H., & Jailani, M. S. (2022). Evaluation of Arabic Learning Program at State Islamic Religious College During COVID-19. *Eurasian Journal of Educational Research*, 2022(99), 189–200. Retrieved from <https://idr.uin-antasari.ac.id/21241/>
- Musthofa, T. (2022). CEFR-Based Policy in Arabic Language Teaching and Cultural Dimension in Indonesian Islamic Higher Education. *Eurasian Journal of Applied Linguistics*, 8(2), 96–107. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1374599.pdf>.
<https://doi.org/10.32601/ejal.911545>
- Ngwa, C. N. (2020). Cameroon english speech: Intelligibility to some chinese speakers of english living in cameroon and pedagogic concerns. *Journal of Language Teaching and Research*, 11(2), 177–183.
<https://doi.org/10.17507/jltr.1102.05>
- Niitemaa, M. L., & Pietilä, P. (2018). Vocabulary skills and online dictionaries: A study on efl learners' receptive vocabulary knowledge and success in searching electronic sources for information. *Journal of Language Teaching and Research*, 9(3), 453–462. <https://doi.org/10.17507/jltr.0903.02>
- Norris, J. (2018). Language Testing and Evaluation: Developing C-tests for estimating proficiency in foreign language research. *Language Testing and Evaluation: Developing C-Tests for Estimating Proficiency in Foreign Language Research*, 1–380. <https://doi.org/10.3726/b13235>
- Pill, J., & Harding, L. (2013). Defining the language assessment literacy gap: Evidence from a parliamentary inquiry. *Language Testing*, 30, 381–402.
<https://doi.org/10.1177/0265532213480337>
- Posner, M. A. (2011). The impact of a proficiency-based assessment and reassessment of learning outcomes system on student achievement and attitudes. *Statistics Education Research Journal*, 10(1), 3–14.
<https://doi.org/10.52041/serj.v10i1.352>
- Rabe-Hemp, C., Woollen, S., & Humiston, G. S. (2009). A comparative analysis of student engagement, learning, and satisfaction in lecture hall and online learning settings. *Quarterly Review of Distance Education*, 10, 207–218.
- Raish, M. (2021). *Issues in Arabic Language Testing and Assessment (Chapter 4) - The Cambridge Handbook of Arabic Linguistics*. Publisher: Cambridge University Press. <https://doi.org/10.1017/9781108277327.005>
- Ren, Y. (2014). A study of the washback effects of the College English Test (band 4) on teaching and learning English at tertiary level in China. *International*

- Journal of Pedagogies and Learning*, 6, 243–259.
<https://doi.org/10.5172/ijpl.2011.6.3.243>
- Ritonga, M. (2020). Analysis of Arabic Language Learning at Higher Education Institutions with Multi-Religion Students. *Universal Journal of Educational Research*, 8(9), 4333–4339. <https://doi.org/10.13189/ujer.2020.080960>
- Ruegg, R., Hoang, H., & Petersen, N. (2024). The influence of English language proficiency test scores on the academic success of ESL undergraduate students. *Educational Research and Evaluation*, 29(3–4), 111–129. <https://doi.org/10.1080/13803611.2024.2314533>
- Sánchez, M. T. (. (2018). Reframing language allocation policy in dual language bilingual education. *Bilingual Research Journal*, 41(1), 37–51. <https://doi.org/10.1080/15235882.2017.1405098>
- Schissel, J. L. (2019). The Construct of Multilingualism in Language Testing. *Language Assessment Quarterly*, 16(4), 373–378. <https://doi.org/10.1080/15434303.2019.1680679>
- Seleksi Timur Tengah - Kemenag.* (n.d.). Retrieved from <http://diktis.kemenag.go.id/timteng12/022/>
- Spolsky, B. (2019). A modified and enriched theory of language policy (and management). *Language Policy*, 18(3), 323–338. <https://doi.org/10.1007/s10993-018-9489-z>
- Wiley, T. G. (2016). Language Policy and Planning in Language Education: Legacies, Consequences, and Possibilities. *Modern Language Journal*, 100, 48–63. <https://doi.org/10.1111/modl.12303>
- Yusri, G. (2013). Cognitive and metacognitive learning strategies among Arabic language students. *Interactive Learning Environments*, 21(3), 290–300. <https://doi.org/10.1080/10494820.2011.555840>