Correlation between Intensive Training Program (ITP) and English and Arabic Proficiency Test (EAPT) results: An analysis of three language skills of postgraduate students

[1]**Ana Mariana***, [2]**Ibnu Rawandhy N. Hula**, [1]**Dahlia Husain**, [3]**Nurul Aini Pakaya**
[1]*English Literature, Faculty of Social Science, Universitas Muhammadiyah Gorontalo, Indonesia*
[2]*Arabic Language Education, Faculty of Education and Teacher Training, IAIN Sultan Amai Gorontalo, Indonesia*
[3]*Arabic Literature, Faculty of Social Science, Universitas Muhammadiyah Gorontalo, Indonesia*

*CorrespondingAuthor
Email:anamariana@umgo.ac.id

Abstract

This study measures the correlation between the Intensive Training Program (ITP) and the English and Arabic Proficiency Test (EAPT) scores among 100 postgraduate students at IAIN Sultan Amai Gorontalo from four non-language programs. A quantitative method with an ex-post facto design was used, and data analysis was conducted using Spearman's correlation test. Data were collected through a three-aspect proficiency test: listening (istima'), structure (tarakib), and reading (qira'ah), using a Computer Assisted Test (CAT) application with 140 items. The test results were processed using score conversion and SPSS Version 30. Before the ITP, none of the students achieved the minimum passing score for the EPT (450), and only six passed the APT. After the training, 18% reached the minimum EPT score, and the number of students who passed the APT increased significantly to 32. Correlation analysis showed that ITP effectively improved listening (correlation 0.582; p=0.002) and structure (correlation 0.592; p=0.002) in EPT, but was less effective for reading (correlation 0.085; p=0.686). In contrast, ITP greatly enhanced all aspects of APT: istima' (correlation 0.972; p=0.000), tarakib (correlation 0.993; p=0.000), and qira'ah (correlation 0.997; p=0.000). The short duration of the training affects its effectiveness, along with students' educational backgrounds and learning motivation. To enhance efficacy, it is recommended to extend the training duration, provide relevant modules for non-language students, enhance motivation through regular practice, and improve reading instruction methods. ITP positively impacts Arabic and English language proficiency, particularly for APT, suggesting that targeted training can significantly benefit non-language students' language skills.

***Keywords:** Intensive Training Program (ITP); English; Arabic Proficiency Test; three language skills*

INTRODUCTION

Foreign language learning, such as Arabic and English, plays an essential role in improving the global communication skills of graduate students. Proficiency in these two languages is not only an indicator of academic quality but also a prerequisite for facing challenges in the era of globalization (Rocha, 2025), (Aljarelah, 2024). The importance of foreign language learning is reflected in the requirement for students to achieve a minimum score in the required language test, which serves as a benchmark for their comprehensive language proficiency(El Meneza, 2020). Success in this test not only demonstrates foreign language proficiency but also facilitates the completion of their studies and career preparation and opens up wider opportunities to compete in the world of work (Alhamami & Almelhi, 2021)after they obtain their master's degree at IAIN Sultan Amai Gorontalo.

As one of the Islamic higher education institutions, the State Islamic Institute (IAIN) Sultan Amai Gorontalo recognizes the importance of Arabic and English language skills for graduate students. To enhance students' linguistic competence prior to the English and Arabic Proficiency Test (EAPT), the institution implements an Intensive Training Program (ITP) (El-Dakhs et al., 2023). This program is designed to equip students with foundational language skills and strategic approaches to tackle EAPT items administered through Computer-Assisted Testing (CAT). Although officially categorized as optional under the Language Development Center's Standard Operating Procedures (SOP), the ITP functions as a de facto mandatory requirement and is conducted regularly to ensure student readiness.

The necessity of such training is underscored by evidence indicating that postgraduate students in similar Islamic higher education contexts often exhibit low proficiency in both spoken and written international languages (Shohamy, 2005). emphasizes that language testing in educational settings must account for learners' diverse linguistic backgrounds and prior exposure, particularly in multilingual and non-native English-speaking environments. This insight is highly relevant to IAIN Sultan Amai Gorontalo, where many graduate students come from non-language disciplines and have limited experience with standardized language assessments such as TOEFL, TOAFL, or IELTS (Boy Jon et al., 2021). Consequently, it is rare to find students who demonstrate standard proficiency in both Arabic and English, highlighting the critical role of preparatory programs like the Intensive Training Program (ITP). Despite its regular implementation, EAPT results show significant variation—some students achieve high scores, while others struggle to meet the minimum standards and must repeat remedial courses. This disparity raises questions about the effectiveness of the ITP in equipping students for bilingual proficiency assessment.

Although intensive training programs have been conducted

regularly, the EATP results of postgraduate students at IAIN Sultan Amai Gorontalo show significant variation. Some students achieved high scores, but others struggled to meet the minimum standards set and had to repeat remedial courses several times. This condition raises questions about the effectiveness of the ITP program in helping students achieve maximum results on the EATP.

Previous studies have explored the effect of language training on students' language abilities. For example, a study by Dobrić et al., (2021) found that intensive English training significantly improves speaking and writing skills. In the context of Arabic language learning, Nufus, Erlina, Koderi et al., (2022) reported that intensive training is more effective when integrated with contextual learning methods, a finding supported by Awad Al-Rohili (2024).

However, most of these studies focus on only one language or evaluate only certain aspects of language proficiency, such as grammar or vocabulary. Other studies mention that intensive language training can positively impact students' language proficiency. For example, a study by Abdulrahman & Kara, (2023) found that intensive English training programs could improve listening and speaking skills by up to 30% in a short period (Short, 2013). Meanwhile, Jendli & Albarakati, (2024) reported that Arabic language training using contextual methods helped students understand language structure and vocabulary more effectively. However, most of these studies focused on only one language or a specific aspect of language proficiency, such as grammar or vocabulary, without combining a holistic approach that encompasses both languages simultaneously.

In addition, literature on the effectiveness of intensive combined training for Arabic and English is minimal, especially in the context of standardized tests such as the English and Arabic Proficiency Test (EAPT). This indicates the need for further research to identify students' bilingual abilities, which is a gap in research on the relationship between intensive combined training (Arabic and English) and comprehensive language proficiency test results, which is still very limited.

This gap forms the basis for this study, which aims to analyze the impact of the Intensive Training Program (ITP) on the results of the English and Arabic Proficiency Test (EAPT) for postgraduate students at IAIN Sultan Amai Gorontalo. This study is highly relevant because most previous studies have focused on one language, either Arabic or English, without considering the interaction between the two in the context of intensive training. Uniquely, this study combines both languages in one analysis, allowing for a more comprehensive understanding of how simultaneous training can affect students' bilingual abilities. With this holistic approach, this study provides insights into the effectiveness of intensive training programs and fills gaps in the

existing academic literature.The novelty of this research lies in the local context of graduate students at IAIN Sultan Amai Gorontalo, who are required to pass the EAPT test as a condition for completing their studies. By considering the specific challenges students face in mastering two languages simultaneously, this research is expected to significantly contribute to the development of more effective language training programs at the institution.

The significance of this research from a scientific perspective is to fill the gap in the literature on intensive language training (ITP) involving two languages simultaneously. From a practical perspective, the results of this study can be used by educational institutions to design more effective and relevant language training programs to meet the needs of postgraduate students in improving their language proficiency. Additionally, it is unknown whether there is a correlation between the implementation of ITP and the EAPT scores obtained by postgraduate students at IAIN Sultan Amai Gorontalo. Therefore, this correlation is considered essential to determine for evaluating and providing feedback on improvements in several aspects, such as 1) the training program structure, 2) training materials, and duration, 3) the difficulty level of EAPT questions, 4) the difficulty level of the questions, and 5) the CAT system currently in use.

The primary objective of this study is to evaluate the extent to which the ITP contributes to the achievement of EATP scores among postgraduate students, supported by data on (1) English and Arabic Proficiency Test (EAPT) scores obtained by students before and after participating in the ITP, and (2) the correlation between the Intensive Training Program (ITP) in English and Arabic and EAPT scores. Additionally, this study aims to identify other factors that may influence EATP results, such as learning motivation, educational background, and training duration. The findings of this study are expected to provide practical recommendations for improving the quality of language training programs at IAIN Sultan Amai Gorontalo, thereby supporting students' success in language proficiency exams.

METHOD

This study used a quantitative approach with an ex post facto design(Em, 2024). This design helps test the causal relationship between independent and dependent variables that have occurred naturally, so researchers cannot control variables that may influence the results(Remler & Van Ryzin, 2024). In this design, the researcher analyzes existing data to find correlational relationships, so that the relationship between two variables (1) Intensive Training Program and (2) EAPT (English and Arabic Proficiency Test) results based on CAT (Computerized Adaptive Testing) can be identified.

The research population consists of 240 students. The study

employs a stratified sample, comprising 41.66% of the population, or 100 students, as outlined in the following table.

Table 1. Population, Sample, and Percentage

No	Department	Populations	Samples	Persentage
1	Islamic Religious Education	64	26	10,83%
2	Family Law	100	42	17,50%
3	Sharia Economics	43	18	7,50%
4	Islamic Education Management	33	14	5,83%
	Total	240	100	41,66%

The research was conducted at the Language Development Center Laboratory of IAIN Sultan Amai Gorontalo. The period from training to testing took 27 JPL, or 1,620 minutes. Data collection techniques used include (1) intensive training document analysis and (2) results from the English and Arabic Proficiency Test (EAPT). The intensive training documents consist of: pre-test scores, post-training scores, and final EAPT scores that have been converted. The minimum conversion score is 450 as the passing requirement, as referenced in the table below.

Table 2. EAPT scores, number of correct answers, and score conversion

No	Predicate	Score EAPT	Number of Correct Answers	Conversion		
				L/I	S/T	R/Q
1	Elementary	310-420	0-18	24-43	20-42	21-38
2	Low Intermediate	420-480	19-28	44-49	44-52	39-46
3	High Intermediate	480-520	29-40	50-57	53-68	47-55
4	Advanced	525-677	41-50	58-68		56-67

Note Conversion: L/I: Listening/istima' S/T: Structure/Tarakib, R/Q: Reading/Qira'ah

The scoring rubric and score conversion for the EAPT were developed by the Language Development Center team in collaboration with language assessment experts, based on a criterion-referenced framework. The conversion scale was adapted from the TOEFL ITP scoring model (Educational Testing Service [ETS] (Kubiszyn & Borich, 2024), which has been widely used in institutional language assessments, and further adjusted through pilot testing and item analysis to align with the institutional proficiency standards of IAIN Sultan Amai Gorontalo. This adaptation is consistent with practices in similar language testing contexts (Lee, 2022). The standardization process aimed to ensure the reliability and validity of test interpretation, particularly in mapping the number of correct answers to proficiency levels (Elementary to Advanced). Listening/Istima', Structure/Tarakib, and Reading/Qira'ah sections were weighted and converted accordingly to produce the final EAPT score.

In this study, the researchers conducted validity and reliability tests on the instruments to ensure the accuracy and consistency of the measurement results. The validity of the instruments was tested through content analysis involving three experts in the field of language assessment and testing. The validators were selected based on the following criteria: (1) hold a doctoral degree (Ph.D.) in applied linguistics or language education, (2) have at least five years of experience in developing language proficiency tests, and (3) are actively involved in language program development at the institutional level.

In addition, construct validity can be established by analyzing the

relationship between test scores and other relevant variables, such as test results from intensive training programs. By ensuring the validity and reliability of the instruments, this study can provide more credible and reliable results in identifying the relationship between the Intensive Training Program (ITP) and the EAPT test results. Reliability is measured to test internal consistency through Cronbach's alpha coefficient analysis (Izah et al., 2024), which is formulated as:

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k}\sigma_{X_i}^2}{\sigma_X^2}\right)$$

k Is the number of item, $\sigma_{X_i}^2$ Is the variance of each item, and σ_X^2 Is the total variance of the overall scores. The reliability test in this study resulted in a Cronbach's alpha value of 0.86, which exceeds the commonly accepted threshold of 0.70. This indicates that the instrument has a high level of internal consistency, meaning that the items used are measuring the same underlying construct in a consistent manner. Therefore, the instrument can be considered reliable for use in assessing the relationship between ITP and EAPT test results.

Data analysis techniques use descriptive statistical techniques. This analysis describes data characteristics and is used to test hypotheses and perform Spearman's correlation test(Yu & Hutson, 2024). The formula used to calculate Spearman's correlation coefficient is:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2 - 1)}$$

d_iIs the difference in rank, and $n$is the number of data pairs. Data processing was performed using SPSS software. The significance level used in this study was 5% (0.05). The results of the validity test showed that all items had a correlation coefficient (r_{count}) greater than the critical value 0. 2787 (r_{table}), indicating that each item was significantly correlated with the total score. This demonstrates that the instrument items effectively measure the intended construct, and thus, the instrument can be considered valid for assessing the relationship between ITP and EAPT test results.

RESULTS AND DISCUSSION

This section presents data related to the research objectives. The data consists of (1) English and Arabic Proficiency Test (EAPT) scores of students before participating in the Intensive Training Program (ITP), (2) English and Arabic Proficiency Test (EAPT) scores of students after participating in the ITP, and (3) the correlation between the Intensive Training Program (ITP) in English and Arabic and the EAPT scores:

Scores obtained before the Intensive Training Program (ITP)

Researchers first conducted a pretest for all students to determine the initial abilities of students who will participate in ITP. A total of 100 students registered for the training. The following are the EPT and APT

pretest scores:

**Table 3:** Recapitulation of Initial EPT Scores before ITP

Criterion	Listening	Structure	Reading	Score EPT
Average	41,62	37,87	33,71	377
Standard Deviation	4,39	2,38	7,09	33,75
Highest Score	49	50	53	450
Lowest Score	32	33	28	320
Number of Passes	2 Students			

Table 3 above shows that the EPT scores are still far from the minimum EPT score of 450. This can be seen from the average score of 377.34 and standard deviation of 33.75. In addition, only two students passed with scores of 450.
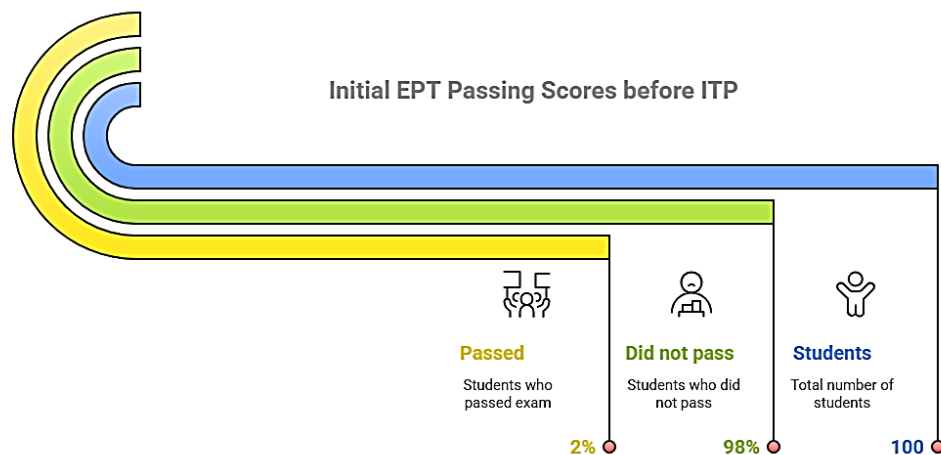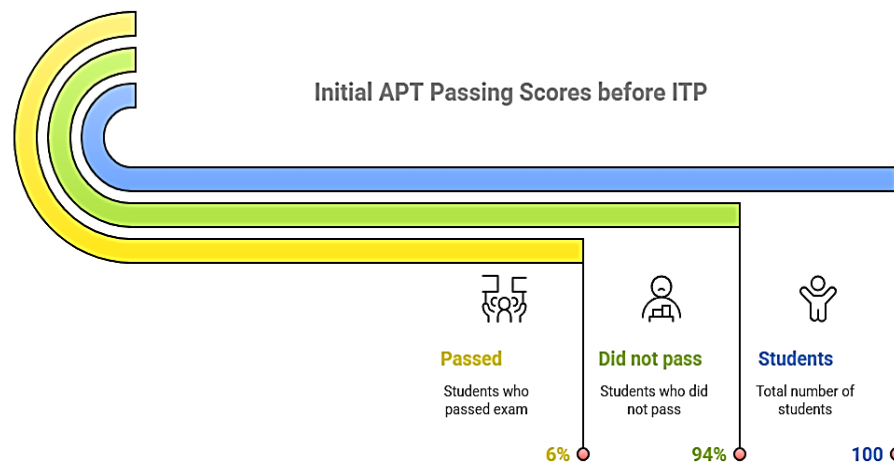

Figure 1. Initial EPT Passing Scores before ITP

Table 4: Summary of Initial APT Scores before ITP

Criterion	Istima'	Tarakib	Qira'ah	Score APT
Average	42,17	38,16	35,5	386
Standard Deviation	4,28	3,37	8,15	39,09
Highest Score	54	52	54	523
Lowest Score	32	29	28	320
Number of Passes	6 Students			

Table 4 above shows that the APT scores are still low, but there are students who were able to achieve scores above the standard of 450. Six students achieved the highest scores of 523, 487, 480, 477, 473, and 463. The lowest score was only 320.

Figure 2. Initial APT Passing Scores before ITP

Based on the results above, the researchers obtained initial data that out of 100 students who took the pretest, two students achieved a minimum score of 450 in the EPT, while six students achieved the minimum passing score in the APT. Based on this, the researcher conducted an Intensive Training Program (ITP) specifically designed to improve students' proficiency and skills in three areas: (1) Listening/Istima', (2) Structure/Tarakib, and (3) Reading/Qira'ah. The training was conducted over 2 days, each requiring 7 JPL.

The objective of the ITP is not only to provide students with foundational knowledge about the three language proficiency sessions required for the CAT-based EAPT test but also to offer tips, tricks, question formats, techniques, and strategies for tackling the EAPT, accompanied by practice exercises and mock tests. This approach ensures students achieve the required minimum score and avoid repeated failures.
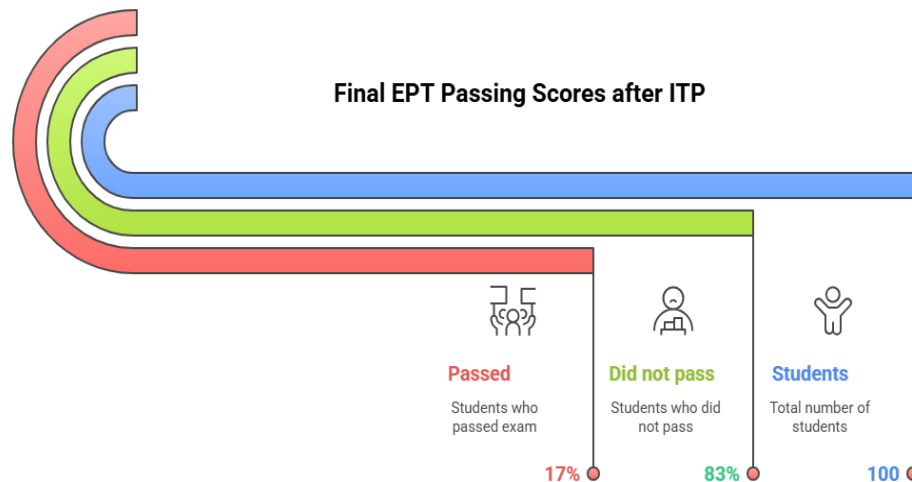
Final EPT score after Intensive Training Program (ITP)

Table 5. Final EPT scores after the Intensive Training Program (ITP)

| Criterion	Listening	Structure	Reading	Score EPT
Average	41,62	38,79	34,11	382
Standard Deviation	4.11	4,22	7,52	42,04
Highest Score	48	56	54	480
Lowest Score	32	33	28	320
Number of Passes		17 Students		

Figure 4. Final EPT Passing Scores after ITP

Table 5 and figure 4 illustrate a notable improvement on the achievement of a minimum score of 450 following students participation in the Intensive Training Program (ITP). The average score increases to 382, along with additional 15 students meeting the minimum standard. In total, 17 students achieved the target score, whereas 83 students remained below the standard. Compared to the initial scores before participating in the ITP, there was a significant increase of 15%.

Table 6. Final APT scores after the Intensive Training Program (ITP)

Criterion	Istima'	Tarakib	Qira'ah	Score APT
Average	44,31	41,08	39,26	416
Standard Deviation	5,63	6,65	9,19	60,77
Highest Score	58	56	54	533
Lowest Score	32	29	28	320
Number of Passes	32 Students			
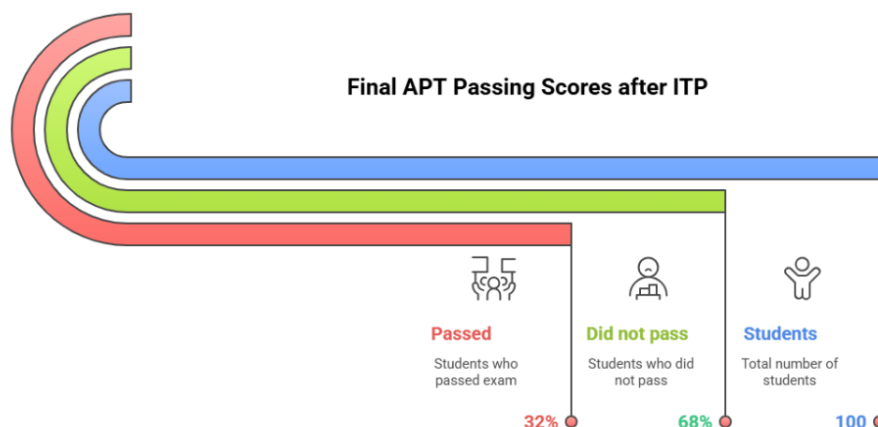
Figure 5. Final APT Passing Scores after ITP

Table 6 and Figure 5 indicate a marked improvement in achieving the minimum score of 450 following students' participation in the Intensive

Training Program (ITP). The average score rose to 516, with 32 students attaining the minimum requirement, while 68 students remained below 450. Compared to the baseline scores prior to the ITP, this reflects a substantial increase of 26%.

Analysis of Results between Intensive Training Program and EAPT Scores

Table 7. EAPT Difference Test after Intensive Training Program (ITP)

Test the Difference in Pretest and Posttest Results			
Proficiency Test	N	Test Normal[c]	Sig.
EPT_Pre & EPT_Post	100	Yes	0,000[a]
APT_Pre & APT_Post	100	No	0,000[b]

[a]Paired Sample t-Test
[b]Wilcoxon Signed Rank Test
[c]Kolmogorov-Smirnov Test

Table 7 above shows the difference test results between the pretest and post-test scores on the EPT and APT, each involving 100 participants. The Kolmogorov-Smirnov test was used to determine the appropriate type of statistical test. The results of the normality test showed that the EPT data (pre and post) were normally distributed, so the Paired Sample t-test was used. Conversely, the APT data were not normally distributed, so the Wilcoxon Signed Rank Test was used. The results of both tests showed a sign.427ificance value of 0.000 (p < 0.05), with the 95% confidence interval confirming the statistical significance of the differences observed, indicating a significant difference between the pretest and posttest scores for both types of tests. Thus, the intervention or treatment administered before the posttest improved participants' abilities in both EPT and APT aspects.

Table 8. EPT Difference Test after Intensive Training Program (ITP)

Test the Difference in Pretest and Posttest Results			
English Proficiency Test (EPT)	N	Test Normal[c]	Sig.
Listening_Pre & Listening_Post	100	No	0,427 [b]
Structure_Pre & Structure_Post	100	No	0,003 [b]
Reading_Pre & Reading_Post	100	Yes	0,273 [a]

[a]Paired Sample t-Test
[b]Wilcoxon Signed Rank Test
[c]Kolmogorov-Smirnov Test

Table 8 presents the results of the pretest and post-test score difference tests on three components of the English Proficiency Test (EPT), namely Listening, Structure, and Reading, each with 100 participants. Before conducting the difference test, a normality test was performed using the Kolmogorov-Smirnov test to determine the appropriate statistical test. The normality test results showed that the data for Listening and Structure were not normally distributed, so the non-parametric Wilcoxon Signed Rank Test was used.

Meanwhile, the data for Reading were normally distributed, so

the Paired Sample t-test was used. The significance values of each difference test showed that all three components had p-values < 0.05 (Listening = 0.427; Structure = 0.003; Reading = 0.273). This means there is a significant difference between the pretest and post-test results in the Structure component, as the significance value is less than 0.05 (p = 0.003). with the 95% confidence interval confirming the statistical significance of this difference. Meanwhile, there is no significant difference between the listening and reading components.

The results from the Structure session showed a significant relationship between the pre- and post-test scores. This indicates that the ITP program effectively improves students' abilities. As for the Listening and Reading sessions, there was no significant relationship between the pre- and post-test scores. This indicates that the ITP program's teaching of Listening and Reading does not have a substantial impact or relationship, so improvements are needed in this area in the future.

In contrast, there was no significant relationship between the pre- and post-test scores in the Listening and Reading components, suggesting that the current ITP curriculum has limited impact on these receptive skills. The lack of improvement in Reading/Qira'ah, in particular, may be attributed to several interrelated factors. First, reading comprehension requires not only vocabulary and grammar knowledge but also background knowledge, inferencing ability, and exposure to diverse text types—competencies that are difficult to develop within a limited 27-JPL (1,620 minutes) training period. Second, many students come from non-language majors and may have limited prior exposure to academic or authentic texts in English and Arabic, placing them at a disadvantage when tackling complex passages in the EAPT.

Furthermore, the current Reading module may emphasize literal comprehension and translation rather than higher-order thinking skills such as critical analysis, skimming, scanning, or guessing meaning from context. This form of instruction may not fully align with the cognitive demands of standardized proficiency tests. Additionally, the absence of extensive reading practice—where learners engage with large volumes of comprehensible input—may limit vocabulary acquisition and fluency development, both of which are crucial for reading improvement (Agustin et al., 2023).

These findings are consistent with prior research indicating that receptive skills (listening and reading) often show slower progress in intensive programs compared to productive or discrete skills (Tseng et al., 2024). Therefore, while the ITP has proven effective in strengthening structural accuracy, future iterations should consider redesigning the Reading and Listening components to include more authentic materials, strategy-based instruction, and extended practice opportunities. Such enhancements could help bridge the gap between training outcomes and real-world language proficiency.

Table 9. APT Difference Test After Intensive Training Program (ITP)

Test the Difference in Pretest and Posttest Results

Arabic Proficiency Test (APT)	N	Test Normal[b]	Sig.
Istima_Pre & Istima_Post	100	No	0,000[a]
Tarakib_Pre & Tarakib_Post	100	No	0,000[a]
Qira'ah_Pre & Qira'ah_Post	100	No	0,000[a]

[a] Wilcoxon Signed Rank Test
[b] Test Kolmogorov-Smirnov

Based on Table 9 above, a normality test was first conducted using the Kolmogorov-Smirnov test. The test results showed that all pretest and posttest data from the three subtests, namely *Istima'*, *Tarakib*, and *Qira'ah*, were not normally distributed. Therefore, the Wilcoxon Signed Rank Test, which is appropriate for non-parametric data, was used to test the difference between pretest and posttest scores. The results of the Wilcoxon test showed that all three subtests had a significance value of 0.000 (p < 0.05), with the 95% confidence interval confirming the statistical significance of these differences. This value indicates a significant difference between each APT subtest's pretest and posttest scores. In other words, there was a substantial improvement in the participants' Arabic language proficiency after the treatment or intervention.

The *Tarakib* subtest significantly improves the students' ability to understand Arabic language structure from before to after the program. Similarly, in the Qira'ah stage, students experienced significant progress compared to the pretest. Table 9 shows a strong and important relationship between the Arabic language proficiency test results before and after the ITP program. Thus, the ITP program was successful in improving students' APT scores.

Learning in the Intensive Training Program (ITP) has generally been conducted and followed well, enabling a correlation between ITP and improvements in students' EPT scores. In the pretest analysis for EPT, two students achieved the minimum score. However, the improvement was still very small. Out of 100 students who participated in ITP, only 17 students achieved the minimum passing score of 450.

Several factors contribute to the failure to achieve the minimum score. First, the training period was relatively short, with each skill receiving only 180 minutes, or the equivalent of 3 hours per skill. If the training period were extended, many students would likely be able to achieve the minimum score of 450. Second, the students who took the EPT were from non-language majors and did not have sufficient basic English skills. To address this issue, providing introductory English modules specifically designed for students from non-English language backgrounds before they take the ITP can help them build a stronger foundation in English. This will give them the basic skills they need to participate in the training more effectively.

In addition, motivation and active participation in learning English are crucial in achieving good results, especially in an Intensive Training Program (ITP) context(Hamadtoh A. Gohar, 2023; Shebani et al., 2025). One way to increase motivation is to encourage students to

practice various tests, such as listening, reading, and sentence structure tests(Ahmed Abdel-Al Ibrahim et al., 2023; Aryadoust et al., 2024; Saito et al., 2023) By practicing frequently, students will understand the format and types of questions they will encounter and increase their confidence when taking exams. This regular practice helps students identify their weaknesses and provides opportunities to improve areas that need improvement. Moreover, consistent exposure to test-like conditions enhances familiarity with time constraints and question patterns, which significantly reduces test anxiety and improves overall performance (Meylinda, 2025).

According to Chanwaiwit & Wiboonwachara, (2025), sustained motivation in language learning can be fostered through goal-setting, immediate feedback, and a sense of progress—elements often embedded in structured practice routines. Furthermore, research by Huang highlights that students who engage in regular mock testing demonstrate higher retention rates and better strategic thinking, particularly in standardized tests like TOEFL or IELTS (Huang & Huang, 2024).

Therefore, integrating frequent practice sessions into ITP curricula not only boosts learners' competence but also strengthens their self-efficacy and perseverance in mastering English proficiency(Rahimi & Mosalli, 2025; Saito et al., 2023). Student engagement in learning can be improved through various learning resources, such as books, applications, or online platforms that offer interactive exercises. Activities like mock exams and quizzes can create a fun competitive atmosphere, encouraging students to participate actively. By involving them in challenging and enjoyable activities, students will be more motivated to learn and practice consistently (Zen et al., 2022).

Support from instructors and peers is also very influential in maintaining motivation (Wang et al., 2023). After practice, instructors who provide positive and constructive feedback can help students feel valued and motivated to continue trying. In addition, forming study groups where students can support each other and share learning strategies can also increase their activity (Scott-Barrett et al., 2023).

With a supportive learning environment, students will be more motivated to practice and strive to achieve higher scores on English tests. Unlike the EPT, in the pretest analysis for the Arabic Proficiency Test (APT), six students achieved scores above the minimum. However, after the training, the number of students who completed the minimum score increased significantly to 32.

Background knowledge in Islamic Studies (Islamic Studies) was the primary factor supporting their APT success. This is because the types of questions in Arabic are closely related to Islamic knowledge, providing an additional advantage for students from the departments of Islamic Education (PAI), Islamic Education Management (MPI), Islamic Economics (ES), and Family Law (HK), even though they are not from

language-related departments.

Other factors contributing to this improvement include interactive and relevant teaching methods and the use of teaching materials appropriate to the Islamic context(Jasadijaya et al., 2023; Nor et al., 2024) In addition, support from lecturers and a conducive academic environment alsoplay an essential role in motivating students to learn and participate actively in training(Ruijuan et al., 2023). This improvement demonstrates that students can succeed in language proficiency exams even without a specialized language education background with the right approach and relevant materials (Li & Pei, 2024).

Description		Arabic		Amount
		L	TL	
English	L	7%	10%	17%
	TL	25%	58%	83%
Amount		32%	68%	100%
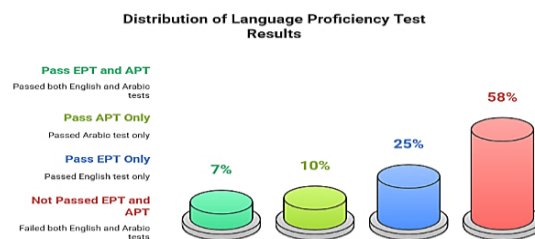
Figure 6. Distribution of Language Proficiency Test Result

Figure 6 present the distribution of participants' passing grades in the English Proficiency Test (EPT) and Arabic Proficiency Test (APT) based on passing (L) and failing (TL) categories. Out of 100 participants, 32% achieved scores above the passing threshold on the ATP, while 17% met the passing criteria on the ETP. A cross-analysis of the results of both tests shows that 7 participants (7%) achieved the passing standards for both the English and Arabic Proficiency Tests. In addition, 10 participants (10%) passed the ATP despite not meeting the criteria for the ETP, and 25 participants (25%) passed the ETP but did not pass the ATP. Fifty-eight participants (58%) did not meet the passing threshold for both tests. These results indicate positive progress among some participants, particularly regarding Arabic language proficiency. Although most participants did not pass both tests, this data reflects the potential for further improvement through strengthening the learning program focused on both language aspects.

These findings also provide an initial constructive overview for evaluating the effectiveness of the implemented learning strategies and identifying more specific areas for development in each language skill. The analysis results of the above factors have been translated into practical recommendations that will help improve the quality of language training programs at IAIN Sultan Amai Gorontalo. Thus, the most critical areas for improvement are the technical implementation of the Intensive Training

Program (ITP), including aspects such as time, training materials, teaching methods, learning motivation, and active participation in various tests.

CONCLUSION

This study concludes that the Intensive Training Program (ITP) significantly enhances Arabic language proficiency, as evidenced by the increase in students achieving the minimum passing score from 6 to 32 participants (a 26% improvement). Strong correlations between pretest and posttest scores in the Arabic components—istima' ($r = 0.972$), tarakib ($r = 0.993$), and qira'ah ($r = 0.997$)—indicate that the teaching methods used are effective for skill development. In contrast, the program's impact on English proficiency is less pronounced, with only 17% of students reaching the minimum score of 450. Improvements were observed in listening and structure, but the reading component showed no significant progress ($r = 0.085$; $p = 0.686$), highlighting a misalignment between training materials and students' needs.

These findings emphasize the importance of ongoing evaluation of language training programs, especially considering the diverse academic backgrounds of students. Recommendations include extending training duration for better comprehension, developing modules tailored for non-language background students, enhancing motivation through regular practice and support, and revising teaching methods to improve reading skills in English. Overall, this study offers practical implications for higher education institutions, suggesting strategies for creating more effective language training programs that support graduate students in achieving proficiency in both Arabic and English, while also paving the way for future research on personalized learning approaches and technology-based instructional materials.

AUTHOR STATEMENTS

Ana Mariana collected and analysed the data and wrote the findings and discussion sections. In addition, She formatted the manuscript according to journal guidelines and carefully proofread all drafts of the article. **Ibnu Rawandhy N. Hula** conducted the tests at the language LAB, classified the students' English and Arabic test data, while **Dahlia Husain** and **Nurul Aini Pakaya**, assisted in analysing the data, co-authored the findings, translated and confirmed the accuracy of the data, and concluded and provided recommendations on the research results.

ACKNOWLEDGEMENTS

the Intensive Training Program (ITP) and the results of the computer-based EAPT at the Language Laboratory.

REFERENCES

Abdulrahman, S. A., & Kara, S. (2023). The Effects of Movie-Enriched Extensive Reading on TOEFL IBT Vocabulary Expansion and TOEFL IBT Speaking Section Score. *Journal of Qualitative Research in Education*, *23*(33), 176–200. https://doi.org/10.14689/enad.33.913

Agustin, L., Wisudaningsih, E. T., & Fatmawati, R. (2023). Exploring How Skimming and Scanning Fosters EFL Students' Reading Comprehension at an English Club Senior High School in Indonesia. *Tamaddun*, *22*(1), 20–27. https://doi.org/10.33096/tamaddun.v22i1.309

Ahmed Abdel-Al Ibrahim, K., Cuba Carbajal, N., Zuta, M. E. C., & Bayat, S. (2023). Collaborative learning, scaffolding-based instruction, and self-assessment: impacts on intermediate EFL learners' reading comprehension, motivation, and anxiety. *Language Testing in Asia*, *13*(1), 1–33. https://doi.org/10.1186/s40468-023-00229-1

Alhamami, M., & Almelhi, A. (2021). English or arabic in healthcare education: Perspectives of healthcare alumni, students, and instructors. *Journal of Multidisciplinary Healthcare*, *14*, 2537–2547. https://doi.org/10.2147/JMDH.S330579

Aljarelah, A. K. (2024). Language and International Communication the Impact of Language on International Understanding and Diplomatic Relations. *The American Journal of Social Science and Education Innovations*, *6*(6), 144–175. https://doi.org/10.37547/tajssei/volume06issue06-23

Aryadoust, V., Zakaria, A., & Jia, Y. (2024). Investigating the affordances of OpenAI's large language model in developing listening assessments. *Computers and Education: Artificial Intelligence*, *6*, 100204. https://doi.org/10.1016/j.caeai.2024.100204

Awad Al-Rohili, M. (2024). An Exploration of Language Learning Strategies by Non-native Arabic Speakers: Identifying Effective Approaches. *Pakistan Journal of Life and Social Sciences (PJLSS)*, *22*(1), 339–356. https://doi.org/10.57239/pjlss-2024-22.1.0026

Boy Jon, R., Embong, R., Purnama, B., & Safar Wadi, A. (2021). The Challenges of English Language Teaching in Indonesia. *International Journal of English and Applied Linguistics (IJEAL)*, *1*(3), 158–168. https://doi.org/10.47709/ijeal.v1i3.1157

Chanwaiwit, P., & Wiboonwachara, L. (2025). Perceptions of CMRU-TEP: English Majors versus Non-English Majors' Voices. *REFLections*, *32*(1), 374–394. https://doi.org/10.61508/refl.v32i1.280209

Dobrić, N., Sigott, G., Ilc, G., Lazović, V., Cesnik, H., & Stopar, A. (2021). Errors as indicators of writing task difficulty at the Slovene general matura in English. *International Journal of Applied Linguistics (United Kingdom)*, *31*(3), 475–491. https://doi.org/10.1111/ijal.12345

El-Dakhs, Dina Abdel Salam, Ahmed, Mervat M, Altarriba, Jeanette, & Sonbul, Suhad. (2023). Differential emotional expression in autobiographical narratives: The case of Arabic–English bilinguals. *International Journal of Bilingualism*, *28*(5),

962–979. https://doi.org/10.1177/13670069231201173

El Meneza, S. (2020). Egyptian neonatal safety training network: A dream to improve patient safety culture in egyptian neonatal intensive care units. *Eastern Mediterranean Health Journal*, *26*(10), 1303–1311. https://doi.org/10.26719/emhj.20.034

Em, S. (2024). Exploring Experimental Research: Methodologies, Designs, and Applications Across Disciplines. *SSRN Electronic Journal*, *March*, 1–9. https://doi.org/10.2139/ssrn.4801767

Hamadtoh A. Gohar, R. (2023). An Extensive Online Listening Program for Preparing Postgraduates to Take the TOEFL Listening Test and Enhancing their Autonomy. *مجلة كلية التربية. بنها*, *34*(136), 1–46. https://doi.org/10.21608/jfeb.2024.262663.1831

Huang, H., & Huang, J. (2024). Assessing the Impact of Disorganized Background Noise on Timed Stress Task Performance Through Attention Using Machine-Learning Based Eye-Tracking Techniques. *AAXiv Is Hiring a DevOps Engineer*, *13*(6), 1–18. https://doi.org/10.48550/arXiv.2410.04208

Izah, S. C., Sylva, L., & Hait, M. (2024). Cronbach's Alpha: A Cornerstone in Ensuring Reliability and Validity in Environmental Health Assessment. *ES Energy and Environment*, *23*, 1–14. https://doi.org/10.30919/esee1057

Jasadijaya, A., Bala, R., & Purnamasari, H. (2023). Enhancing Learning Achievement through Effective Use of Teaching Aids in Islamic Education. *International Journal of Learning Reformation in Elementary Education*, *2*(02), 104–113. https://doi.org/10.56741/ijlree.v2i02.322

Jendli, A., & Albarakati, M. (2024). Exploring Motivational Dynamics: The Role of Oral Activities in Improving Arab Students' Learning of English. *International Journal of Learning, Teaching and Educational Research*, *23*(3), 131–149. https://doi.org/10.26803/ijlter.23.3.7

Kubiszyn, T., & Borich, G. D. (2024). *Educational testing and measurement.* John Wiley & Sons.

Lee, S. (2022). Evaluating the New TOEFL ITP® Speaking Test: Insights From Field Test Takers. *ETS Research Report Series*, *2022*(1), 1–19. https://doi.org/10.1002/ets2.12352

Li, M., & Pei, L. (2024). Exploring challenges in academic language-related skills of EFL learners in Chinese EMI settings. *Acta Psychologica*, *247*, 104309. https://doi.org/10.1016/j.actpsy.2024.104309

Meylinda, L. (2025). Exploring Students' Perception on Difficulties of Structure And Written Expression in Toefl Test: A Case Study. *JEELL (Journal of English Education, Linguistics and Literature)*, *12*(1), 27–38. https://doi.org/10.32682/jeell.v1102/27

Nor, A., Yusuf, M., & Arabi, I. (2024). Strategies for Improving the Professionalism of Islamic Education Teachers at University. *Tafkir: Interdisciplinary Journal of Islamic Education*, *5*(1), 40–61. https://doi.org/10.31538/tijie.v5i1.774

Nufus,erlina, koderi, H., Utama, M., & Ramadhan, C. (2022). Development of Tarkib Teaching Materials Based on Motion Graphic in Islamic Junior High School. *Jurnal Al Bayan: Jurnal Jurusan Pendidikan Bahasa Arab*, *14*(1), 40–57. https://doi.org/10.24042/albayan.v

Rahimi, A. R., & Mosalli, Z. (2025). Language learners' surface, deep, and organizing approaches to ChatGPT-assisted language learning: What contextual, individual,

and ChatGPT-related factors contribute? *Smart Learning Environments*, *12*(1). https://doi.org/10.1186/s40561-025-00368-3

Remler, D. K., & Van Ryzin, G. G. (2024). Control, Exogeneity, and Directness: Understanding and Designing Quasi- and Natural Experiments. *American Journal of Evaluation.* https://doi.org/10.1177/10982140241265708

Rocha, B. I. (2025). Language Assistant Programs and CLIL: Developing and Validating Questionnaires to Determine Program Training Needs. *Journal of Language Teaching and Research*, *16*(2), 359–370. https://doi.org/10.17507/jltr.1602.01

Ruijuan, L., Srikhoa, S., & Jantharajit, N. (2023). Blending of collaborative and active learning instructional methods to improve academic performance and self-motivation of vocational students. *Asian Journal of Education and Training*, *9*(4), 130–135. https://doi.org/10.20448/edu.v9i4.5211

Saito, K., Uchihara, T., Takizawa, K., & Suzukida, Y. (2023). Individual differences in L2 listening proficiency revisited: Roles of form, meaning, and use aspects of phonological vocabulary knowledge. *Studies in Second Language Acquisition*, 26–52. https://doi.org/10.1017/S027226312300044X

Scott-Barrett, J., Johnston, S.-K., Denton-Calabrese, T., McGrane, J. A., & Hopfenbeck, T. N. (2023). Nurturing curiosity and creativity in primary school classrooms. *Teaching and Teacher Education*, *135*, 104356. https://doi.org/https://doi.org/10.1016/j.tate.2023.104356

Shebani, Z., Aldhafri, S., & Alsaidi, F. (2025). The effect of parental involvement on academic passion: the mediating role of student motivation in learning English online. *International Journal of Adolescence and Youth*, *30*(1). https://doi.org/10.1080/02673843.2025.2467109

Shohamy, E. (2005). Language policy: Hidden agendas and new approaches. In *Language Policy: Hidden Agendas and New Approaches* (Issue May). https://doi.org/10.4324/9780203387962

Short, D. (2013). Training and Sustaining Effective Teachers of Sheltered Instruction. *Theory Into Practice*, *52*(2), 118–127. https://doi.org/10.1080/00405841.2013.770329

Tseng, W. T., Liu, Y. T., Hsu, Y. T., & Chu, H. C. (2024). Revisiting the effectiveness of study abroad language programs: A multi-level meta-analysis. *Language Teaching Research*, *28*(1), 156–200. https://doi.org/10.1177/1362168820988423

Wang, Y., Jiang, G., Yao, Z., & Liu, L. (2023). The influence of teacher–student relationship on Chinese high school students' academic motivation for the ideological and political subject: the mediating role of academic emotions. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1329439

Yu, H., & Hutson, A. D. (2024). A robust Spearman correlation coefficient permutation test. *Communications in Statistics - Theory and Methods*, *53*(6), 2141–2153. https://doi.org/10.1080/03610926.2022.2121144

Zen, Z., Reflianto, Syamsuar, & Ariani, F. (2022). Academic achievement: the effect of project-based online learning method and student engagement. *Heliyon*, *8*(11), e11509. https://doi.org/10.1016/j.heliyon.2022.e11509

Conflict of Interest Statement: The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.