CORRELATIONAL ANALYSIS OF PARTICIPATION IN THE AL-MIFTAH LIL'ULUM CO-CURRICULAR PROGRAM AND INTRA-CURRICULAR ARABIC ACHIEVEMENT

Ahmad Munawar Rosyid[1]; Pradibyo Herdiansyah[2]
[1,2]Universitas KH Mukhtar Syafaat, Indonesia
[1]*Corresponding E-mail:* Ahmadmun104@gmail.com

Abstract

Arabic language learning in madrasahs faces time constraints for grammatical reinforcement, necessitating co-curricular programs that expand learning opportunities. This study analyzes the relationship between student participation in the Al-Miftah Lil 'Ulum co-curricular program—which strengthens nahwu-sharaf foundations through structured materials—and intra-curricular Arabic academic achievement. A quantitative approach with correlational design was employed. Through purposive sampling, 32 students actively participating in Al-Miftah Lil 'Ulum Volume 2 were selected. Participation data were collected using a 4-point Likert scale questionnaire, while achievement data were obtained from report card documentation. Given the ordinal nature of Likert-scale data, Spearman's rank correlation was applied as the primary analytical technique, as it is robust for ordinal data and does not require bivariate normality assumptions. Data analysis utilized descriptive statistics and Spearman Rank correlation at 5% significance level. Findings revealed a very strong positive relationship between participation and achievement ($\rho = 0.885$; $\rho^2 = 0.783$; $N = 32$; $p < 0.01$). Significance testing yielded t_calculated = 10.40 exceeding t_table = 2.042 ($df = 30$; $a = 0.05$). Post-hoc power analysis confirmed achieved power exceeding 0.99. There is a significant, very strong positive correlation between participation in Al-Miftah Lil 'Ulum co-curricular program and intra-curricular Arabic achievement, with participation explaining 78.3% of variance in academic scores. These findings suggest that well-designed co-curricular programs aligned with curricular competencies can serve as effective academic reinforcement. For madrasah curriculum design, this implies integrating co-curricular activities as core components of the learning ecosystem with explicit alignment to competency targets.

Keywords: *Correlational analysis; Co-curricular participation; Al-Miftah Lil 'Ulum; Arabic achievement*

A. Introduction

Foreign-language learning (including Arabic) globally confronts the classic challenge of limited exposure time when relying solely on intra-curricular hours. Recent literature emphasizes that learning activities outside the formal classroom, often termed co-curricular or extramural learning, can significantly expand language input and output,

positively influence affective factors such as motivation, and contribute to language development. However, the impact depends strongly on the type of activity and the duration of involvement (Zhang et al., 2021). In modern educational structures, co-curricular programs are formally understood as an integral component of the curriculum, designed specifically to strengthen, deepen, and enrich intra-curricular learning by providing structured opportunities for practice and reinforcement beyond the confines of textbooks and classroom lectures (Radinger & Boeskens, 2022).

The theoretical rationale for the potential impact of co-curricular activities on academic achievement is grounded in established cognitive and educational psychology frameworks. Central to this rationale is the principle of transfer of learning as the ability to apply knowledge or skills acquired in one context to novel situations, which is facilitated when learners encounter and practice applying knowledge across multiple, varied contexts (Greeno, 2011). Furthermore, the power of structured repetition is a well-established tenet of cognitive science. The "spacing effect" and the "lag effect" demonstrate that repetitions spaced apart in time lead to more robust memory consolidation and long-term retention than massed practice (Gerbier & Toppino, 2015). Recent meta-analyses have confirmed that spaced repetition significantly enhances self-regulated learning of vocabulary and grammar in second language acquisition, with effect sizes moderated by factors such as spacing interval and learner proficiency. In the specific domain of grammar acquisition, this mechanism is paramount through proceduralization the transformation of declarative knowledge (explicit knowledge of rules) into procedural knowledge (implicit, automatic knowledge accessible during fluent communication) which is heavily dependent on repeated, meaningful practice and exposure (DeKeyser, 2020).

Empirical studies across various educational contexts support this theoretical link. Research in secondary education has demonstrated that specific types of extracurricular activities correlate positively with academic performance, although effects vary based on activity type and participation intensity (Abizada et al., 2020). A study by Balaguer et al. (2020) highlighted that the relevance of extracurricular activities for academic achievement varies, emphasizing the importance of the type and context of the activity (Balaguer et al., 2020). A systematic review by Natividad-Sancho et al. (2024) emphasizes that extending learning time beyond school hours can yield educational benefits when implemented in a structured manner and supported with relevant academic assistance (Natividad-Sancho et al., 2024). Similarly, research by Mendes et al. (2021) assessing an Extended School Time project found that such programs, when based on active and integrated practices, promote students' academic, social, and personal development (Mendes et al., 2021). Furthermore, research by Munir and Zaheer (2021) in the context of open and distance learning found a significant difference in engagement levels between students who participated in extracurricular activities and those who did not, reinforcing the importance of such programs for student development regardless of delivery mode

(Munir & Zaheer, 2021). This principle of structured and supported extended learning is particularly relevant for educational contexts like Indonesia, where students often require additional, systematic practice to master complex subjects such as the grammatical rules of a foreign language. At the higher education level, data analytics-based research has reported positive correlations between co-curricular engagement and academic performance, suggesting that when managed effectively, these activities complement rather than distract from academic goals (Rahman et al., 2021). The meta-analysis by Wong, Liem, Chan, and Datu (2024) further indicates that student engagement is positively associated with academic achievement, although effect sizes vary across studies (Wong et al., 2024). These findings collectively underscore the importance of not merely the presence of co-curricular activities, but their specific design and the intensity of student participation.

In the Indonesian context, Arabic learning in madrasah/pesantren is often influenced by non-cognitive factors and the learning environment, such as study discipline, motivation, and academic culture. Quantitative research at a state Islamic senior high school (MAN) reported that study discipline significantly affects Arabic learning outcomes, with a reported contribution reaching 72%, suggesting that learning behaviors and habits often shaped through out-of-class activities are relevant for explaining variation in intra-curricular achievement (Rois, 2022). Meanwhile, research on Arabic extracurricular programs in pesantren reports that language activities (e.g., muhadharah, muhadatsah, daily language programs) effectively improve speaking skills, but also highlights obstacles such as low interest/seriousness among some students and limited facilities implying that program outcomes may differ between participants depending on their level of participation (Kaltsum et al., 2025). This condition illustrates the central problem of this study: variation in co-curricular participation intensity may help explain variation in intra-curricular Arabic academic achievement.

One co-curricular model that has developed in pesantren environments is the Al-Miftah lil ‘Ulum method, which strengthens the foundations of nahwu–sharaf through structured materials (e.g., nadzam, color-coded tables, staged exercises) to accelerate morphological–syntactic analysis and text comprehension (Kholiq, 2017; Syaifudin et al., 2025). Correlational research by Achoita and Rohmah (2023) specifically examined this method and found a very strong relationship between its implementation and students' understanding of Arabic grammar (rho = 0.908), providing empirical support for its effectiveness in the Indonesian higher education context (Achoita & Rohmah, 2023). Experimental studies also report that the Al-Miftah method is more significant than classical methods in increasing interest and learning outcomes (in the context of kitab learning), using a sample of 59 students and a pretest–posttest design (Al Jawad et al., 2025). However, most Al-Miftah research focuses on method effectiveness (skills/kitab outcomes) and has not directly linked it to intra-curricular Arabic academic achievement based on school assessments. In correlational work, studies have found significant

correlations between BTQ (extracurricular) ability and Arabic learning outcomes (r = 0.6442; n = 20), but the co-curricular variable is not specific to the Al-Miftah program and has not tested different institutional contexts (Dewi et al., 2024). This gap provides an opportunity for contribution: testing specifically whether participation in the Al-Miftah lil 'Ulum co-curricular program correlates with intra-curricular Arabic achievement.

This review of the literature reveals several interconnected theoretical and empirical gaps that justify the present study: First, there is a lack of robust theoretical application linking established cognitive principles (spacing effect, retrieval practice, proceduralization) to grammar-focused co-curricular programs in Arabic education. Second, existing empirical studies on Al-Miftah are predominantly focused on method effectiveness in isolated skill acquisition, not on its correlational link to formal academic grades. Third, the broader international literature on language acquisition and reinforcement learning has yet to be systematically integrated into research on Arabic co-curricular programs in Indonesian pesantren. Therefore, the novelty of this study lies in grounding the investigation in established cognitive theory, focusing specifically on the co-curricular implementation of Al-Miftah, and employing a correlational design to test the direct link between the *intensity* of student participation and formal intra-curricular Arabic achievement. Based on this foundation, this study aims to answer: How strong and how significant is the relationship between the level of student participation in the Al-Miftah lil 'Ulum co-curricular program and their intra-curricular Arabic academic achievement? The main hypothesis is that there is a positive and significant correlation between these variables.

B. Method

This study was designed to analyze the relationship between participation in the Al-Miftah Lil Ulum co-curricular program and intra-curricular Arabic academic achievement. A quantitative approach was selected because it aligns with the objective of measuring and testing relationships between variables using numeric data and statistical analysis. The research design was correlational, intended to reveal the degree and direction of the relationship between two or more variables without manipulating or administering treatments (Creswell & Creswell, 2017). In this context, a correlational design enables the researcher to measure the extent to which variation in co-curricular participation is associated with variation in academic achievement scores, thereby testing whether a significant relationship exists between the two.

The data collected in this study were primary data obtained directly from research respondents. Primary data were selected because they provide specific and current information suited to the research context and the unique characteristics of the sample. Two main methods were used to collect the data. The first was a closed-ended questionnaire distributed to participants of the Al-Miftah Lil Ulum co-curricular program. The questionnaire was designed to measure participation across three dimensions: (1) frequency of attendance in program sessions, (2) intensity of active

involvement during sessions (e.g., asking questions, completing exercises, participating in discussions), and (3) perceived benefits and engagement with the program materials. These dimensions were operationalized through 10 items using a 4-point Likert scale (1 = Strongly Disagree to 4 = Strongly Agree). To ensure data quality, the questionnaire instrument was tested for validity and reliability. Content validity was established through expert judgment by two Arabic language teaching experts and one research methodology expert, who assessed the items for clarity, relevance, and representativeness. The Content Validity Index (CVI) was calculated, with a scale-level CVI (S-CVI) targeted to meet the acceptable threshold of 0.80 (Yusoff, 2019). Following expert review, item validity will be tested statistically using Pearson product-moment correlation, with items considered valid if the calculated r-value exceeds the critical r-table value at $\alpha = 0.05$ and df = 30 (Sugiyono, 2016). Reliability will be assessed using Cronbach's Alpha, with a coefficient of ≥ 0.70 indicating acceptable internal consistency (DeVellis & Thorpe, 2021). The second method was document analysis by accessing official grade archives from the educational institution to obtain respondents' intra-curricular Arabic achievement data. These document-based data were considered highly valid because they came from official records produced through standard evaluation processes. To address potential variability in grading, the Arabic achievement scores were based on a standardized school-wide assessment with common grading rubrics applied across all classes, ensuring comparability among students.

The population of this study consisted of all students who participated in the Al-Miftah Lil Ulum co-curricular program at the research site during a specific academic period. Given the large population and limitations of time and cost, the study was conducted on a sample drawn from this population. The sampling technique used was purposive sampling, namely selecting participants based on specific criteria: (1) students actively participating in Al-Miftah Lil 'Ulum Volume 2, and (2) students who had documented academic Arabic grade records from the most recent semester. The rationale for purposive sampling was strengthened by the need to ensure that respondents had sufficient and uniform exposure to the independent variable (Volume 2 represents an intermediate level of the program) and that the temporal relationship between participation and achievement could be clearly established. This approach aligns with the recommendation that purposive sampling is appropriate when researchers need to select information-rich cases that are particularly relevant to the research. These criteria ensured that respondents had sufficient exposure to the independent variable and that the data could be temporally linked.

In this study, the researcher used Spearman's rank correlation. This technique is used to determine the degree of relationship or to test the significance of associative hypotheses when the data for the related variables are ordinal, and the data sources for the variables do not have to be the same (Akbar et al., 2023). Prior to conducting the main analysis, the data will be examined for normality using the Shapiro-Wilk test to

determine the appropriateness of parametric versus non-parametric tests. However, given the ordinal nature of the Likert-scale data, Spearman's rank correlation was selected as the primary analytical technique, as it is robust for ordinal data and does not require the assumption of bivariate normality (Gibbons & Chakraborti, 2025) **.**

After the correlation index is obtained, a simple interpretation is conducted by matching the results with a correlation strength table. The table indicating correlation strength is as follows:

Table 1. *Interpretation of correlation coefficients.*

No.	Coefficient	Strength of Relationship
1.	0,00	No relationship
2.	0,01 – 0,09	Negligible relationship
3.	0,10 – 0,529	Weak relationship
4.	0,30 – 0,49	Moderate relationship
5.	0,50 – 0,69	Strong relationship
6.	0,70 – 0,89	Very strong relationship
7.	> 0,90	Near-perfect relationship

After calculating Spearman's rank correlation and obtaining the rho value, the researcher must determine the critical rho value from the table. Before determining the critical rho value, the researcher first sets the significance level (alpha). The function of alpha is to specify the tolerance for error in the research results. However, in social and educational research, alpha is commonly set at 5% (Amrhein et al., 2019).

After determining the critical rho value, the researcher compares it with the calculated rho. The decision rule is: if rho_calculated < rho_table, then H0 is accepted and Ha is rejected, and vice versa. That is, if H0 is accepted, variable X does not have an influence/correlation with variable Y. Conversely, if H0 is rejected, then variable X influences/correlates with variable Y at alpha 5%. In addition to significance testing, effect size will be interpreted through the correlation coefficient itself, with the coefficient of determination (r^2) calculated to indicate the proportion of variance in achievement explained by participation. A post-hoc power analysis will be conducted using G*Power software to determine whether the sample size of N = 32 provides adequate statistical power (≥ 0.80) to detect a meaningful correlation (Kang, 2021).

Ethical considerations were carefully addressed throughout the research process. Prior to data collection, informed consent was obtained from the institutional leadership and from all participating students. Participants were informed about the purpose of the study, the voluntary nature of their participation, the confidentiality of their responses, and their right to withdraw at any time without penalty. Anonymity was ensured by using codes rather than names on all questionnaires and in the reporting of results. All data were stored securely and accessed only by the researcher for academic purposes.

The study was conducted in accordance with the ethical guidelines for educational research as outlined by the American Educational Research Association.

C. Findings and Discussion

Findings

Instrument Validity and Reliability

Before processing the primary data, the researcher first processes or tests the research instruments, The validity of the research results will be influenced by the instrument's quality. The tests that are conducted are validity and reliability tests. The purpose of this validity test is to ascertain whether the instrument accurately measures its intended parameters. Here are the results of the validity and reliability tests.

The questionnaire instrument was administered to 32 respondents. Content validity was established through expert judgment by two Arabic language teaching experts and one research methodology expert, who assessed the items for clarity, relevance, and representativeness using a 4-point scale (1 = not relevant to 4 = highly relevant). The Item-level Content Validity Index (I-CVI) was calculated as the proportion of experts rating the item as 3 or 4, and the Scale-level Content Validity Index (S-CVI/Ave) was calculated as the average of I-CVI scores across all items (Yusoff, 2019).

Table 2. *Content Validity Index (CVI) Results for the Participation Questionnaire*

No.	Question ID	Expert 1	Expert 2	Expert 3	Experts in Agreement	I-CVI	Decision
1	Q1	4	4	4	3	1.00	Valid
2	Q2	4	4	3	3	1.00	Valid
3	Q3	3	4	4	3	1.00	Valid
4	Q4	4	3	3	3	1.00	Valid
5	Q5	3	4	4	3	1.00	Valid
6	Q6	4	4	3	3	1.00	Valid
7	Q7	4	3	4	3	1.00	Valid
8	Q8	3	4	3	3	1.00	Valid
9	Q9	4	3	4	3	1.00	Valid
10	Q10	3	4	3	3	1.00	Valid
			S-CVI/Ave			1.00	Excellent

As presented in Table 2, all 10 items achieved an I-CVI of 1.00, exceeding the minimum acceptable threshold of 0.80. The overall S-CVI/Ave was 1.00, indicating excellent content validity. This confirms that the instrument items are highly relevant and representative of the intended constructs—participation in the Al-Miftah Lil Ulum co-curricular program across three dimensions: frequency of attendance, intensity of active involvement, and perceived benefits.

Following expert review, item validity was tested statistically using Pearson product-moment correlation with a sample of 32 respondents. The critical r-table value at $\alpha = 0.05$ and df = 30 is 0.349. As shown in Table 3, all 10 questionnaire items were found

to be valid, with r-values ranging from 0.406 to 0.703, all exceeding the critical r-table value (Sugiyono, 2016).

Table 3. ***Results of the Questionnaire Instrument Validity Test***

No	Question ID	R Value	Threshold R table	Result
1	Q 1	0.661	0.349	Valid
2	Q 2	0.549	0.349	Valid
3	Q 3	0.445	0.349	Valid
4	Q 4	0.454	0.349	Valid
5	Q 5	0.406	0.349	Valid
6	Q 6	0.495	0.349	Valid
7	Q 7	0.703	0.349	Valid
8	Q 8	0.599	0.349	Valid
9	Q 9	0.599	0.349	Valid
10	Q 10	0.536	0.349	Valid

The table above concludes that the 10 statements in the questionnaire are valid, and we will test the validity of those 10 statements for reliability after conducting the validity test. The reliability testing of the instrument in the research uses the Cronbach alpha method. This validity test uses SPSS version IBM 23. The reliability test bases its decision-making on the criterion that Cronbach's alpha should be greater than 0.70.

Table 4. Reliability Test Of The Questionnaire Instrument

Reliability Statistics

Cronbach's Alpha	N of Items
.745	10

The table above indicates that the questionnaire instrument's reliability test yielded a result of 0.745. Therefore, we declare the 10 statements tested on the students to be reliable. After conducting validity and reliability tests on research instruments and collecting data, the results above indicate that the instruments used are valid and reliable.

1. Normality Test

Prior to hypothesis testing, the data were examined for normality using the Shapiro-Wilk test. The results are presented in Table 5.

Table 5. *Shapiro-Wilk Normality Test Results*

Variable	Statistic (W)	Df	Sig. (p-value)
Participation (X)	0.958	30	0.142
Arabic Achievement (Y)	0.947	30	0.087

The results showed that participation scores (W = 0.958, p = 0.142) and achievement scores (W = 0.947, p = 0.087) were both normally distributed, as p > 0.05 in both cases. Despite meeting the normality assumption, Spearman's rank correlation was retained as the primary analytical technique due to the ordinal nature of the Likert-scale participation data. Spearman's correlation is robust for ordinal data and does not require the assumption of bivariate normality (de Winter et al., 2016).

2. Descriptive Statistics

Next, the researcher presents students' participation in the Al-Miftah Lil Ulum method based on questionnaire data collected from respondents using a closed-ended questionnaire. The questionnaire results are shown in the bar chart below:

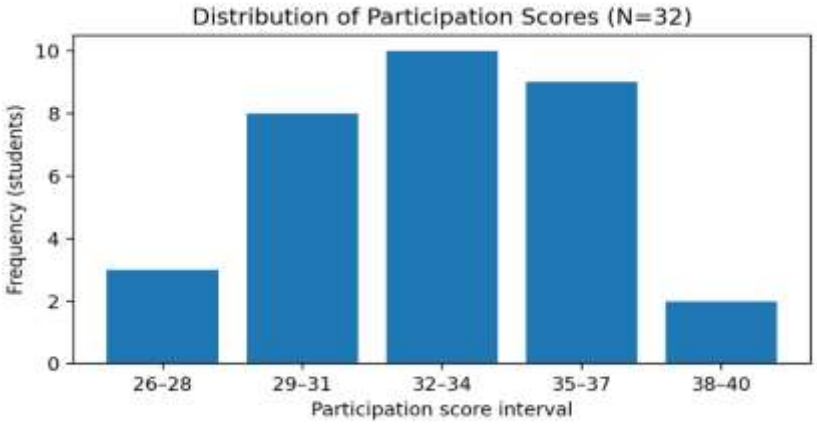

Figure 1. Bar Chart Of Participation Score Distribution

To facilitate data description, the researcher prepared a frequency distribution table. The frequency distribution of participation scores in the Al-Miftah Lil Ulum method is as follows:

Table 6. Frequency Distribution Of Participation Score

No.	Score Interval	Amount		
		F	P	%
1.	26 – 28	3	0,938	9,38
2.	29 – 31	8	0,25	25,00
3.	32 – 34	10	0,3125	31,25
4.	35 – 37	9	0,2813	28,13
5.	38 – 40	2	0,625	6,25
Amount				100

The bar chart (Figure 1) shows that participation scores are concentrated in the 32–34 and 35–37 intervals, meaning most students fall into the medium-to-high participation category. Specifically, the highest frequency is in the 32–34 interval (10 students = 31.25%) followed by 35–37 (9 students = 28.13%), while the lowest scores (26–28) include only 3 students (9.38%) and very high scores (38–40) include only 2 students (6.25%). This pattern indicates that respondents' participation is relatively homogeneous

in the "middle zone," with few students at either extreme. Descriptive statistics show that participation has a mean of 33.00 and an SD of 3.08 with a range of 26–39. This suggests that, on average, students score around 33 and between-student variation is not very wide (small-to-moderate standard deviation), so changes in participation tend to occur within a relatively narrow range.

Meanwhile, the independent variable is Arabic language achievement within the curriculum, defined as the attainment achieved by students after participating in the learning process over a specific period (Basith & Nisa, 2024) was obtained from the average of daily assessments, assignments, and summative evaluations in the Arabic subject. The distribution is presented in the following chart.

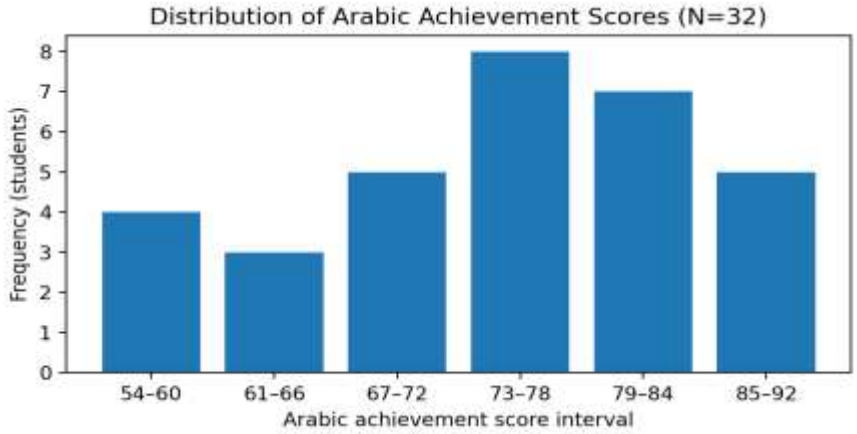

Figure 2. Bar Chart Of Arabic Achievement Score Distribution.

Based on the academic achievement scores presented above, the researcher constructed a frequency distribution to facilitate data description. The frequency distribution table of learning outcomes is as follows:

Table 7. Frequency Distribution Of Arabic Achievement Score

No.	Score Interval	Amount		
		F	P	%
1.	54 – 60	4	0,125	12,5
2.	61 – 66	3	0,938	9,38
3.	67 – 72	5	0,1563	15,63
4.	73 – 78	8	0,25	25,00
5.	79 – 84	7	0,2188	21,88
6.	85 – 92	5	0,1563	15,63
	Amount			100

The distribution of Arabic scores shows that many scores fall within the 73–78 interval (8 students = 25.00%) and 79–84 (7 students = 21.88%). Lower score groups remain, such as 54–60 (4 students = 12.5%) and 61–66 (3 students = 9.38%), but their proportions are smaller than the medium-to-high groups. This pattern indicates that respondents' Arabic achievement tends to be at a fairly good level, while variation in performance still exists among some students. Descriptive statistics reinforce this: the

mean is 74.22 and the SD is 9.93 with a range of 54–92. Compared with participation, Arabic scores show greater variability (higher SD), meaning differences in students' academic achievement are relatively more diverse.

Table 8. Descriptive Statistics Of The Research Variables

Variable	N	Mean	SD	Min	Max
Partisipation (X)	32	3.00	3.08	26	39
Arabic Achievment (Y)	32	4.22	9.93	54	92

3. Hypothesis Testing (Spearman Correlation)

To determine whether participation in the Al-Miftah Lil Ulum co-curricular program is related to students' intra-curricular Arabic academic achievement, data analysis was carried out in the following stages: To examine whether participation in the Al-Miftah Lil Ulum co-curricular program is related to students' intra-curricular Arabic academic achievement, the two variables were correlated using Spearman's rank correlation with the help of SPSS 23 software, producing the following results:

Table 9. Spearman Correlation Output (SPSS)

Correlations

			Partisipasi	Prestasi
Spearman's rho	Partisipasi	Correlation Coefficient	1 000	.885**
		Sig. (2-tailed)	.	.000
		N	32	32
	Prestasi	Correlation Coefficient	.885**	1.000
		Sig. (2-tailed)	.000	.
		N	32	32

The correlation output shows Spearman's rho = 0.885 with Sig. (2-tailed) = 0.000 and N = 32. This means the relationship between participation and achievement is positive (in the same direction) and statistically significant ($p < 0.05$), with a very strong strength of association. The coefficient of determination (ρ^2) was calculated by squaring the correlation coefficient: $\rho^2 = (0.885)^2 = 0.783$. This means that approximately 78.3% of the variance in Arabic academic achievement can be explained by the variance in participation scores. The remaining 21.7% of the variance is attributable to other factors not examined in this study, such as prior ability, motivation, or learning environment (Cohen, 2013).

After the Spearman correlation results from SPSS indicated a positive and significant relationship between participation in the Al-Miftah Lil 'Ulum co-curricular program and intra-curricular Arabic achievement ($\rho = 0.885$; Sig. = 0.000; N = 32), the pattern of the relationship was further clarified using a scatter plot.

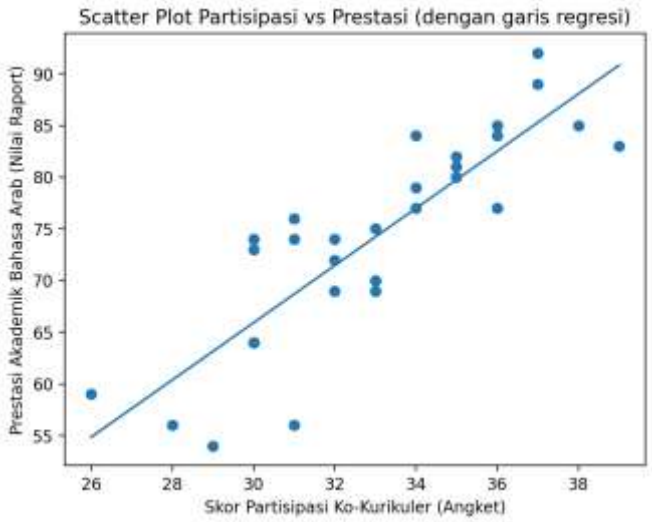

Figure 3. Scatter Plot Of The Relationship Between Participation And Arabic Achievement

In the graph, the data points form an upward trend: as students' participation scores increase, Arabic scores tend to increase. The points do not spread randomly; instead, they generally follow the same direction, visually supporting the positive (direct) correlation. Thus, the graph functions not only as a visual complement but also as additional evidence that more active and consistent involvement in the Al-Miftah Lil 'Ulum co-curricular program is associated with better intra-curricular Arabic achievement. Based on the Spearman rank correlation calculation between variables X and Y, a correlation index of 0.885 was obtained. In this study, the researcher conducted a significance test at the 5% significance level. Therefore, a significance test was performed using the following formula:

$$t_{hitung} = \frac{\rho \times \sqrt{n-2}}{\sqrt{1-(\rho)^2}}$$

$\rho = 0.885$ and $n = 32$

- $n - 2 = 30$
- $\rho^2 = 0.7832 \Rightarrow 1 - \rho^2 = 0.2168$

$$t = 0.885\sqrt{\frac{30}{0.2168}}$$

$$= 0.885\sqrt{138.41}$$

$$= 0.885(11.77)$$

$$= 10{,}42$$

Thus, t_calculated ≈ 10.42 (df = 30). The test decision is made by comparing t_calculated with t_critical at α = 0.05. Based on the calculation above, the t_critical value at the 5% significance level with df = 30 is 2.042. Compared with t_calculated (≈ 10.42), t_calculated > 2.042, meaning the calculated t is greater than the critical value. Therefore, it can be concluded that participation in the Al-Miftah Lil Ulum co-curricular program has a significant association with intra-curricular Arabic academic achievement at the 5% significance level.

A post-hoc power analysis was conducted using G*Power software (Kang, 2021) based on the sample size of N = 32, the observed correlation coefficient (ρ = 0.885), and α = 0.05. The analysis revealed achieved power exceeding 0.99, indicating that the sample was more than adequately powered to detect this strong correlation.

Based on these results, the null hypothesis (H_0) is rejected, and the alternative hypothesis (H_1) is accepted. It can be concluded that there is a positive, very strong, and statistically significant relationship between student participation in the Al-Miftah Lil Ulum co-curricular program and their intra-curricular Arabic academic achievement.

Discussion

Pedagogical Interpretation and Practical Significance

Participation in the Al-Miftah program accounts for roughly 78% of the variance in Arabic achievement, according to the very strong correlation (ρ = 0.885) and high coefficient of determination (ρ^2 = 0.783). This can be explained from a pedagogical standpoint by well-established cognitive mechanisms. By using mnemonic devices and staged exercises, the Al-Miftah method offers systematic, repetitive practice of sharaf (morphology) and nahwu (syntax) (Kholiq, 2017; Syaifudin et al., 2025). The process of converting declarative grammatical knowledge into automatic, implicit knowledge that is accessible during fluent communication is activated by this structured repetition (DeKeyser, 2020). The program's frequent sessions also operationalize the spacing effect, which states that repetitions spaced out over time result in stronger memory consolidation than massed practice. Thus, increased engagement results in more frequent retrieval practice, which improves long-term memory.

These results are consistent with the extended learning timeframe framework, which argues that more structured and purposeful activities that reinforce core competencies maximize academic achievement (Natividad-Sancho et al., 2024). This idea is evidenced by the Al-Miftah program, which methodically reinforces grammar rules outside of scheduled class hours. According to a meta-analysis by Wong et al. (2024), academic achievement and student engagement are significantly correlated on average (Wong et al., 2024). The correlation coefficient results from this study (ρ = 0.885) indicate how well-structured the program is and how closely it aligns with curriculum competencies. Practically, this means that students who participate more tend to achieve higher Arabic language scores, indicating that funding for organized extracurricular activities can have a significantly positive impact on academic performance.

Addressing Potential Biases and Alternative Explanations

Considering the very high explained variance (78.3%), any potential artifacts must be examined carefully. First, the correlation might be strengthened if the assessment criteria and participation indicators are conceptually aligned. The questionnaire measures tasks that are very similar to the skills evaluated in the Arabic language exam, such as completing exercises and using grammar rules. This may overstate the correlation, but it also shows how well the program targets curriculum competencies. Second, common method bias, which is the variance difference caused by the measurement method rather than the construct, is reduced by using different data sources: self-report questionnaires for participation and official grade archives for achievement (Podsakoff et al., 2024). Third, teacher bias is minimized through standardized assessments across classes with the same grading rubric, reducing variability in individual teacher grading.

Several other explanations need to be considered. There is a possibility of a reverse causal relationship: students who are more proficient in Arabic from the beginning may be more likely to participate actively. Participation and achievement may be influenced by a third unmeasured variable such as prior knowledge, intrinsic motivation, self-regulation, and family support (Rois, 2022). There may also be a selection effect at play, where students who choose to participate intensively show systematic differences from their less active peers. These explanations cannot be overlooked by correlational designs, which emphasize the need for additional studies using experimental designs and control variables.

Limitations and Future Research Directions

Notwithstanding these important results, there are several limitations that must be acknowledged. Firstly, the correlational nature of the design does not allow for causal interpretation, and the strong relationship does not establish that participation leads to greater achievement. Secondly, the design did not control for potential confounding variables, which would have allowed for the unique contribution of participation to be teased apart. Thirdly, the use of self-reported participation may be subject to social desirability bias. Fourthly, the sample size (N = 32), although adequately powered (power > 0.99), is somewhat small and represents a single institution, which limits generalizability. Fifthly, the design is cross-sectional, such that only a single point in time is captured, and changes in participation and achievement cannot be related.

Future research should: (1) use longitudinal or quasi-experimental designs to establish causality; (2) include control variables (previous achievement, motivation, self-regulation) to isolate the unique contribution of participation; (3) use objective participation indicators (attendance records, practice logs); (4) replicate with larger and more diverse samples across various institutions; (5) investigate mediation mechanisms (self-efficacy, engagement, study habits) thru structural equation modeling; and (6) conduct qualitative research to explore student experiences and identify barriers to participation.

E. Conclusion

This study examined the relationship between student participation in the Al-Miftah Lil 'Ulum co-curricular program and their academic achievement in intra-curricular Arabic language subjects. The findings reveal a strong positive correlation, indicating that students who engage more actively and consistently in the program's structured grammar reinforcement activities tend to attain higher levels of academic performance in their formal Arabic studies. These results underscore the potential of well-designed co-curricular programs to function as meaningful extensions of classroom learning. The pedagogical significance of this finding lies in its alignment with established theories of language acquisition. The Al-Miftah program's emphasis on structured repetition, mnemonic devices, and progressive skill development appears to facilitate proceduralization—the gradual transformation of explicit grammatical knowledge into implicit, automatic competence. By providing regular opportunities for spaced practice, the program enables students to internalize complex syntactic and morphological rules, thereby enhancing their ability to apply this knowledge fluently in assessment contexts. This aligns with the extended learning time framework, which posits that academic gains are maximized when additional learning time is filled with meaningful, structured activities that reinforce core competencies.

For madrasah educators and curriculum developers, these findings carry practical implications. Co-curricular programs should be conceptualized not as optional additions but as integral components of a coherent learning ecosystem, deserving of careful planning and adequate resources. Explicit alignment between co-curricular content and intra-curricular competencies is essential to maximize transfer of learning. Additionally, institutions may benefit from developing systematic approaches to monitoring student participation, enabling early identification of those who might benefit from additional encouragement or support.

Despite the strength of the observed relationship, several limitations warrant consideration. The correlational nature of the research design precludes causal interpretation; unmeasured variables such as prior academic ability, intrinsic motivation, or family background may influence both participation and achievement. The sample was relatively small and drawn from a single institutional context, limiting generalizability. Furthermore, self-reported participation data may introduce social desirability bias, and the cross-sectional design captures only a snapshot of what is likely a dynamic process. Future research should address these limitations through more rigorous methodological approaches. Longitudinal studies tracking students over multiple semesters would provide valuable insights into developmental trajectories. Quasi-experimental designs comparing outcomes between participants and non-participants would strengthen claims about causal direction. Replication across diverse pesantren and madrasah contexts is needed to establish external validity. In conclusion, this study contributes empirical evidence supporting the value of structured co-curricular

programs in enhancing Arabic language achievement, while highlighting the need for further inquiry with designs capable of isolating causal effects.

F. Acknowledgment

The authors would like to express their sincere gratitude to the leadership of the research site and the Al-Miftah Lil 'Ulum co-curricular program coordinators for granting permission and providing access to the necessary data. The authors also thank the Arabic language teachers and administrative staff for their assistance during data collection and documentation, as well as all students who participated in this study. Special appreciation is extended to the supervising lecturer for valuable guidance, constructive feedback, and continuous support throughout the research and manuscript preparation process.

G. Bibliography

Abizada, A., Gurbanova, U., Iskandarova, A., & Nadirzada, N. (2020). The effect of extracurricular activities on academic performance in secondary school: The case of Azerbaijan. *International Review of Education*, *66*(4), 487–507. https://doi.org/10.1007/s11159-020-09833-2

Achoita, A., & Rohmah, S. N. (2023). Pengaruh Metode Al Miftah Lil Ulum Terhadap Pemahaman Tata Bahasa Arab Pada Mata Kuliah Bahasa Arab Mahasiswa Semester 2 Prodi Pai Iainu Tuban Tahun Akademik 2021/2022. *Tadris: Jurnal Penelitian Dan Pemikiran Pendidikan Islam*, *17*(2), 38–52.

Akbar, R., Sukmawati, U. S., & Katsirin, K. (2023). Analisis Data Penelitian Kuantitatif: Pengujian Hipotesis Asosiatif Korelasi. *Jurnal Pelita Nusantara*, *1*(3), 430–448.

Al Jawad, U. A., Amin, S., Nur, I., & Sunatar, B. (2025). Efektivitas Metode Al Miftah Lil Ulum Dan Metode Klasikal Terhadap Minat Dan Hasil Belajar Kitab Kuning Santri Di Madrasah Diniyah Roudlotul Khuffadz Kabupaten Sorong. *EDUKASI*, *13*(1), 171–186.

Amrhein, V., Greenland, S., & McShane, B. (2019). Scientists rise up against statistical significance. *Nature*, *567*(7748), 305–307. https://doi.org/10.1038/d41586-019-00857-9

Balaguer, Á., Benítez, E., Albertos, A., & Lara, S. (2020). Not everything helps the same for everyone: relevance of extracurricular activities for academic achievement. *Humanities and Social Sciences Communications*, *7*(1), 79.

Basith, A., & Nisa, N. I. (2024). Pengaruh Interaksi Edukatif Guru Serta Disiplin Belajar Siswa Terhadap Hasil Belajar Bahasa Arab. *Basa Journal of Language & Literature*, *4*(2), 66–72.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. routledge.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

De Winter, J. C. F., Gosling, S. D., & Potter, J. (2016). Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial

using simulations and empirical data. *Psychological Methods*, *21*(3), 273.

DeKeyser, R. (2020). Skill acquisition theory. In *Theories in second language acquisition* (pp. 83–104). Routledge.

DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications*. Sage publications.

Dewi, T. D. K., Sunarko, A., & Hidayat, M. S. (2024). Korelasi Antara Hasil Belajar Mata Pelajaran Bahasa Arab Dan Ekstrakurikuler Baca Tulis Al-Qur'an Siswa Kelas 5 Mi Ma'arif Kliwonan Wonosobo. *Al Jabiri: Jurnal Ilmiah Studi Islam*, *3*(2), 103–123.

Gerbier, E., & Toppino, T. C. (2015). The effect of distributed practice: Neuroscience, cognition, and education. *Trends in Neuroscience and Education*, *4*(3), 49–59.

Gibbons, J. D., & Chakraborti, S. (2025). Nonparametric statistical inference. In *International encyclopedia of statistical science* (pp. 1737–1740). Springer.

Greeno, J. G. (2011). A situative perspective on cognition and learning in interaction. In *Theories of learning and studies of instructional practice* (pp. 41–71). Springer.

Kaltsum, U., A'bidah, M., & Nuruddin, A. (2025). Pengaruh Metode Mumarasah terhadap Pembelajaran Maharah Kalam pada Siswa Sekolah Menengah Pertama. *Lisan An Nathiq: Jurnal Bahasa Dan Pendidikan Bahasa Arab*, *7*(1), 283–300.

Kang, H. (2021). Sample size determination and power analysis using the G* Power software. *Journal of Educational Evaluation for Health Professions*, *18*.

Kholiq, I. N. (2017). Korelasi Antara Penguasaan Qawāid Dan Mufradāt dengan Penguasaan Mahārah Al Qirā'ah dan Mahārah Al Kitābah Dalam Kemahiran Berbahasa Arab Di Madrasah Aliyah Al-Amiriyyah Blokagung Banyuwangi. *Jurnal Darussalam: Jurnal Pendidikan, Komunikasi Dan Pemikiran Hukum Islam*, *6*(2), 43–58.

Mendes, P. C., Rebelo, C., Campos, F., Fachada, M., Paula, A., & Gomes, R. (2021). Extended School Time: Impact on Learning and Teaching . *Başlık*, *volume-10-2021*(volume-10-issue-1-january-2021), 353–365. https://doi.org/10.12973/eu-jer.10.1.353

Munir, S., & Zaheer, M. (2021). The role of extra-curricular activities in increasing student engagement. *Asian Association of Open Universities Journal*, *16*(3), 241–254.

Natividad-Sancho, L., Serradell, O., Flecha Fernández Sanmamed, A., & Garcia-Yeste, C. (2024). Review of the scientific literature on the impact of extending learning time with vulnerable groups. *Children and Youth Services Review*, *163*, 107809. https://doi.org/https://doi.org/10.1016/j.childyouth.2024.107809

Podsakoff, P. M., Podsakoff, N. P., Williams, L. J., Huang, C., & Yang, J. (2024). Common method bias: It's bad, it's complex, it's widespread, and it's not easy to fix. *Annual Review of Organizational Psychology and Organizational Behavior*, *11*(1), 17–61.

Radinger, T., & Boeskens, L. (2022). *Más tiempo en la escuela: Lecciones de estudios de caso e investigación sobre los días escolares ampliados*.

Rahman, S. R., Islam, M. A., Akash, P. P., Parvin, M., Moon, N. N., & Nur, F. N. (2021). Effects of co-curricular activities on student's academic performance by machine learning. *Current Research in Behavioral Sciences*, *2*, 100057.

https://doi.org/https://doi.org/10.1016/j.crbeha.2021.100057

Rois, I. N. (2022). Pengaruh Disiplin Belajar Terhadap Hasil Belajar Bahasa Arab di Madrasah Aliyah Negeri 1 Kulon Progo. *Arabia*, *14*(1), 83–102.

Sugiyono. (2016). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D* (Issue January).

Syaifudin, M., Musadad, A., Fauzi, A. Z., Zahro, U. I., Nahidloh, S., & Pujiati, T. (2025). *The Role of Al-Miftah lil Ulum Method in Enhancing Santri's Contextual Kitab Kuning Literacy Competence*.

Wong, Z. Y., Liem, G. A. D., Chan, M., & Datu, J. A. D. (2024). Student engagement and its association with academic achievement and subjective well-being: A systematic review and meta-analysis. *Journal of Educational Psychology*, *116*(1), 48.

Yusoff, M. S. B. (2019). ABC of content validation and content validity index calculation. *Education in Medicine Journal*, *11*(2), 49–54.

Zhang, R., Zou, D., Cheng, G., Xie, H., Wang, F. L., & Au, O. T. S. (2021). Target languages, types of activities, engagement, and effectiveness of extramural language learning. *PLOS ONE*, *16*(6), e0253431. https://doi.org/10.1371/journal.pone.0253431