Alibbaa': Jurnal Pendidikan Bahasa Arab
Special Issue Vol. 01 2026
P-ISSN: 2721-1606 | E-ISSN: 2716-4985
doi: https://doi.org/10.19105/ajpba.vi.24350

Enhancing Maharah Istima' Through Roleplay and Focus Group Discussion: Implications for AI-Supported Language Learning

Mushthafa Shadiq Al-Rafi'i[1], Nissa Media Utami[2], Anneu Rosyanti Suryana[3], Rizzaldy Satria Wiwaha[4]
[1,2,3,4] *Universitas Islam Negeri Sunan Gunung Djati Bandung, Indonesia*
Corresponding E-mail: shadiqar7@gmail.com

Abstract

Arabic language learning in Indonesia continues to face challenges, particularly in developing students' Maharah Istima' (listening skills), which are often hindered by limited vocabulary, rapid speech input, and the dominance of passive instructional approaches. This study aims to examine the effectiveness of integrating Roleplay and Focus Group Discussion (FGD) in improving students' listening ability within a classroom context. This research employed a quantitative approach using a quasi-experimental design with a non-equivalent control group. The participants consisted of 20 fifth-grade students at Al-Irhaam Global Islamic School Bandung, divided into an experimental group receiving Roleplay and FGD-based instruction and a control group receiving conventional audio-based instruction. Data were collected through pre-test and post-test and analyzed using Shapiro-Wilk normality test, homogeneity test, paired sample t-test, and N-Gain analysis. The findings indicate a statistically significant improvement in the experimental group ($p < 0.05$), with mean scores increasing from 72.0 to 88.0 and an N-Gain value of 0.63 (moderate category). Compared to the control group, the results suggest that interactive and collaborative learning activities contribute more effectively to the development of listening skills. These findings highlight the pedagogical potential of integrating Roleplay and FGD in promoting active engagement and deeper processing of language input. While this study is limited to a small-scale classroom context, it offers practical implications for designing interactive listening instruction

and suggests opportunities for future integration with AI-supported language learning tools.

Keywords: *Arabic Language Learning, Artificial Intelligence, Focus Group Discussion (FGD), Maharah Istima', Roleplay*

Abstrak

Pembelajaran bahasa Arab di Indonesia masih menghadapi berbagai tantangan, khususnya dalam pengembangan keterampilan menyimak (Maharah Istima'), yang sering terhambat oleh keterbatasan kosakata, cepatnya input ujaran, serta dominasi pendekatan pembelajaran yang bersifat pasif. Penelitian ini bertujuan untuk mengkaji efektivitas integrasi metode Roleplay dan Focus Group Discussion (FGD) dalam meningkatkan kemampuan menyimak siswa dalam konteks pembelajaran di kelas. Penelitian ini menggunakan pendekatan kuantitatif dengan desain kuasi-eksperimen (quasi-experimental design) tipe non-equivalent control group. Partisipan penelitian terdiri atas 20 siswa kelas V di Al-Irhaam Global Islamic School Bandung yang dibagi ke dalam kelompok eksperimen yang memperoleh pembelajaran berbasis Roleplay dan FGD serta kelompok kontrol yang memperoleh pembelajaran konvensional berbasis audio. Data dikumpulkan melalui pre-test dan post-test, kemudian dianalisis menggunakan uji normalitas Shapiro-Wilk, uji homogenitas, paired sample t-test, dan analisis N-Gain. Hasil penelitian menunjukkan adanya peningkatan yang signifikan secara statistik pada kelompok eksperimen ($p < 0,05$), dengan rata-rata skor meningkat dari 72,0 menjadi 88,0 dan nilai N-Gain sebesar 0,63 dalam kategori sedang. Dibandingkan dengan kelompok kontrol, hasil penelitian menunjukkan bahwa aktivitas pembelajaran yang bersifat interaktif dan kolaboratif lebih efektif dalam mengembangkan keterampilan menyimak siswa. Temuan ini menegaskan potensi pedagogis integrasi Roleplay dan FGD dalam mendorong keterlibatan aktif serta pemrosesan input bahasa yang lebih mendalam. Meskipun penelitian ini masih terbatas pada konteks kelas berskala kecil, hasil penelitian memberikan implikasi praktis bagi pengembangan pembelajaran menyimak yang interaktif serta membuka peluang integrasi dengan teknologi pembelajaran bahasa berbasis kecerdasan buatan (Artificial Intelligence/AI) pada penelitian selanjutnya.

Kata Kunci: *Pembelajaran Bahasa Arab; Artificial Intelligence; Focus Group Discussion (FGD); Maharah Istima'; Roleplay*

Introduction

Language learning fundamentally involves the development of communicative competence, which requires the integration of receptive and productive skills.[1] In the context of Arabic as a foreign language, *maharah istima'* (listening) and *maharah kalam* (speaking) are inherently interdependent, with listening serving as the foundational basis for the development of speaking ability.[2] However, despite its foundational role, listening instruction in many Indonesian classrooms remains pedagogically underdeveloped. As noted by Sari and Kholifah, difficulties in *maharah istima'* extend beyond linguistic limitations and involve cognitive, pedagogical, and affective dimensions, indicating that current instructional practices have not sufficiently facilitated meaningful and active processing of language input.[3]

From a theoretical perspective, listening in a second language is a complex cognitive activity that involves decoding auditory signals, constructing meaning, and employing metacognitive strategies simultaneously.[4] However, existing studies in Arabic language learning contexts tend to focus predominantly on identifying learners' difficulties such as limited vocabulary, rapid speech rate, and low concentration.[5] Without adequately addressing how instructional design can support deeper cognitive processing during listening. This tendency reflects a broader pedagogical limitation, where listening is treated as a passive activity centered on comprehension outcomes rather than as an active process involving interpretation, inference, and meaning construction.

In parallel, advancements in educational technology, particularly Artificial Intelligence (AI), have introduced new possibilities for

---

[1] Ali, M. (2020). Pembelajaran bahasa indonesia dan sastra (basastra) di sekolah dasar. *Pernik*, *3*(1), 35-44

[2] Ubaidillah, M. I., Masripah, M., & Holis, A. (2025). Kemampuan menyimak sebagai pondasi pengembangan keterampilan berbahasa siswa kelas rendah pada mata pelajaran bahasa indonesia. *Jurnal pendidikan dan pembelajaran indonesia (JPPI)*, *5*(1), 439-448.

[3] Sari, R., & Kholifah, S. (2025). Dampak Rekonstruksi Kurikulum Bahasa Arab Terhadap Peningkatan Kompetensi Guru Dalam Pembelajaran Istima'. *Al-Lahjah: Jurnal Pendidikan, Bahasa Arab, Dan Kajian Linguistik Arab*, *8*(1), 900-908.

[4] Graham, S., & Zhang, P. A. (2024). Learning vocabulary through listening: The role of strategy use and linguistic proficiency. *AILA Review*, *37*(2), 241-265.

[5] Hamidah, H., & Marsiah, M. (2020). Pembelajaran maharah al-istima'dengan memanfaatkan media youtube: problematika dan solusi. *Al-Ta'rib: Jurnal Ilmiah Program Studi Pendidikan Bahasa Arab IAIN Palangka Raya*, *8*(2), 147-160.

providing adaptive and authentic language input.[6] While these developments are theoretically promising, their practical integration into Arabic language instruction remains limited and often disconnected from pedagogical frameworks that emphasize interaction and learner engagement. This gap suggests that technological innovation alone is insufficient without corresponding instructional strategies that actively engage learners in processing input through meaningful interaction.

Within the framework of communicative language teaching, social interaction plays a central role in second language acquisition, particularly through the process of negotiation of meaning among learners.[7] In this regard, collaborative learning approaches such as Focus Group Discussion (FGD) have been shown to enhance learner participation and facilitate shared meaning-making processes. However, while FGD promotes interaction, it does not inherently provide structured communicative input unless it is integrated with context-rich language activities.

Previous studies have explored the role of interactive and collaborative approaches in second and foreign language learning. Research by Utari et al., demonstrated that roleplay effectively improves learners' speaking ability by increasing participation and communicative engagement.[8] Similar findings were reported by Mulyadi et al., who found that technology-enhanced task-based learning significantly improved learners' listening and speaking performance through interactive communicative activities.[9] In addition, Ahangari and Heidari revealed that collaborative listening activities contributed positively to learners' listening comprehension and vocabulary acquisition. These studies indicate that interaction-based learning can promote more active

[6] Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society*, *26*(1), 112-131.

[7] Dao, P. (2024). *Learner engagement in online second language classrooms*. London: Palgrave Macmillan.

[8] Utari, R. F., May, A., & Hikmah, H. (2024). Pengembangan Metode *Role Play* Untuk Meningkatkan Kemampuan Berbicara Pada Mata Pelajaran Bahasa Arab. *INTIFA: Journal of Education and Language*, *1*(2).

[9] Mulyadi, D., Wijayatiningsih, T. D., Singh, C. K. S., & Prastikawati, E. F. (2021). Effects of technology-enhanced task-based language teaching on learners' listening comprehension and speaking performance. *International Journal of Instruction, 14*(3), 717–736.

language processing and learner engagement.[10] However, most of these studies primarily focus on productive skills, particularly speaking, while the role of interaction in supporting receptive skills such as listening remains underexplored.

Furthermore, studies on listening comprehension emphasize that successful listening is not merely determined by exposure to language input, but also by learners' cognitive and metacognitive engagement during the listening process. Wang and MacIntyre highlighted the importance of metacognitive awareness and emotional engagement in second language listening comprehension.[11] Meanwhile, Tai and Chen found that immersive and interactive learning environments could enhance listening comprehension by providing contextualized and authentic input.[12] Similarly, Zhang and Zou argued that multimedia and interactive input facilitate deeper language processing in second language learning.[13] Despite these findings, previous studies have rarely integrated roleplay and collaborative discussion within a unified instructional framework specifically designed for maharah istima' in Arabic language learning. In addition, the integration of interactive pedagogy with emerging technologies such as AI remains limited in Arabic listening instruction contexts. Therefore, this study attempts to address these gaps by integrating roleplay and Focus Group Discussion (FGD) as an interaction-based approach to improve students' maharah istima'.

This separation reveals a critical pedagogical gap. Roleplay provides contextualized and meaningful input but is rarely designed to optimize listening processes, while FGD facilitates interaction and reflection but lacks structured communicative stimuli. As a result, neither approach independently addresses the dual demands of listening development, which require both meaningful input exposure and active

---

[10] Ahangari, S., & Heidari, N. (2021). The impact of collaborative listening to podcasts on high school learners' listening comprehension and vocabulary learning. *System, 101*, 102588.

[11] Wang, L., & MacIntyre, P. D. (2021). Second language listening comprehension: The role of anxiety and enjoyment in listening metacognitive awareness. *Studies in Second Language Learning and Teaching, 11*(4), 491–515.

[12] Tai, T.-Y., & Chen, H.-H. (2021). The impact of immersive virtual reality on EFL learners' listening comprehension. *Journal of Educational Computing Research, 59*(7), 1272–1293.

[13] Zhang, R., & Zou, D. (2022). A state-of-the-art review of the modes and effectiveness of multimedia input for second and foreign language learning. *Computer Assisted Language Learning, 35*(9), 2790–2816.

processing through interaction. Therefore, the issue is not merely the absence of interactive methods, but the lack of an integrated instructional design that systematically connects input, interaction, and cognitive processing.

Based on this gap, this study proposes the integration of roleplay and Focus Group Discussion (FGD) as a unified pedagogical approach to enhance *maharah istima'*. This integration is not intended as a simple combination of two methods, but as a structured strategy in which roleplay functions as a source of contextualized and comprehensible input, while FGD serves as a mechanism for negotiating meaning and deepening comprehension through collaborative discussion.

Accordingly, this study aims to (1) analyze the implementation of the integrated roleplay and FGD approach in teaching *maharah istima'*, and (2) examine its effect on students' listening ability. The novelty of this study lies in repositioning roleplay from a method primarily associated with productive skills into an integrative approach that simultaneously supports receptive and productive language processes. This reconceptualization is theoretically significant as it challenges the conventional separation between listening and speaking in language pedagogy, and empirically contributes to the development of interaction based instructional models for Arabic language learning.

Method

This study employed a quantitative approach using a quasi-experimental design, specifically a non-equivalent control group design. This design was selected due to practical constraints in educational settings where random assignment is not feasible.[14] Despite the absence of full randomization, this design enables a systematic comparison of changes in students' abilities between the experimental and control groups, making it appropriate for examining the effect of instructional interventions in real classroom contexts.[15]

The study was conducted at Al-Irhaam Global Islamic School, Bandung, involving 20 fifth-grade students selected through purposive sampling. The relatively small sample size positions this study as a pilot

[14] Sukmadinata, N. S. (2012). Pendekatan Penelitian Pendidikan. Bandung: PT. Remaja Rosdakarya.

[15] Saputri, L., & Mardiati, M. (2025). DESAIN PENELITIAN QUASI EKSPERIMEN DALAM PENELITIAN PENDIDIKAN: KAJIAN PUSTAKA. *Jurnal Serunai Matematika*, *17*(2).

study aimed at exploring the initial effectiveness of integrating roleplay and Focus Group Discussion (FGD) within a limited classroom context. Therefore, the findings are not intended for broad generalization but rather to provide preliminary empirical evidence for future studies with larger and more diverse samples.

In its implementation, Class V Salman was designated as the experimental group, while Class V Abdurrahman served as the control group. Both groups were administered a pre-test (O1 and O3) to assess their initial listening abilities. The experimental group then received treatment in the form of instruction integrating roleplay and Focus Group Discussion (FGD) (X), while the control group received conventional audio-based instruction. After the intervention, both groups were given a post-test (O2 and O4) to measure changes in listening ability. The research design is presented in Table 1.

Table 1. *Non-Equivalent Control Group Design*

A	**O1**	**X**	**O2**
B	**O3**		**O4**

To strengthen internal validity, several control measures were implemented, including the use of comparable instructional materials, equal duration of treatment, and similar classroom conditions across both groups. In addition, the time interval between the pre-test and post-test was kept consistent to minimize potential threats such as maturation and history effects. The pre-test functioned as a baseline measure, allowing observed differences in outcomes to be more reliably attributed to the instructional treatment.

Furthermore, to ensure treatment fidelity, the instructional procedures in the experimental group were implemented based on a systematically designed learning scenario. Each session followed a consistent sequence: (1) introduction of learning objectives and stimulus materials, (2) roleplay performance in groups, (3) Focus Group Discussion (FGD) to analyze and interpret the dialogue, and (4) teacher-led reflection. This structured procedure was applied uniformly across all sessions to maintain consistency in the delivery of the treatment.

The primary instrument used in this study was a listening comprehension test designed to measure students' *maharah istima'*. The test consisted of 10 multiple-choice items assessing students' ability to identify main ideas, understand specific information, and infer meaning from spoken Arabic input. The items were developed based on the instructional materials used in the classroom and aligned with the learning objectives.

To ensure content validity, the instrument was reviewed and validated by an Arabic language teacher with expertise in teaching *maharah istima'*, focusing on the appropriateness of linguistic level, content relevance, and clarity of the items. Given the limited sample size and the exploratory nature of this study, the validation process relied on expert judgment rather than statistical validation. Although reliability testing was not conducted, the consistency of the instrument was maintained through alignment with instructional objectives and expert validation.

Data were analyzed using SPSS version 26. The analytical procedures included: (1) the Shapiro-Wilk test to assess data normality, selected due to the small sample size ($n < 50$), (2) Levene's test to examine the homogeneity of variance between groups, (3) paired sample t-test to determine the significance of differences between pre-test and post-test scores within each group, and (4) N-Gain analysis to evaluate the magnitude of learning improvement. These procedures were employed to ensure that the data met the assumptions of parametric testing and to provide both statistical and practical interpretations of the findings.

Results and Discussion

The researcher conducted a normality test to find out if the distribution of the collected data was normal. Pre-test and post-test results in the experimental and control classes were tested using the Shapiro-Wilk method. According to Ismail et al, this method is considered suitable for studies with a sample count below 50, the data is considered normally distributed if there is a significance value of more than 0.05, while a significance value of less than 0.05 indicates that the data is not normally distributed.[16]

[16] Ismail, M. L., Dinangsit, D., & Muhtar, T. PENGARUH METODE LATIHAN SHADOW TERHADAP PUKULAN BACKHAND DRIVE DALAM TENIS MEJA. SpoRTIVE, 3(1), 361-370.

Tests of Normality

	Kelas	Kolmogorov-Smirnov[a] Statistic	df	Sig.	Shapiro-Wilk Statistic	df	Sig.
Hasil	Pretest Kontrol (Konvensional)	.188	10	.200*	.937	10	.516
	Postest Kontrol (Konvensional)	.240	10	.107	.886	10	.152
	Pretest Eksperimen (RP & FGD)	.203	10	.200*	.916	10	.321
	Postest Eksperimen (RP & FGD)	.286	10	.020	.885	10	.149

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Figure 1. Normality results *of pre-test* and *post-test* control and experimental classes

The Shapiro-Wilk test results indicate that all significance values exceeded the threshold of 0.05, suggesting that there is no statistically significant evidence to reject the assumption of normality. However, this does not confirm that the data are perfectly normally distributed, but rather indicates that no strong deviation from normality was detected. In addition to the p-values, the Shapiro-Wilk statistics (W) ranged from 0.885 to 0.937 across all groups. These values, being relatively close to 1, suggest that the data distributions approximate normality, although slight deviations may be present, particularly in the post-test scores where the W values are comparatively lower.

Furthermore, the consistency of both p-values and W statistics across the experimental and control groups indicates that the data distributions are relatively comparable in shape. This supports the validity of subsequent comparisons between groups. However, it should be noted that the sample size in this study is relatively small (n = 10 per group), which may limit the sensitivity of the Shapiro-Wilk test in detecting deviations from normality. Therefore, the normality assumption is considered sufficiently met for parametric analysis, but the results should be interpreted with caution.

Test of Homogeneity of Variance

		Levene Statistic	df1	df2	Sig.
Hasil RP & FGD	Based on Mean	.051	1	18	.825
	Based on Median	.106	1	18	.749
	Based on Median and with adjusted df	.106	1	17.978	.749
	Based on trimmed mean	.040	1	18	.843

Figure 2. Homogeneity test results

Before making a comparison between two or more groups, a variance homogeneity test must be carried out to ensure that the differences that

occur are not caused by inhomogeneity of basic data between groups.[17] The homogeneity of variance was examined using Levene's test based on multiple estimators, including mean, median, and trimmed mean. The results indicate that all significance values ranged from 0.749 to 0.843, exceeding the threshold of 0.05. This suggests that there is no statistically significant difference in variance between the experimental and control groups.

Beyond the significance values, the Levene statistics themselves were notably low (ranging from 0.040 to 0.106), indicating that the absolute differences in variance between groups were minimal. This implies that the variability of scores across groups is not only statistically comparable but also practically similar.

Furthermore, the consistency of results across different Levene test estimators (mean, median, and trimmed mean) demonstrates that the homogeneity assumption is robust, even when accounting for potential deviations such as non-normality or outliers. This strengthens the reliability of the variance equality assumption.

Given that both normality and homogeneity assumptions are satisfied, the data meet the required conditions for conducting parametric statistical analyses. However, considering the relatively small sample size, these findings should still be interpreted with caution.

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	PreEksperimen - PostEksperimen	-16.00000	11.73788	3.71184	-24.39677	-7.60323	-4.311	9	.002
Pair 2	PreKontrol - PostKontrol	-10.00000	17.63834	5.57773	-22.61771	2.61771	-1.793	9	.107

Figure 3. Paired sample t-test *results*

The paired sample t-test results reveal different patterns of improvement between the experimental and control groups. In the experimental group, the mean difference between pre-test and post-test scores was -16.00, indicating a substantial increase in students' listening performance. The standard deviation of 11.73 suggests that the improvement was relatively consistent among participants. The t-value of -4.311 reflects a strong effect, and the 95% confidence interval [-

[17] Usmadi, U. (2020). Pengujian persyaratan analisis (Uji homogenitas dan uji normalitas). Inovasi Pendidikan, 7(1).

24.39, -7.60] does not cross zero, indicating that the observed improvement is both statistically robust and consistently positive.

In contrast, the control group showed a mean difference of -10.00, indicating a moderate increase in scores. However, the higher standard deviation (17.63) suggests greater variability in students' performance gains. The t-value of -1.793 indicates a weaker effect, and the 95% confidence interval [-22.61, 2.61] includes zero, implying that the improvement is not consistent and may not be reliably distinguished from no effect.

Comparatively, while both groups experienced score improvements, the experimental group demonstrated not only a greater magnitude of gain but also a more stable and consistent pattern of improvement. This suggests that the integration of roleplay and Focus Group Discussion (FGD) contributes not only to higher achievement but also to more uniform learning outcomes among students.

However, given the relatively small sample size, these findings should be interpreted cautiously, and further studies with larger samples are recommended to confirm the observed patterns.

The N-Gain Score test is used to determine the effectiveness of the treatment given on student learning outcomes. After getting the *results of the pre-test* and *post-test*. The researcher tried to measure the increase in the impact *of the Roleplay* and FGD methods on the learning *of Maharah Istima'*, this is to determine the effectiveness of these methods.

The N-Gain formula is:

$$N\text{-}Gain = \frac{Post\ Test \text{ Score} - Pre\ Test \text{ Score}}{\text{Ideal Score} - Pre\ Test \text{ Score}}$$

Kategori Tafsiran Efektivitas N-Gain

Presentase (%)	Tafsiran
< 40	Tidak Efektif
40 – 55	Kurang Efektif
56 – 75	Cukup Efektif
> 76	Efektif

Sumber: Hake,R.R, 1999

Figure 4. Effectiveness criteria table image

And the following are the results of the N-Gain Score Test:

Descriptives

	Kelas			Statistic	Std. Error
N_GainPersen	Eksperimen	Mean		62.9630	7.42128
		95% Confidence Interval for Mean	Lower Bound	45.8495	
			Upper Bound	80.0765	
		5% Trimmed Mean		62.1811	
		Median		50.0000	
		Variance		495.679	
		Std. Deviation		22.26385	
		Minimum		40.00	
		Maximum		100.00	
		Range		60.00	
		Interquartile Range		33.33	
		Skewness		1.177	.717
		Kurtosis		-.042	1.400
	Kontrol	Mean		15.3439	19.06770
		95% Confidence Interval for Mean	Lower Bound	-28.6263	
			Upper Bound	59.3141	
		5% Trimmed Mean		18.6361	
		Median		25.0000	
		Variance		3272.195	
		Std. Deviation		57.20311	
		Minimum		-100.00	
		Maximum		71.43	
		Range		171.43	
		Interquartile Range		83.33	
		Skewness		-1.220	.717
		Kurtosis		.808	1.400

Figure 5. N-Gain Score Test Results

The N-Gain Score test is used to measure the level of effectiveness of treatment in improving student learning outcomes. Based on the calculations, an average N-Gain value of 0.63 (63%) was obtained in the experimental class. Referring to the criteria of Hake, the value is included in the medium category (quite effective). The N-Gain analysis reveals a substantial difference in learning improvement between the experimental and control groups.

The experimental group achieved a mean N-Gain score of 62.96, which falls into the moderate effectiveness category. The standard deviation of 22.26 indicates a relatively consistent improvement among students. The confidence interval [45.85, 80.08] remains entirely positive, suggesting that the observed gains are stable and reliably above zero.

In contrast, the control group obtained a mean N-Gain score of 15.34, indicating low effectiveness. However, the standard deviation is notably high (57.20), reflecting substantial variability in students' performance. The confidence interval [-28.63, 59.31] crosses zero, suggesting that the improvement is inconsistent and may not be reliably distinguished from no improvement.

Further analysis of data distribution shows that the experimental group has a positively skewed distribution (skewness = 1.177), indicating that most students experienced moderate to high gains, with a few achieving very high improvement. Conversely, the control group

exhibits a negatively skewed distribution (skewness = -1.220), indicating that a considerable number of students experienced low or even negative gains, as reflected in the minimum value of -100.

Additionally, the range and interquartile range of the control group are substantially larger than those of the experimental group, indicating a wider dispersion of scores and less consistent learning outcomes. Overall, these findings suggest that the integration of roleplay and Focus Group Discussion (FGD) not only enhances the magnitude of learning gains but also contributes to more stable and equitable learning outcomes across students.

Discussion

The findings of this study show that the improvement of maharah istima' ability in the experimental class is not only quantitative, but also reflects changes in the way students process language input.[18] Compared to control classes that rely on a passive audio-based approach, the integration of roleplay methods and Focus Group Discussion (FGD) provides a more interactive learning experience, thus encouraging students' active involvement in the language interpretation process. This finding supports previous studies emphasizing that collaborative and interactive listening activities facilitate deeper comprehension and more meaningful processing of language input.[19]

Theoretically, these findings are in line with the concept *of comprehensible input* put forward by Stephen Krashen, but at the same time expand on it. If in the classical view input is understood as something that is accepted by the learner, then the results of this study show that the effectiveness of input is highly dependent on the process of social interaction that accompanies it.[20] Language understanding is not only formed from exposure to inputs, but also from the negotiation of meaning that occurs through discussion and communication simulations. These findings are in line with the study of interactions in

[18] Shamsi, E., & Bozorgian, H. (2024). Collaborative listening using multimedia through metacognitive instruction: A case study with less-skilled and more-skilled EFL learners. *Asian-Pacific Journal of Second and Foreign Language Education, 9*(25).

[19] Lam, D. M. K., & Goh, C. C. M. (2026). Listening beyond comprehension: Towards a taxonomy of second language interactive listening skills. *RELC Journal*.

[20] Krashen, S. D. (1985). The input hypothesis: Issues and implications.

second language acquisition which emphasizes the importance of negotiation of meaning in improving language understanding.[21]

The integration of roleplay and FGD in this study plays a role in activating more complex cognitive mechanisms, such as selective attention, meaning inference, and information elaboration. Roleplay provides an authentic communication context, while FGD serve as reflective spaces to clarify and deepen understanding. The combination of the two creates a learning cycle that is not only receptive, but also constructive.[22] This is in line with the findings of Norbert Schmitt and I.S.P. Nation who stated that language learning that involves active interaction can improve the quality of input processing in more depth.[23]

These findings contribute to the development of the pedagogical function of the roleplay method. So far, roleplay has often been positioned as a strategy to improve maharah kalam or productive language skills.[24] However, the results of this study show that when integrated with collaborative approaches such as Focus Group Discussion (FGD), roleplay is also effective in developing receptive skills, particularly maharah istima'. This finding is consistent with previous studies emphasizing that drama and roleplay-based instruction can facilitate integrated language learning processes through interaction, contextual communication, and collaborative meaning making.[25] Thus, this study confirms that roleplay has an integrative function in language learning, not only as a medium for language production but also as a means of strengthening listening comprehension and language processing.[26]

---

[21] Sato, M. (2017). Interaction mindsets, interactional behaviors, and L2 development: An affective-social-cognitive model. *Language learning*, *67*(2), 249-283.

[22] Vandergrift, L., & Cross, J. (2021). The role of metacognitive awareness in second language listening comprehension. *Computers & Education, 172*, 104259.

[23] Newton, J. M., & Nation, I. S. (2020). *Teaching ESL/EFL listening and speaking*. Routledge.

[24] Chen, Y. L. (2022). The effects of role-play and collaborative interaction on EFL learners' communicative competence. *Education Sciences, 12*(6), 376.

[25] Sato, M., & Ballinger, S. (2022). *Peer interaction and second language learning: Pedagogical potential and research agenda*. John Benjamins Publishing Company.

[26] Galante, A., & Thomson, R. I. (2017). The effectiveness of drama as an instructional approach for the development of second language oral fluency, comprehensibility, and accentedness. *TESOL Quarterly, 51*(1), 115–142

In the perspective of 21st century learning, these results also confirm the importance of an approach that emphasizes student interaction, collaboration, and active engagement, as stated by Bernie Trilling and Charles Fadel.[27] Roleplay activities encourage contextual communication, while FGDs strengthen critical thinking skills through a discussion process, making them both relevant to the demands of contemporary pedagogy.

Related to technological developments, especially Artificial Intelligence (AI), this study shows the potential for complementary integration. AI can be used as a supporting tool, for example in providing adaptive audio material, developing roleplay scenarios, and facilitating discussions. This is in line with the findings of Dirk Ifenthaler and Sabine Zawacki-Richter, who affirm that AI in education plays a role as a learning support system that is able to increase personalization and learning effectiveness.[28] Thus, AI does not replace the role of learning methods, but enriches the quality of learning inputs and interactions.

However, this study has several limitations, particularly the relatively small sample size and the research scope, which was limited to a single institution. Consequently, the findings should be interpreted cautiously and cannot be generalized broadly across different educational contexts. In addition, the study primarily focused on examining the effectiveness of roleplay and Focus Group Discussion (FGD) within a conventional classroom setting, without systematically integrating digital or AI-assisted learning environments. Therefore, future research is recommended to involve larger and more diverse samples, as well as to explore the integration of interactive learning strategies with emerging technologies such as AI-supported language learning tools in order to examine their potential contributions to maharah istima' development more comprehensively.[29]

---

[27] Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times*. John Wiley & Sons

[28] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators?. *International journal of educational technology in higher education*, *16*(1), 39

[29] Viberg, O., Khalil, M., & Baars, M. (2024). Self-regulated learning and AI in education: A systematic review and future directions. *Computers and Education: Artificial Intelligence, 5*, 100172.

Conclusion

Based on the results of the above research, it can be concluded that the application of the Roleplay and Focus Group Discussion (FGD) methods in the learning of Maharah Istima' has a significant influence and shows higher effectiveness compared to conventional methods. This is proven through the results of the N-Gain Score test, where the Roleplay and FGD methods obtained a score of 63% which is included in the category of quite effective. In addition, there is a real difference between the pre-test and post-test results. In the class with the application of the Roleplay and FGD methods, the average pre-test score of 72.0 increased to 88.0 in the post-test, with a difference of 16.0. Meanwhile, in the class with the conventional method, the average pre-test score of 71.0 increased to 81.0, with a difference of 10.0.

The findings of this study empirically answer the research questions that have been formulated previously. First, related to the application of methods, the integration of roleplay and Focus Group Discussion (FGD) has been proven to be carried out systematically through learning stages involving communication simulation activities, group discussions, and reflection, thereby creating an interactive and participatory learning environment. Second, related to the effectiveness of the method, the results of statistical analysis showed a significant increase in the ability of students in the experimental class, which was shown by the significance value of *the paired sample t-test* ($p < 0.05$) and the N-Gain value of 0.63 in the medium category. In addition, the improvement that occurred in the experimental class was also higher than in the control class, which shows that the method applied has a real contribution to improving students' listening skills.

Based on these findings, it can be concluded that the integration of roleplay and FGD methods is an effective learning strategy in improving maharah istima', especially in the context of learning Arabic as a foreign language. Therefore, this method is recommended to be applied by Arabic teachers, especially in educational environments that have limited learning media, because it is able to create a more active, interactive, and contextual learning atmosphere.

In addition, as a further development, this learning approach has the potential to be combined with the use of Artificial Intelligence (AI)-based technology, such as *text-to-speech* or *AI-based language tools*, to enrich the variety of language inputs and improve the quality of learning

Maharah Istima' [30]. For future researchers, it is recommended to develop research on the integration of Roleplay methods, FGD, and AI technology in Arabic language learning more broadly, involving a more diverse context, education level, and sample size in order to obtain more comprehensive results.

[30] Teng, M. F., & Yue, M. (2023). Metacognitive writing strategies, critical thinking skills, and academic writing performance: A structural equation modeling approach. *Metacognition and Learning*, *18*(1), 237-260.

REFERENCES

Ahangari, S., & Heidari, N. (2021). The impact of collaborative listening to podcasts on high school learners' listening comprehension and vocabulary learning. *System, 101*, 102588.

Ali, M. (2020). Pembelajaran bahasa indonesia dan sastra (basastra) di sekolah dasar. *Pernik*, *3*(1), 35-44.

Chen, Y. L. (2022). The effects of role-play and collaborative interaction on EFL learners' communicative competence. *Education Sciences, 12*(6), 376.

Dao, P. (2024). *Learner engagement in online second language classrooms*. London: Palgrave Macmillan.

Erlina, A., Sukardi, S., & Wahidah, A. (2024). Pengaruh Model Pembelajaran FGD (*Focus Group Discussion*) Berbantuan Poster Terhadap Partisipasi Belajar Siswa Pada Mata Pelajaran Sosiologi. *Jurnal Ilmiah Profesi Pendidikan*, *9*(4), 2428-2435.

Galante, A., & Thomson, R. I. (2017). The effectiveness of drama as an instructional approach for the development of second language oral fluency, comprehensibility, and accentedness. *TESOL Quarterly, 51*(1), 115–142

Graham, S., & Zhang, P. A. (2024). Learning vocabulary through listening: The role of strategy use and linguistic proficiency. *AILA Review*, *37*(2), 241-265.

Hamidah, H., & Marsiah, M. (2020). Pembelajaran maharah al-istima'dengan memanfaatkan media youtube: problematika dan solusi. *Al-Ta'rib: Jurnal Ilmiah Program Studi Pendidikan Bahasa Arab IAIN Palangka Raya*, *8*(2), 147-160.

Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2023). Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society*, *26*(1), 112-131.

Ismail, M. L., Dinangsit, D., & Muhtar, T. PENGARUH METODE LATIHAN SHADOW TERHADAP PUKULAN BACKHAND DRIVE DALAM TENIS MEJA. SpoRTIVE, 3(1), 361-370.

Krashen, S. D. (1985). The input hypothesis: Issues and implications.

Lam, D. M. K., & Goh, C. C. M. (2026). Listening beyond comprehension: Towards a taxonomy of second language interactive listening skills. *RELC Journal*.

Mulyadi, D., Wijayatiningsih, T. D., Singh, C. K. S., & Prastikawati, E. F. (2021). Effects of technology-enhanced task-based language teaching on learners' listening comprehension and speaking

performance. *International Journal of Instruction, 14*(3), 717–736.

Newton, J. M., & Nation, I. S. (2020). *Teaching ESL/EFL listening and speaking*. Routledge.

Saputri, L., & Mardiati, M. (2025). DESAIN PENELITIAN QUASI EKSPERIMEN DALAM PENELITIAN PENDIDIKAN: KAJIAN PUSTAKA. *Jurnal Serunai Matematika*, *17*(2).

Sari, R., & Kholifah, S. (2025). Dampak Rekonstruksi Kurikulum Bahasa Arab Terhadap Peningkatan Kompetensi Guru Dalam Pembelajaran Istima'. *Al-Lahjah: Jurnal Pendidikan, Bahasa Arab, Dan Kajian Linguistik Arab*, *8*(1), 900-908.

Sato, M. (2017). Interaction mindsets, interactional behaviors, and L2 development: An affective-social-cognitive model. *Language learning*, *67*(2), 249-283.

Sato, M., & Ballinger, S. (2022). *Peer interaction and second language learning: Pedagogical potential and research agenda*. John Benjamins Publishing Company.

Shamsi, E., & Bozorgian, H. (2024). Collaborative listening using multimedia through metacognitive instruction: A case study with less-skilled and more-skilled EFL learners. *Asian-Pacific Journal of Second and Foreign Language Education, 9*(25).

Sukmadinata, N. S. (2012). Pendekatan Penelitian Pendidikan. Bandung: PT. Remaja Rosdakarya.

Tai, T.-Y., & Chen, H.-H. (2021). The impact of immersive virtual reality on EFL learners' listening comprehension. *Journal of Educational Computing Research, 59*(7), 1272–1293.

Teng, M. F., & Yue, M. (2023). Metacognitive writing strategies, critical thinking skills, and academic writing performance: A structural equation modeling approach. *Metacognition and Learning*, *18*(1), 237-260.

Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times*. John Wiley & Sons.

Ubaidillah, M. I., Masripah, M., & Holis, A. (2025). Kemampuan menyimak sebagai pondasi pengembangan keterampilan berbahasa siswa kelas rendah pada mata pelajaran bahasa indonesia. *Jurnal pendidikan dan pembelajaran indonesia (JPPI)*, *5*(1), 439-448.

Usmadi, U. (2020). Pengujian persyaratan analisis (Uji homogenitas dan uji normalitas). Inovasi Pendidikan, 7(1).

Utari, R. F., May, A., & Hikmah, H. (2024). Pengembangan Metode *Role Play* Untuk Meningkatkan Kemampuan Berbicara Pada Mata Pelajaran Bahasa Arab. *INTIFA: Journal of Education and Language*, *1*(2).

Vandergrift, L., & Cross, J. (2021). The role of metacognitive awareness in second language listening comprehension. *Computers & Education, 172*, 104259.

Viberg, O., Khalil, M., & Baars, M. (2024). Self-regulated learning and AI in education: A systematic review and future directions. *Computers and Education: Artificial Intelligence, 5*, 100172.

Wang, L., & MacIntyre, P. D. (2021). Second language listening comprehension: The role of anxiety and enjoyment in listening metacognitive awareness. *Studies in Second Language Learning and Teaching, 11*(4), 491–515.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–where are the educators?. *International journal of educational technology in higher education*, *16*(1), 39.

Zhang, R., & Zou, D. (2022). A state-of-the-art review of the modes and effectiveness of multimedia input for second and foreign language learning. *Computer Assisted Language Learning, 35*(9), 2790–2816.