PIJAR: Jurnal Pendidikan dan Pengajaran
Vol. 4, No.2 April 2026), pp.1090-1102
E-ISSN 2964-7460: 10.58540/pijar.v4i2.2043

THE EFFECTIVENESS OF THE WINDOW SHOPPING STRATEGY IN IMPROVING STUDENTS' LEARNING OUTCOMES IN DURŪS AL-LUGHAH AT PONDOK PESANTREN MIFTAHUL HUDA SUKO IN THE ACADEMIC YEAR 2025-2026

Fitri Setyo Rini[1], Ifa Rodifah Nur[2], Vava Ulil Albab[3]

[1,2,3]Department of Arabic Language Education, Faculty of Tarbiyah, University of Darussalam Gontor, Ponorogo, Indonesia; fitrisetyo2@unida.gontor.ac.id

ARTICLE INFO

Article history:

Received 2026-04-15
Revised 2026-04-27
Accepted 2026-04-30

ABSTRAK

Pembelajaran bahasa Arab pada kelas dua Pondok Pesantren Miftahul Huda Suko masih didominasi oleh metode ceramah konvensional yang bersifat pasif dan berpusat pada guru sehingga banyak siswa kesulitan memahami materi dan kurang termotivasi. Penelitian ini bertujuan mengetahui efektivitas strategi Window Shopping dalam meningkatkan hasil belajar siswa pada mata pelajaran Durūs al-Lughah. Penelitian menggunakan metode kuantitatif kuasi-eksperimen dengan desain time series yang melibatkan 25 siswa yang dipilih melalui purposive sampling. Data dikumpulkan melalui tes hasil belajar yang diberikan empat kali sebelum penerapan strategi (pretest) dan empat kali setelahnya (posttest), lalu dianalisis menggunakan SPSS 27 melalui uji Friedman, ANOVA pengukuran berulang dengan koreksi Greenhouse-Geisser, dan uji t sampel berpasangan. Hasil menunjukkan nilai rata-rata meningkat dari 63,60 pada pretest menjadi 82,90 pada posttest dengan selisih 19,30 poin. Nilai signifikansi (Sig. = 0,001 < 0,05) menunjukkan peningkatan yang signifikan secara statistik, dan ukuran efek (Cohen's d = 2,095) berada pada kategori sangat besar. Standar deviasi juga menurun setelah perlakuan, menandakan hasil belajar yang lebih homogen. Dengan demikian, strategi Window Shopping terbukti efektif dalam pembelajaran bahasa Arab karena meningkatkan hasil belajar sekaligus menumbuhkan keaktifan, kerja sama, dan kemampuan berpikir kritis siswa.

Kata Kunci: Strategi Window Shopping, Pembelajaran Bahasa Arab, Hasil Belajar, Desain Time Series, Pondok Pesantren.

ABSTRACT

Arabic instruction in the second grade of Pondok Pesantren Miftahul Huda Suko is still dominated by the conventional lecture method, which is passive and teacher-centered, causing many students to struggle to understand the material and to lack motivation. This study aims to determine the effectiveness of the Window Shopping strategy in improving students' learning outcomes in Durūs al-Lughah (Arabic language lessons). It employed a quantitative quasi-experimental method with a time series design, involving 25 second-grade students selected through purposive sampling. Data were

collected through achievement tests administered four times before the strategy (pretest) and four times after it (posttest), and analyzed with SPSS 27 using the Friedman test, repeated-measures ANOVA with the Greenhouse-Geisser correction, and a paired-sample t-test. The results showed that the mean score increased from 63.60 in the pretest to 82.90 in the posttest, a gain of 19.30 points. The significance value (Sig. = 0.001 < 0.05) indicated a statistically significant improvement, and the effect size (Cohen's d = 2.095) fell within the very large category. The standard deviation also declined after the treatment, indicating more homogeneous outcomes. Thus, the Window Shopping strategy is proven effective in Arabic language learning, as it improves learning outcomes while fostering students' activeness, cooperation, and critical thinking skills.

Keyword: Window Shopping strategy, Arabic language learning, learning outcomes, time series design, Islamic boarding school.

This is an open access article under the CC BY license.

Corresponding Author:
Fitri Setyo Rini
University of Darussalam Gontor, Ponorogo, Indonesia; fitrisetyo2@unida.gontor.ac.id

1. Introduction

Arabic is one of the world's major languages, possessing great civilizational, cultural, and religious value; it is the language of the Qur'an, the Prophetic traditions, and the primary sources of the Islamic heritage. Az-Zuhri (2007) emphasized that Arabic is not merely a means of communication, but a key to understanding the Islamic heritage and the religious sciences, which makes learning it a necessity for every Muslim. Beyond the Muslim world, Arabic is also one of the official languages of the United Nations and is spoken by hundreds of millions of people, which underlines its global and academic relevance. For learners in Indonesian pesantren, however, Arabic is above all the gateway to the classical religious sciences tafsīr, ḥadīth, fiqh, and balāghah so that mastery of the language is inseparable from the deeper aims of religious education and the formation of a well-rounded Muslim personality.

The curriculum of Arabic language teaching has developed considerably in recent decades, particularly with the shift toward communicative and competency-based instruction that aims to enable students to use Arabic in their daily lives (Rahman, 2020). With the emergence of modern educational technology, the methods of teaching Arabic have also become more varied, relying on digital media and interactive applications. Nevertheless, classroom reality shows that many students in Islamic boarding schools still suffer from weak language competence, which indicates that technology alone cannot guarantee successful learning. Hidayah (2018) noted that the absence of variety in teaching methods negatively affects students' motivation to learn, and Mulyasa (2016) pointed out that adherence to

traditional techniques such as rote instruction and direct lecturing weakens students' ability to interact with the learning content, making them feel bored and isolated from the learning environment. Modern instructional strategies therefore play an essential role in improving Arabic learning, especially when they reflect active methods that stimulate students to become active learners rather than passive recipients.

The theoretical foundation of this study rests on the paradigm of active and cooperative learning, which emerged as a response to the limitations of teacher-centered instruction. Constructivist learning theory holds that knowledge is not transmitted intact from teacher to student but is actively constructed by the learner through experience and social interaction (Creswell, 2012). Within this paradigm, cooperative learning in which students work together in structured activities toward shared goals has been shown to raise achievement, motivation, and interpersonal skills across a wide range of subjects. The notion of the zone of proximal development further suggests that learners advance further when they interact with peers and more capable others than when they work in isolation. These principles are especially pertinent to language learning, where meaning is negotiated through communication and where exposure to varied input together with repeated productive practice is essential for the internalization of vocabulary and grammatical rules.

The Window Shopping strategy is a cooperative learning technique that operationalizes these principles in a concrete classroom routine. In this strategy, the learning material is distributed across several stations or "stalls" arranged around the classroom, and students organized in small groups circulate from one station to another as though browsing shop windows, observing the displayed content, discussing it, and recording what they find (Apriana, 2020; Nengsih, 2022). Typically one member of each group remains at the home station to explain the group's work to visitors, while the others move to inspect and evaluate the work of the other groups, after which the roles rotate so that every student both presents and observes. The strategy thus combines physical movement, visual presentation, peer explanation, and note-taking, engaging several channels of learning at once. Because students must present their own understanding and, at the same time, interrogate that of others, the technique naturally elicits the higher-order processes of analysis, comparison, and evaluation, while the informal, game-like atmosphere reduces the anxiety that often accompanies foreign-language use.

A growing body of Indonesian classroom research has documented the benefits of the Window Shopping strategy across subjects. Apriana (2020) and Rasyid (2021) reported that it improved students' social-studies achievement at the junior-secondary and primary levels respectively, while Farida (2023) found that it increased pupils' activeness and learning outcomes in primary school. In mathematics, Nengsih (2022) observed higher learning activity when the strategy was applied to solid-geometry material. Within Arabic language education specifically, Husna (2021) demonstrated a positive effect on students' speaking skills, Wardah (2021) showed that the strategy improved comprehension and application of grammar (naḥw), and Ibrahim et al. (2022) reported gains in students' confidence in using Arabic. Taken

together, these studies indicate that the strategy is broadly effective; yet most of them address skills other than the integrated language lessons (Durūs al-Lughah) that form the core of the pesantren curriculum, and few employ a repeated-measurement design capable of tracing the trajectory of improvement over time.

In the context of this study, learning outcomes are understood as the measurable result of the instructional process, spanning the cognitive domain (the comprehension, application, and analysis of linguistic material), the affective domain (attitudes and motivation toward the language), and the psychomotor domain (the productive use of the language in speaking and writing). Durūs al-Lughah, as an integrated subject, draws on all three domains at once, which makes it both a demanding area to teach and a sensitive indicator of the quality of instruction. An effective strategy for this subject must therefore do more than transmit information; it must engage students affectively and give them repeated opportunities to produce the language precisely the conditions that an active, cooperative approach is designed to create.

At Pondok Pesantren Miftahul Huda Suko, preliminary observation revealed a persistent gap between the aims of the Arabic curriculum and the students' actual attainment. In the second grade, instruction in Durūs al-Lughah was still dominated by lecture and memorization, with limited opportunity for students to use the language actively or to learn from one another. Many students' scores in earlier assessments did not exceed about 75 out of 100, and classroom engagement was low: students appeared passive, easily bored, and reluctant to participate. Asrori (2023) stressed that a positive, participatory classroom climate is a precondition for effective language learning, a condition that the prevailing teacher-centered routine at the school did not adequately provide. This situation motivated the search for a more active strategy capable of raising both achievement and engagement simultaneously.

Although the effectiveness of the Window Shopping strategy has been examined in several settings, a clear gap remains. Prior studies have largely relied on single pretest–posttest comparisons, which cannot show whether an observed improvement is stable or merely a momentary fluctuation, and few have focused on Durūs al-Lughah in the pesantren context. The present study addresses this gap by adopting a time-series design with four pretests and four posttests, allowing the pattern of change to be traced systematically before and after the intervention. Accordingly, the study is guided by the following question: to what extent is the Window Shopping strategy effective in improving second-grade students' learning outcomes in Durūs al-Lughah at Pondok Pesantren Miftahul Huda Suko, compared with their performance under conventional instruction?

The objective of this study is therefore to measure the effectiveness of the Window Shopping strategy in improving the learning outcomes of Durūs al-Lughah among second-grade students of Pondok Pesantren Miftahul Huda Suko in the academic year 2025-2026. Theoretically, the study contributes to the literature on cooperative learning by testing, through a repeated-measurement design, the proposition that structured social interaction enhances language-learning outcomes. Practically, it is expected to benefit students by

fostering activeness, cooperation, and critical thinking; to provide teachers with a concrete and replicable model for cooperation-based instruction; and to help the institution formulate more effective and innovative educational policies. The findings may likewise serve as a reference for future researchers seeking to adapt active-learning strategies to other language skills and educational levels.

2. Method

This study is a quantitative quasi-experimental study using a time series design. It examines the effect of the independent variable (X), namely the use of the Window Shopping strategy, on the dependent variable (Y), the students' learning outcomes in Durūs al-Lughah. The study was conducted in the second grade of Pondok Pesantren Miftahul Huda Suko, Kediri, East Java, in the academic year 2025-2026. This approach was chosen because it corresponds to the objective of the research, as it makes it possible to trace changes in students' learning outcomes over a certain period of time.

Research Design and Procedure

In the time series design, repeated measurements are taken on a single experimental group before and after the treatment. The design can be represented as O1 O2 O3 O4 X O5 O6 O7 O8, where O1–O4 are the four pretests, X is the treatment (the implementation of the Window Shopping strategy), and O5–O8 are the four posttests. The pretests were administered for four consecutive weeks to measure the students' level before the strategy was applied, while the four posttests traced the gradual development of their learning outcomes after the intervention.

The implementation began by determining the linguistic content to be taught, which was then divided into a number of learning stations; each station represented a part of the topic and could contain a short text, a picture, a stimulating question, or a simple activity. The students were asked to "shop by looking," moving from one station to another, observing the content, reading the information, and taking notes or answering the questions on their own worksheets. After the shopping round, the students returned to their seats and a whole-class discussion led by the teacher was organized, in which the students shared what they had observed and learned from each station, and the main points were elaborated according to their responses. During this process, the researcher observed the students' interaction, their interest in the content, and the ways in which they expressed their understanding.

Population and Sample

The population of this study comprised all second-grade students of Pondok Pesantren Miftahul Huda Suko in the academic year 2025-2026, numbering 33 students in one class (Sugiyono, 2019). The planned sample size was calculated using the Slovin formula with a margin of error of 10%, and the sample was then determined through purposive sampling, so that 25 students were selected as the single experimental class. Each of them took four pretests and four posttests, before and after the implementation of the Window Shopping strategy.

Instruments, Validity, and Data Analysis

The data were collected through tests, observation, and interviews. The test, a set of written questions on the topic of the verb (fi'l), was used to measure the students' learning outcomes in Durūs al-Lughah; observation was carried out to see how the learning process was conducted in the classroom; and an interview was held with the teacher of the subject. The empirical validity of the 17 test items was examined with SPSS 27, and all items were found to be valid (r-count > r-table). The reliability test produced a Cronbach's alpha of 0.875 > 0.60, so the instrument was declared reliable.

The data were analyzed through descriptive statistics, followed by the Shapiro-Wilk normality test, the Friedman test, Mauchly's test of sphericity with repeated-measures ANOVA using the Greenhouse-Geisser correction, and a paired-sample t-test, complemented by the calculation of the effect size using Cohen's d. The statistical hypotheses were formulated as follows: Ha, there is an effect of the Window Shopping strategy on improving the learning outcomes of Durūs al-Lughah; and Ho, there is no effect of the Window Shopping strategy on improving the learning outcomes of Durūs al-Lughah.

3. Result And Discussion

Result

In accordance with the time series design, four pretests were administered for four consecutive weeks before the implementation of the Window Shopping strategy, followed by four posttests after it. The mean scores of the students in the first to the fourth pretest were relatively stable in the range of 62–65, indicating that before the treatment the students' learning outcomes tended to be constant. After the strategy was implemented, the mean scores rose sharply and then continued to increase gradually until they reached about 87 in the fourth posttest. This pattern a sharp rise immediately after the intervention followed by a continuous, more moderate increase indicates that the Window Shopping strategy had a positive effect on the students' learning outcomes in Durūs al-Lughah, as illustrated in Figure 1 and Figure 2.

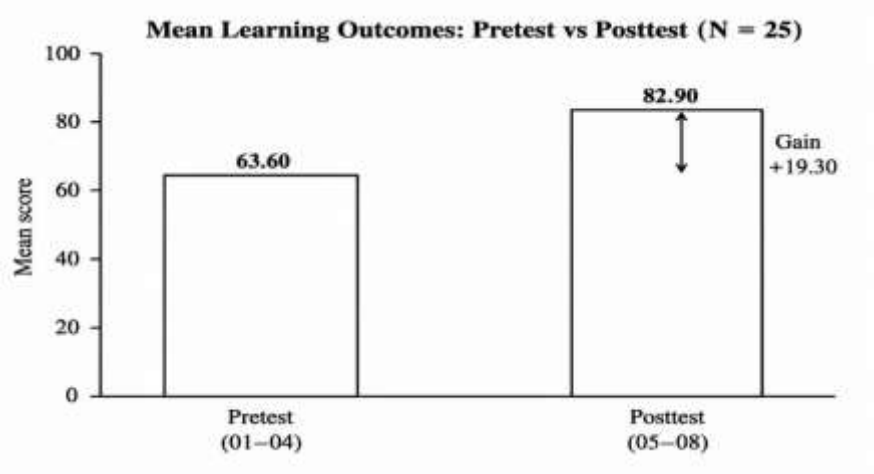

Figure 1. Mean learning outcomes before (pretest) and after (posttest) the Window Shopping strategy (N = 25).

The per-measurement descriptive statistics are presented in Table 1. Across the four pretests the means remained close to one another (65.00, 63.50, 63.00, and 62.90), confirming that the group was stable prior to the intervention. From the first posttest onward the means climbed steadily (78.00, 82.00, 84.60, and 87.00), so that the distance between the lowest posttest mean and the highest pretest mean already exceeded thirteen points. The consistency of the pretest scores strengthens the internal validity of the design, because it makes it unlikely that the subsequent rise was caused by maturation or by repeated testing alone rather than by the treatment.

Table 1. Descriptive Statistics of the Eight Measurements (O1–O8)

Measurement	Phase	N	Mean	Std. Deviation
O1	Pretest	25	65.00	11.83
O2	Pretest	25	63.50	11.40
O3	Pretest	25	63.00	11.10
O4	Pretest	25	62.90	10.95
O5	Posttest	25	78.00	8.90
O6	Posttest	25	82.00	7.60
O7	Posttest	25	84.60	6.80
O8	Posttest	25	87.00	6.10

Table 1 also shows a clear reduction in the standard deviation over the course of the study. The dispersion of the scores was highest in the first pretest (SD = 11.83) and declined steadily to 6.10 in the final posttest, as visualized in Figure 3. This means that the students' scores not only increased on average but also became more homogeneous, so that the gap between higher- and lower-achieving students narrowed after the strategy was applied. The strategy therefore appears to have benefited weaker students in particular, lifting the whole group toward a more uniform level of attainment.

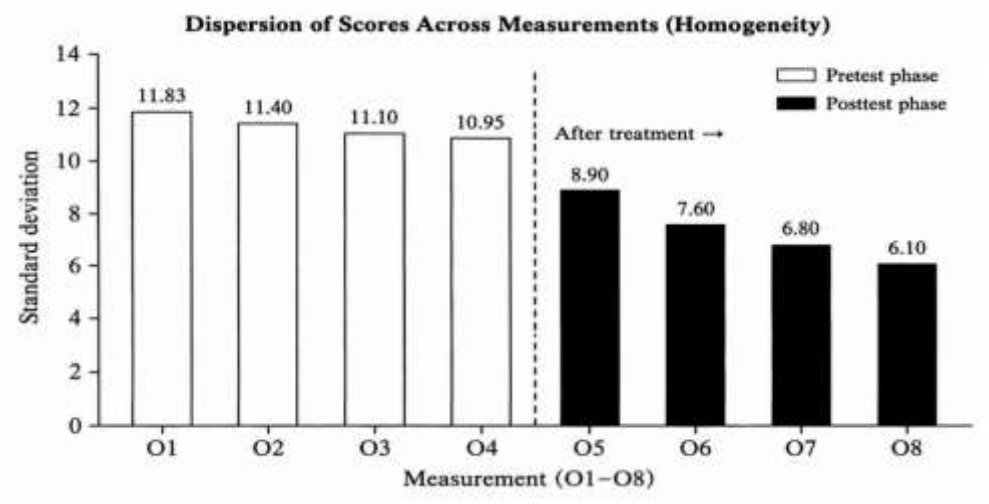

Figure 2. Dispersion (Standard Deviation) Of The Scores Across The Eight Measurements; The Spread Narrows After The Treatment, Indicating Greater Homogeneity

Considered as two phases, the descriptive comparison in Table 2 summarizes the overall movement. The overall mean of the pretest scores was 63.60, while the mean of the posttest scores rose to 82.90, an increase of 19.30 points, equivalent to a relative gain of roughly thirty percent over the pretest level. With a sample of 25 students (N = 25), this magnitude of change is substantial for a single class and provides the descriptive basis for the inferential tests that follow.

Table 2. Descriptive Comparison of the Pretest and Posttest Phases

Measurement phase	N	Mean	Note
Pretest (O1–O4)	25	63.60	Scores stable within the 62–65 range
Posttest (O5–O8)	25	82.90	Rose to ≈ 87 at the fourth posttest
Gain (posttest − pretest)	—	19.30	Statistically significant (p = 0.001)

Before the inferential analysis, the quality of the measurement instrument was confirmed, as summarized in Table 3. All 17 test items were empirically valid (r-count > r-table), and the reliability coefficient was high (Cronbach's alpha = 0.875, well above the 0.60 threshold). These values indicate that the differences observed between the pretest and posttest scores can be attributed to genuine changes in the students' ability rather than to measurement error or an unstable instrument.

Table 3. Validity and Reliability of the Test Instrument

Aspect	Result
Number of test items	17
Valid items	17 of 17 (r-count > r-table)
Reliability (Cronbach's alpha)	0.875 (> 0.60 → reliable)

The Shapiro-Wilk normality test indicated that the data were not normally distributed (Table 4); therefore, the non-parametric Friedman test was applied to the repeated measurements within the single group, and the parametric analyses were interpreted with the appropriate corrections. The Friedman test produced a Chi-Square value of 119.850 with df = 7 and a significance value of less than 0.001 < 0.05, so it can be concluded that there are significant differences among the eight measurements.

Table 4. Result of the Shapiro-Wilk Normality Test

Test	Result	Implication
Shapiro-Wilk	p < 0.05 (not normally distributed)	Non-parametric / corrected tests applied

Furthermore, Mauchly's test of sphericity produced a significance value of less than 0.001 < 0.05, meaning that the assumption of sphericity was not fulfilled; consequently, the repeated-measures ANOVA was interpreted through the Greenhouse-Geisser correction, which showed F = 54.587 with a significance value of less than 0.001, confirming that there are statistically significant differences among the measurement stages. The research hypothesis was then tested using a paired-sample t-test to analyze the difference in learning outcomes

before and after the implementation of the strategy. The test produced a Sig. (2-tailed) value of 0.001, which is smaller than the significance level of 0.05; accordingly, the null hypothesis (Ho) was rejected and the alternative hypothesis (Ha) was accepted. In addition, the effect size calculated using Cohen's d reached 2.095 in absolute value, which falls within the "very large" category, meaning that the intervention produced not only a statistically significant difference but also a practically substantial one. A complete summary of the inferential tests is presented in Table 5.

Table 5. Summary of the Inferential Statistical Analyses

Statistical test	Result	Decision
Friedman test	Chi-Square = 119.850; df = 7; p < 0.001	Significant differences among the 8 measurements
Mauchly's test of sphericity	p < 0.001	Sphericity not met → Greenhouse-Geisser used
Repeated-measures ANOVA (GG)	F = 54.587; p < 0.001	Significant differences among stages
Paired-sample t-test	Sig. (2-tailed) = 0.001 < 0.05	Ho rejected, Ha accepted
Effect size (Cohen's d)	2.095	Very large practical effect

Beyond the numerical results, the classroom observations conducted during the intervention documented a visible change in the learning atmosphere. During the shopping rounds the students moved purposefully between the stations, discussed the displayed material in their groups, and questioned one another about the meaning and use of the verbs under study. Learners who had previously been reluctant to speak were drawn into the activity by the need to explain their own station to visiting classmates, and the noise of the room was, for the most part, the noise of on-task discussion. These qualitative observations complement the test data by showing that the rise in scores was accompanied by a genuine increase in participation, cooperation, and willingness to use Arabic.

Taken together, the descriptive and inferential results converge on a single conclusion. The stable pretest baseline, the sharp and sustained rise across the four posttests, the narrowing standard deviation, the significant Friedman and repeated-measures ANOVA results, the significant paired-sample t-test, and the very large effect size all point in the same direction: the alternative hypothesis (Ha) is supported and the null hypothesis (Ho) is rejected. The convergence of several independent indicators, rather than reliance on a single test, lends additional confidence to the finding that the Window Shopping strategy improved the students' learning outcomes in Durūs al-Lughah.

Discussion

The results of this study confirm that the Window Shopping strategy significantly improved the students' learning outcomes in Durūs al-Lughah, answering the research question posed in the introduction. The rise of the mean score from 63.60 to 82.90, together with a paired-sample significance value of 0.001 and a very large effect size (Cohen's d = 2.095), demonstrates that the improvement is both statistically reliable and practically

meaningful. These findings are consistent with the theory of Gerlach and Ely (1980), which asserts that instructional strategies encompass organized activities between the teacher and the students to achieve learning objectives, and that active, interaction-stimulating strategies improve learning outcomes.

Several mechanisms can explain why the strategy was effective. First, the Window Shopping routine keeps students physically and cognitively active. Instead of listening passively, learners move between stations, read, discuss, and record what they observe, so that attention is renewed at each new station and the material is processed repeatedly in different forms. This repeated, active processing is precisely the condition that the literature associates with durable learning, and it accounts for the steady, cumulative rise visible across the four posttests rather than a single isolated jump.

Second, the strategy is inherently social. By requiring students to explain their group's work to visitors and to evaluate the work of other groups, it turns the classroom into a network of peer teaching in which stronger students consolidate their knowledge by articulating it and weaker students receive explanations in language closer to their own. This interpretation is supported by the marked reduction in the standard deviation after the treatment: the narrowing spread suggests that lower-achieving students benefited disproportionately, which is what one would expect if peer explanation and collaboration were helping the group converge toward a common level of understanding.

Third, the visual and game-like character of the activity lowers the affective barriers that often hinder foreign-language learning. Presenting content at stations, using pictures and short texts, and framing the task as "shopping" reduce the anxiety associated with speaking and using Arabic, and increase enjoyment and willingness to participate. The classroom observations recorded during the study, in which students became more active, more cooperative, and more willing to express their understanding, are consistent with this account and echo Asrori's (2023) emphasis on a positive classroom climate as a driver of effective learning.

The findings also align closely with previous empirical work. Husna (2021) reported a positive effect of the Window Shopping strategy on students' Arabic speaking skills, and Wardah (2021) found it effective for grammar (naḥw) learning, while Ibrahim et al. (2022) documented increased confidence in using Arabic. Beyond Arabic, Apriana (2020), Rasyid (2021), Farida (2023), and Nengsih (2022) reported comparable gains in social studies and mathematics. The present study extends this line of research to Durūs al-Lughah, the integrated language lessons at the heart of the pesantren curriculum, and thereby shows that the strategy's benefits are not confined to a single language skill but extend to comprehensive language instruction.

A distinctive contribution of this study lies in its time-series design. Because the students were measured four times before and four times after the intervention, it is possible to see that the pretest scores were stable and that the improvement appeared only after the treatment and then persisted across successive posttests. This pattern makes it far less plausible that the

gain was a chance fluctuation, a practice effect from repeated testing, or the result of maturation, and it therefore strengthens the causal interpretation that the Window Shopping strategy—rather than the mere passage of time was responsible for the improvement. Few earlier studies of the strategy have been able to offer this kind of temporal evidence.

These results carry practical implications for the different stakeholders in Arabic language education. For teachers, the strategy offers a concrete, low-cost, and replicable model for turning a passive lesson into an active one, using only simple materials arranged as stations. For students, it fosters not only higher achievement but also the transferable competencies of cooperation, communication, and critical thinking. For the institution, the results provide evidence to support the wider adoption of cooperation-based methods and to justify investment in teacher training and in the modest facilities such activities require. In this sense the strategy addresses simultaneously the twin problems identified at the outset low achievement and low engagement.

When set against the conventional lecture method that previously dominated the classroom, the advantage of the Window Shopping strategy becomes clearer. Under lecture-based instruction the students had remained largely passive and their scores had stagnated below expectations; under the active, station-based routine the same students improved rapidly and consistently. This contrast supports the broader theoretical claim of cooperative learning that the organization of social interaction and not merely the content delivered is a decisive factor in achievement. In practical terms, it suggests that comparable gains may be within reach in other pesantren classrooms wherever teachers are willing to restructure the lesson around student activity rather than teacher transmission.

The strong effect size observed in this study (Cohen's d = 2.095) also deserves a cautious interpretation. Effect sizes of this magnitude are uncommon in ordinary classroom research and are partly a product of the single-group design, the homogeneous and highly controlled pesantren setting, and the focused nature of the tested material. Rather than treating the exact figure as a fixed estimate of the strategy's power, it is more prudent to read it as strong evidence that the intervention produced a real and educationally important change, whose precise magnitude in other settings would need to be established by studies with control groups and more diverse samples. This reading keeps the enthusiasm generated by the result within the bounds of methodological caution.

Nevertheless, the study has limitations that should temper the interpretation of its findings. It employed a single experimental group without a separate control class, so the comparison rests on the group's own earlier performance rather than on a parallel group taught conventionally; the sample was relatively small (25 students) and drawn from one grade in one pesantren, which limits generalizability; and the material was confined to a single topic (the verb, fi'l). Future research could therefore employ a control-group design with larger and more varied samples, extend the strategy to other topics and language skills, and combine the quantitative measures with richer qualitative data on classroom interaction. Such

studies would help to establish how far the strong effect observed here can be reproduced in other contexts and educational levels.

4. Conclusion

Learning outcomes are the final result of what a student acquires after going through a series of learning processes, covering the cognitive domain related to mental processes such as understanding, application, and analysis, together with the affective and psychomotor domains. Based on the analysis, the Window Shopping strategy has a clear effect on the learning outcomes of Durūs al-Lughah among the second-grade students of Pondok Pesantren Miftahul Huda Suko. Because the Shapiro-Wilk test showed that the data were not normally distributed, the analysis was carried out through the Friedman test and repeated-measures ANOVA with the Greenhouse-Geisser correction, all of which confirmed significant differences among the measurements. The Sig. (2-tailed) value of the paired-sample t-test was $0.001 < 0.05$, the students' mean score rose from 63.60 in the pretest to 82.90 in the posttest—a difference of 19.30 points the dispersion of the scores narrowed, and the effect size (Cohen's $d = 2.095$) was in the very large category. It can therefore be concluded that the Window Shopping strategy is effective in improving students' learning outcomes in Durūs al-Lughah. It is recommended that teachers make use of this strategy in teaching Arabic language lessons, that students pay full attention during its implementation, that the institution facilitate and support teachers in applying it, and that future researchers examine it more deeply preferably with a control group and larger samples in different contexts and educational levels.

REFERENSI

Abdullah, A. F., Rachman, F., & Muhyiddin, L. (2026). Integrating Wordwall in Arabic language learning: Effectiveness, challenges and contextual factors. *JEThAL: Journal of Educational Technology and Arabic Learning*, *1*(1), 1-22.

Annas, I. K., & Abdullah, A. F. (2026). Transformation of Arabic language learning based on augmented reality: Building an immersive and interactive learning experience in Islamic boarding school-based elementary schools. *PIJAR: Jurnal Pendidikan dan Pengajaran*, *4*(2), 1047-1056.

Apriana, B. N. (2020). Model Cooperative Learning Tipe Window Shopping untuk Meningkatkan Hasil Belajar IPS pada Siswa Kelas IX-B SMP Negeri 1 Wanasaba. Jurnal Ilmiah WUNY, 2(2).

Asrori, A. (2023). Membangun Lingkungan Kelas Positif dan Pengaruhnya terhadap Pembelajaran. Yogyakarta: Pustaka Pelajar.

Az-Zuhri, M. (2007). Makānat al-Lughah al-'Arabiyyah fī Fahm al-'Ulūm al-Islāmiyyah. Dār al-Fikr al-'Arabī.

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Creswell, J. W. (2012). Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research (4th ed.). Boston, MA: Pearson Education.

Farida, N. (2023). Peningkatan Aktivitas dan Hasil Belajar Siswa SD Melalui Model Window Shopping. Jurnal Pembelajaran, Bimbingan, dan Pengelolaan Pendidikan, 3(11), 1056–1065.

Field, A. (2013). Discovering Statistics Using IBM SPSS Statistics (4th ed.). London: SAGE Publications.

Gerlach, V. S., & Ely, D. P. (1980). Teaching & Media: A Systematic Approach. Englewood Cliffs: Prentice-Hall.

Greenhouse, S. W., & Geisser, S. (1959). On Methods in the Analysis of Profile Data. Psychometrika, 24(2), 95–112. https://doi.org/10.1007/BF02289823

Hidayah, M. (2018). Pengaruh Metode Pembelajaran terhadap Motivasi Belajar Siswa. Jakarta: Penerbit Manar.

Husna, L. (2021). Pengaruh Strategi Window Shopping terhadap Keterampilan Berbicara Bahasa Arab Siswa. Jurnal Pendidikan Bahasa Arab, 7(2), 115–126.

Ibrahim, M., Aufa, I., Handayani, L. W., Hasana, N., & Az-Zahra, S. R. (2022). Strategi Pembelajaran Bahasa Arab dengan Menggunakan Model Window Shopping, 4.

Mauchly, J. W. (1940). Significance Test for Sphericity of a Normal n-Variate Distribution. The Annals of Mathematical Statistics, 11(2), 204–209. https://doi.org/10.1214/aoms/1177731915

Mulyasa, E. (2016). Strategi Pembelajaran Aktif dan Kolaboratif. Bandung: Alfabeta.

Mustafa, M. Z. (2017). Istirātījiyyāt at-Tadrīs at-Ta'āwunī. Dār at-Ta'līm al-Jadīd.

Nengsih, S. R. (2022). Penerapan Model Pembelajaran Window Shopping dalam Meningkatkan Aktivitas Belajar Siswa pada Materi Bangun Ruang Sisi Lengkung. Jurnal Alpha Euclid Edu, 3(1).

Nuruddin, M. M., Khotimah, S. H., Kurniawan, M. I., & Tawakal, M. G. (2026). Navigating Organizational Conflict in Higher Education: The Mediating Role of Transformative Leadership in Enhancing Public Service Quality. *Qoumun: Journal of Social and Humanities*, *2*(1), 11-23.

Pallant, J. (2020). SPSS Survival Manual (7th ed.). New York: McGraw-Hill Education.

Rahman, A. H. (2020). Pengembangan Pembelajaran Bahasa Arab di Era Digital. Bandung: Alfabeta.

Rasyid, M. (2021). Penerapan Model Window Shopping untuk Meningkatkan Hasil Belajar Siswa pada Pembelajaran IPS di Sekolah Dasar. Diadik: Jurnal Ilmiah Teknologi Pendidikan, 11(2), 164–173.

Sugiyono. (2019). Metode Penelitian Pendidikan: Pendekatan Kuantitatif, Kualitatif, dan R&D. Bandung: Alfabeta.

Wardah, F. (2021). Efektivitas Metode Window Shopping dalam Pembelajaran Nahwu. Jurnal Pendidikan Bahasa, 6(2).