JURNAL BAHASA DAN SASTRA PUSAKA CENDEKIA Vol. 02, No. 01, April 2026, pp. 24~34
DOI: https://doi.org/10.65427/puscen.v2i1.22

Language Games in Primary Arabic Learning: A Pre-Experimental Study

Irawati[1], Anshar Zulhelmi[2]
[1] Madrasah Supervisor, Ministry of Religious Affairs Office, Banda Aceh, Indonesia
[2] Department of Arabic Language and Literature, Faculty of Adab and Humanities, State Islamic University Ar-Raniry Banda Aceh
Email: irawatiwati478@gmail.com*, anshar.zulhelmi@ar-raniry.ac.id

Article Info

Article history:

Received 09/03, 2026
Revised 06/04, 2026
Accepted 30/04, 2026

Keywords:

Arabic Learning
Game-Based Learning
Language Games
Primary Education
Learning Achievement

ABSTRACT

This study re-examines classroom language games in primary Arabic instruction at Madrasah Ibtidaiyah Negeri 1 Banda Aceh and evaluates changes in students' short-term learning achievement. The revised analysis uses a pre-experimental one-group pretest-posttest design supported by classroom observation. The analyzed intervention group comprised 18 students, with achievement measured before and after game-based Arabic learning. Mean scores increased from 28.33 (SD = 11.38) to 72.50 (SD = 11.01), producing a mean gain of 44.17 points; all students improved. Because gain scores showed marginal non-normality, a Wilcoxon signed-rank test was used as the primary inferential analysis and indicated a significant difference ($W = 0$, $p < .001$). A paired-samples t test produced the same conclusion, $t(17) = 23.16$, $p < .001$. Classroom observations also suggested greater participation and enthusiasm, although motivation was not measured with a validated scale. The findings indicate a strong association between the intervention period and improved achievement, but causal claims remain limited by the absence of a control group, the small sample, and incomplete documentation of the intervention protocol

Penelitian ini menelaah kembali penggunaan permainan bahasa dalam pembelajaran bahasa Arab tingkat dasar di Madrasah Ibtidaiyah Negeri 1 Banda Aceh dan mengevaluasi perubahan hasil belajar jangka pendek peserta didik. Analisis revisi menggunakan desain pra-eksperimental satu kelompok dengan pretest-posttest yang didukung observasi kelas. Kelompok intervensi yang dianalisis berjumlah 18 peserta didik dan hasil belajar diukur sebelum serta sesudah pembelajaran bahasa Arab berbasis permainan. Nilai rata-rata meningkat dari 28,33 (SD = 11,38) menjadi 72,50 (SD = 11,01), dengan kenaikan rata-rata 44,17 poin; seluruh peserta didik mengalami peningkatan. Karena skor selisih menunjukkan penyimpangan marginal dari normalitas, uji Wilcoxon digunakan sebagai analisis inferensial utama dan menunjukkan perbedaan signifikan ($W = 0$; $p < 0{,}001$). Uji t berpasangan memberikan kesimpulan yang sama, $t(17) = 23{,}16$; $p < 0{,}001$. Observasi kelas juga mengindikasikan partisipasi dan antusiasme yang lebih tinggi, tetapi motivasi tidak diukur menggunakan skala tervalidasi. Temuan menunjukkan hubungan kuat antara periode intervensi dan peningkatan hasil belajar, namun klaim kausal dibatasi oleh ketiadaan kelompok kontrol, ukuran sampel kecil, dan dokumentasi intervensi yang belum lengkap

1. INTRODUCTION

Arabic at the primary level requires instructional designs that accommodate children's developing attention, confidence, memory, and need for meaningful interaction. Recent work on young Arabic learners identifies recurring pedagogical problems such as monotonous teaching, limited interactive media, low motivation, language anxiety, and mismatches between learning demands and children's cognitive characteristics (Rahmatika & Syafe'i, 2025). These challenges make the choice of instructional strategy especially important in madrasah contexts, where Arabic is both an academic subject and a language connected to religious literacy.

Game-based learning offers one response to these challenges by organizing practice around goals, rules, feedback, challenge, repetition, and active participation. It should be distinguished from gamification: game-based learning uses a game or game-like activity as the learning environment itself, whereas gamification adds selected game elements to a non-game activity. Conceptually, well-designed games can coordinate cognitive, motivational, affective, and social processes, but learning benefits depend on alignment between game mechanics, instructional objectives, feedback, and teacher facilitation (Bado, 2022; Plass et al., 2015). Evidence from primary education likewise suggests that game-based learning can support engagement and learning, although methodological quality and the type of outcomes measured vary substantially across studies (Guan et al., 2024).

Arabic-language education has increasingly adopted both digital and non-digital games. Rezi et al. (2024) report that gamification can support motivation and achievement (Rezi et al., 2024), while Hasibuan et al. (2024) show how a structured Truth-or-Dare and spin-wheel activity can increase speaking participation and peer interaction (Hasibuan et al., 2024). At the elementary level, Ardinal (2024) developed an interactive Gartic-based medium (Ardinal, 2024), and Qomariyah and Bahrodin (2024) examined an adventure game

designed to support critical thinking in Arabic learning (Qomariyah & Bahrodin, 2024). More recently, Al Farisi et al. (2025) found higher Arabic learning outcomes in a game-based experimental class than in a comparison class (Al Farisi et al., 2025), while Aini and Listiani (2025) demonstrated the potential of game-based assessment for Year One primary learners (Listiani & Ashwary, 2025). These studies collectively indicate that games can make Arabic practice more participatory, but they also show that effectiveness depends on learner age, task design, assessment quality, and classroom context.

Within Jurnal Bahasa dan Sastra Pusaka Cendekia, Rahmatika and Syafe'i (2025) emphasize that young Arabic learners need enjoyable, interactive, and developmentally appropriate learning environments (Rahmatika & Syafe'i, 2025), whereas Az-Zahra and Syafe'i (2025) identify language games as one activity that can support communicative confidence and classroom participation (Az-Zahra & Syafe'i, 2025). These findings provide a relevant pedagogical basis for revisiting older classroom datasets on language games with more transparent statistical reporting and more cautious claims about motivation and causality.

The source manuscript for the present study reports classroom implementation of language games at Madrasah Ibtidaiyah Negeri 1 Banda Aceh, together with individual pretest and posttest scores. However, the original methodological description combines a 'procedural' approach with a qualitative label even though the principal evidence is numerical pretest-posttest data. It also infers improved motivation from observation without a standardized motivation instrument. To strengthen methodological coherence, the present revision reframes the study as a pre-experimental one-group pretest-posttest investigation supplemented by classroom observation and reanalyzes the available scores using appropriate descriptive and inferential procedures.

Accordingly, this study addresses two questions: (1) how did students' Arabic learning scores change after the period in which language games were integrated into instruction, and (2) what classroom engagement patterns were observed during the implementation? The study's contribution is deliberately limited: it provides a transparent reanalysis of a small primary-school dataset and distinguishes measured achievement from observed engagement. It does not claim that language games alone caused the improvement, because no control group, random assignment, or validated motivation scale was available.

Language Games and Young Arabic Learners

Language games are structured learning activities in which learners use target-language forms to achieve a goal under explicit or implicit rules. Their instructional value is not simply that they are enjoyable. A productive game can increase the amount of practice, lower the perceived cost of making errors, provide rapid feedback, distribute participation, and create repeated opportunities to retrieve vocabulary or language patterns. In communicative Arabic learning, these functions can be connected to listening, speaking, reading, and writing tasks, provided that the game remains subordinate to a clear language objective rather than becoming entertainment detached from learning.

For young learners, the teacher's role remains central. Games require age-appropriate rules, manageable cognitive load, equitable turn-taking, explicit links to lesson targets, and a post-game phase in which the teacher consolidates vocabulary or structures practiced during play. Without such design, competition may distract from learning, dominant pupils may monopolize participation, or students may remember the activity more clearly than the language content. Thus, the relevant question is not whether games are inherently effective, but under what instructional conditions they support measurable language learning.

2. METHOD

This revision treats the study as a pre-experimental one-group pretest-posttest design supplemented by classroom observation. The design is appropriate for describing within-group change across an instructional period, but it cannot establish causal effectiveness with the same strength as a randomized or controlled quasi-experiment. The study was conducted at Madrasah Ibtidaiyah Negeri 1 Banda Aceh. The source manuscript reports 68 students across three grade levels: 18 in Grade One, 24 in Grade Two, and 26 in Grade Three. The score table contains 18 participants, corresponding to the reported Grade One cohort; the revised report therefore analyzes these 18 students as the intervention group.

The intervention consisted of teacher-designed language games integrated into Arabic instruction. The source manuscript describes the games as classroom activities intended to support practice in the four language skills and to create an enjoyable learning atmosphere. However, the archived manuscript does not provide a sufficiently detailed intervention protocol specifying the exact game names, number and duration of sessions, lesson sequence, or fidelity procedures. For that reason, the revised study does not attribute improvement to any single game type and treats incomplete intervention documentation as a limitation that must be addressed in future replication.

Achievement was measured using a pretest administered before the game-based instructional period and a posttest administered afterward. The archived report does not provide item-level test specifications, evidence of content validity, or reliability coefficients; therefore, the present analysis is restricted to the available total scores and avoids claims about specific language subskills. Classroom observation was used to record participation, responsiveness, and visible enthusiasm during lessons. Because no validated motivation questionnaire was administered, observation is interpreted as evidence of classroom engagement rather than a direct quantitative measure of motivation.

Data analysis proceeded in four stages. First, participant identifiers were anonymized and the raw pretest and posttest values were checked for arithmetic consistency. Second, descriptive statistics were calculated for each measurement and for gain scores. Third, the distribution of gain scores was assessed using the Shapiro-Wilk test. Because the normality test was marginally significant ($W = .896$, $p = .048$), the Wilcoxon signed-rank test was used as the primary inferential analysis. Fourth, a paired-samples t test was reported as a sensitivity analysis to show whether the conclusion changed under a parametric assumption. Statistical significance was evaluated at alpha = .05. No missing scores were present in the 18-participant dataset.

For publication ethics, the revised manuscript removes the student names that appeared in the source table and uses anonymous participant codes. The source manuscript does not document formal ethics approval, parental consent, assent procedures, or data-sharing permission. These items must be verified by the authors and the institution before journal submission, particularly because the participants were minors.

3. RESULTS AND DISCUSSION

a. Result

1. Descriptive Change in Achievement

All 18 participants obtained higher posttest scores than pretest scores. The pretest mean was 28.33 (SD = 11.38), whereas the posttest mean was 72.50 (SD = 11.01). The mean within-student gain was 44.17 points (SD = 8.09), with gains ranging from 35 to 65 points. The original source table contained an arithmetic inconsistency in one reported gain; recalculation from the raw pretest and posttest scores corrects this value and produces the summary shown in Table 1.

Table 1. Descriptive statistics for pretest, posttest, and gain scores

Measure	N	Mean	SD	Minimum	Maximum
Pretest	18	28.33	11.38	10	50
Posttest	18	72.50	11.01	55	90

Gain	18	44.17	8.09	35	65

Table 2. Anonymized participant-level score changes

Participant	Pre	Post	Gain	Participant	Pre	Post	Gain
P01	10	60	50	P10	50	90	40
P02	40	75	35	P11	25	70	45
P03	30	80	50	P12	30	80	50
P04	25	70	45	P13	20	85	65
P05	45	90	45	P14	15	70	55
P06	30	65	35	P15	30	65	35
P07	20	60	40	P16	15	55	40
P08	15	55	40	P17	40	75	35
P09	40	80	40	P18	30	80	50

2. Inferential Analysis

The Shapiro-Wilk test indicated a marginal departure from normality in the gain scores ($W = .896$, $p = .048$). The Wilcoxon signed-rank test was therefore treated as the primary test. Because every posttest score exceeded the corresponding pretest score, the test yielded $W = 0$ and a statistically significant difference ($p < .001$). A paired-samples t test produced the same substantive conclusion, $t(17) = 23.16$, $p < .001$. The standardized paired effect based on the mean and standard deviation of the difference scores was very large (Cohen's $dz = 5.46$). This effect size should nevertheless be interpreted cautiously because one-group pretest-posttest designs can inflate apparent intervention effects through testing, maturation, history, regression, or other uncontrolled influences.

Table 3. Inferential analysis of pretest-posttest change

Analysis	Statistic	p value	Interpretation
Shapiro-Wilk on gain scores	$W = .896$	.048	Marginal non-normality
Wilcoxon signed-rank	$W = 0$	$< .001$	Posttest scores higher
Paired-samples t test	$t(17) = 23.16$	$< .001$	Same substantive conclusion
Standardized paired effect	$dz = 5.46$	—	Very large; interpret cautiously

3. Classroom Engagement Observations

The source manuscript reports that students participated more actively and appeared more enthusiastic during Arabic lessons that incorporated language games. Observable indicators included willingness to take part in tasks, responsiveness to teacher instructions, and reduced monotony in classroom activity. These observations are pedagogically relevant, but they should not be presented as direct evidence of increased motivation because the study did not use a validated motivation scale, structured observation rubric, inter-rater agreement, or repeated measurement of affective outcomes. The revised interpretation therefore uses the term observed engagement rather than measured motivation.

b. Discussion

The most robust empirical result is the large within-group increase in Arabic achievement scores. The mean rose by 44.17 points, and every student improved. This pattern is consistent with recent Arabic game-based learning research. Al Farisi et al. (2025), using a quasi-experimental comparison, reported stronger post-intervention outcomes in an elementary Arabic class exposed to game-based learning than in a control class (Al Farisi et al., 2025). Aini and Listiani (2025) likewise show that game-oriented assessment can be adapted to Year One primary learners (Aini & Listiani, 2025), while Ardinal (2024) illustrates how interactive game media can support vocabulary-oriented Arabic learning in an integrated Islamic elementary school context (Ardinal, 2024). The present findings therefore fit a wider pattern in which structured play can increase opportunities for practice and make language tasks more accessible to younger learners.

The improvement can be interpreted through several mechanisms proposed in game-based learning literature. Games can create repeated retrieval opportunities, immediate or visible feedback, short goal cycles, peer interaction, and lower-stakes practice. These characteristics may help learners sustain attention and rehearse language forms more frequently than in teacher-dominated instruction (Bado, 2022; Plass et al., 2015). Hasibuan et al. (2024), for example, found that a game combining a spin wheel with Truth-or-Dare prompts increased speaking turns and interaction (Hasibuan et al., 2024). Although the present archive does not document comparable interaction counts, the reported classroom observations of increased participation are compatible with this mechanism.

The age of learners also matters. Rahmatika and Syafe'i (2025) argue that Arabic teaching for young learners should address cognitive development, low motivation, anxiety, and the need for interactive media (Rahmatika & Syafe'i, 2025). Games may be especially useful when they reduce the abstraction of language learning and convert practice into concrete, time-bounded actions. However, young learners also require careful scaffolding. Rules that are too complex, excessive competition, or insufficient debriefing can shift attention away from the language target. The teacher must therefore connect each activity to a clear linguistic objective and consolidate learning after play.

The observed increase in participation should be separated analytically from the measured achievement increase. The source manuscript originally treated student enthusiasm as evidence that motivation had improved. That inference is plausible but not demonstrated. Motivation is a multidimensional construct and normally requires a validated questionnaire, interview protocol, or systematic observation instrument. Rezi et al. (2024) discuss motivation and achievement in Arabic gamification, but their survey-based design differs from the present study (Rezi et al., 2024). The current data support a narrower claim: students appeared more engaged during game-based classroom activities, while achievement scores increased substantially across the same period.

Methodologically, the absence of a comparison group is the principal limitation. A one-group pretest-posttest design cannot rule out alternative explanations such as familiarity with test format, concurrent instruction, maturation, or differences in test difficulty. The very large standardized effect should therefore not be interpreted as a pure causal effect of language games. This is important because stronger designs in the field, such as the quasi-experimental study by Al Farisi et al. (2025), provide a more defensible basis for causal comparison (Al Farisi et al., 2025). The current study is best understood as evidence of promising change associated with an instructional period rather than definitive proof that games alone produced the gains.

A second limitation concerns intervention and measurement transparency. The archived report does not preserve sufficient detail about game types, session duration, teacher scripts, implementation fidelity, or the psychometric quality of the achievement test. This restricts replication and makes it impossible to identify which game features were most productive. Future studies should predefine the intervention protocol, specify targeted language skills, validate test items, collect standardized motivation or engagement measures, and include delayed posttests to examine retention. Studies may also compare competitive and cooperative games, digital and non-digital formats, or different levels of rule complexity.

Despite these constraints, the study has practical value for primary Arabic teachers. It shows how an existing classroom dataset can be reported more transparently by anonymizing students, correcting arithmetic inconsistencies, distinguishing measured from observed outcomes, and applying statistical procedures aligned with the data distribution. Pedagogically, language games appear most defensible when used as structured practice rather than as stand-alone entertainment. Their role is to increase meaningful language use, feedback, and participation while remaining integrated with explicit instructional goals.

4. CONCLUSION

The reanalysis shows a substantial increase in Arabic achievement scores among 18 primary-school students after a period of instruction incorporating language games. Mean scores increased from 28.33 to 72.50, all students improved, and both nonparametric and parametric analyses indicated a statistically significant pretest-posttest difference. Classroom observations also suggested greater participation and enthusiasm during game-based activities. These results support the practical potential of language games as a structured component of Arabic instruction for young learners.

The findings, however, should be interpreted as an association rather than definitive causal evidence. The study used a single group, involved a small sample, lacked a validated

measure of motivation, and did not preserve a sufficiently detailed intervention protocol or test-validation record. Future research should use a control or comparison group, document the games and implementation procedures precisely, validate achievement and affective instruments, and assess retention over time. Before publication, the authors should also verify ethics documentation, participant consent procedures, and the provenance and originality of the dataset.

AUTHOR CONTRIBUTIONS

The authors should confirm and finalize their individual contributions using the CRediT taxonomy before submission. The source manuscript identifies Irawati and Anshar Zulhelmi as authors but does not provide a verified contribution statement.

FUNDING

Funding information was not reported in the source manuscript and should be confirmed by the authors before submission.

ETHICS STATEMENT

The revised manuscript anonymizes participant-level data. Formal ethics approval, parental or guardian consent, participant assent, and authorization for secondary reporting of the archived dataset were not documented in the source manuscript and must be verified before submission.

CONFLICT OF INTEREST

The authors should provide and confirm a conflict-of-interest declaration before submission.

DATA AVAILABILITY

The participant-level pretest and posttest scores analyzed in this revision were reproduced from the source manuscript and are reported anonymously in Table 2. Any broader sharing of the original dataset should follow institutional and ethical requirements for research involving minors.

REFERENCES

Aini, N., & Listiani, H. (2025). Assessing Arabic learning through a game-based evaluation tool for Year One primary school students. *UPGRADE: Journal of Education*, *3*(1), 22–33. https://doi.org/10.30812/g3yk7e05

Al Farisi, M. F. Y., Mustamin, M., & Masdar, M. I. (2025). Efektivitas pembelajaran berbasis permainan dalam meningkatkan hasil belajar bahasa Arab. *El-Syaker:*

Samarinda International Journal of Language Studies, *2*(1), 1–16. https://doi.org/10.64093/esijls.v2i1.281

Ardinal, E. (2024). Development of Gartic platform as Arabic learning media based on interactive quiz at integrated Islamic elementary school. *Lughawiyah: Journal of Arabic Education and Linguistics*, *6*(2), 100–119. https://doi.org/10.31958/lughawiyah.v6i2.13412

Az-Zahra, M., & Syafe'i, I. (2025). Qualitative analysis of the effectiveness of the communicative learning model in Arabic language instruction. *Jurnal Bahasa Dan Sastra Pusaka Cendekia*, *1*(2), 86–97. https://doi.org/10.65427/puscen.v1i2.11

Bado, N. (2022). Game-based learning pedagogy: A review of the literature. *Interactive Learning Environments*, *30*(5), 936–948. https://doi.org/10.1080/10494820.2019.1683587

Guan, X., Sun, C., Hwang, G.-J., Xue, K., & Wang, Z. (2024). Applying game-based learning in primary education: A systematic review of journal publications from 2010 to 2020. *Interactive Learning Environments*, *32*(2), 534–556. https://doi.org/10.1080/10494820.2022.2091611

Hasibuan, R., Jundi, M., Ali, I., Febriani, S. R., & Bedra, K. G. (2024). Game-based language learning: Implementing Arabic speaking proficiency through Truth or Dare with Spin Wheel. *Lughawiyah: Journal of Arabic Education and Linguistics*, *6*(1), 16–33. https://doi.org/10.31958/lughawiyah.v6i1.12224

Listiani, A., & Ashwary, F. (2025). Analisis kontrastif: Perbandingan struktur kalimat pada penerjemahan teks bahasa Arab Al-Jumlah Al-Mufidah ke dalam bahasa Indonesia. *Madah: Jurnal Bahasa Dan Sastra*, *16*(1), 96–109. https://doi.org/10.31503/madah.v16i1.983

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of game-based learning. *Educational Psychologist*, *50*(4), 258–283. https://doi.org/10.1080/00461520.2015.1122533

Qomariyah, L., & Bahrodin, A. (2024). Developing an adventure game for teaching Arabic

to develop critical thinking skills of elementary school students. *Ijaz Arabi Journal of Arabic Learning*, *7*(3), 1014–1029. https://doi.org/10.18860/ijazarabi.v7i3.26098

Rahmatika, A., & Syafe'i, I. (2025). Challenges in teaching Arabic to early childhood learners: Pedagogical and psycholinguistic perspectives. *Jurnal Bahasa Dan Sastra Pusaka Cendekia*, *1*(2), 77–85. https://doi.org/10.65427/puscen.v1i2.10

Rezi, M., Noor, A. F. M., & Nicholas, T. (2024). Utilization of gamification in Arabic language learning to increase student motivation and achievement. *Lughawiyah: Journal of Arabic Education and Linguistics*, *6*(1), 1–15. https://doi.org/10.31958/lughawiyah.v6i1.12559