JOURNAL OF INNOVATION AND RESEARCH IN PRIMARY EDUCATION | ISSN 2829-775X
Vol. 5, No. 1, 2026, 685-696
https://doi.org/10.56916/jirpe.v5i1.2858

# The Effectiveness of the Word Square Learning Model on Reading Ability: A Quasi-Experimental Study of Fourth-Grade Students in Indonesian Language Instruction

**Ayu Amalia Shaltina**
Tadulako University, Palu, Indonesia

**Herlina**
Tadulako University, Palu, Indonesia

**Sisriawan Lapasere**
Tadulako University, Palu, Indonesia

**Muhammad Aqil**
Tadulako University, Palu, Indonesia

**Putriwanti**
Tadulako University, Palu, Indonesia

**Surahman**
Tadulako University, Palu, Indonesia

***Corresponding Author:** ayuweb018@gmail.com

**Keywords**
Word Square model
reading ability
game-based learning
elementary education
Indonesian language instruction

**Article History**
Received 2025-11-11
Accepted 2026-01-16



**Abstract**
Reading ability constitutes a fundamental skill essential for academic success, yet many elementary students experience persistent difficulties in developing adequate reading competencies. The Word Square model, a game-based puzzle intervention, offers a promising approach to enhance reading instruction through active word recognition and contextual learning. This study investigated the effectiveness of the Word Square learning model on reading ability among fourth-grade elementary students in Indonesian language instruction. A quasi-experimental design with nonequivalent control groups was employed, involving 44 students divided into experimental (n=22) and control (n=22) groups. The experimental group received Word Square instruction while the control group participated in conventional teaching. Reading ability was assessed through pretest and posttest measurements. Data were analyzed using descriptive statistics, Shapiro-Wilk normality test, Levene's homogeneity test, independent samples t-test, and normalized gain (N-gain) analysis. The experimental group achieved significantly higher posttest scores (M=89.64) compared to the control group (M=70.86), with t=13.636, p<0.001. The N-gain coefficient of 0.7325 indicated high effectiveness, demonstrating substantial improvement relative to students' growth potential. The Word Square learning model significantly enhances reading ability among fourth-grade students, providing robust empirical support for game-based, student-centered pedagogical approaches in elementary literacy instruction. These findings offer practical implications for educators seeking evidence-based strategies to improve reading outcomes.

## INTRODUCTION

Reading ability constitutes a fundamental skill essential for academic success and lifelong learning, serving as the cornerstone of educational development across all disciplines. As a primary literacy competence, reading enables students to access, comprehend, and critically analyze information presented in various forms of written communication (Lyon, 1998; Fauziah, 2024). Research demonstrates that early reading abilities predict long-term academic achievement, with skills measured in kindergarten correlating significantly with reading performance through fourth and sixth grades (Senechal & LeFevre, 2002; Wagner et al., 1997). The significance of reading proficiency

extends beyond mere decoding of written symbols; it encompasses the cognitive processes of understanding, interpreting, and synthesizing textual information to construct meaningful knowledge (Sugiarti et al., 2012). Furthermore, reading functions as a mechanism for enhancing both verbal and non-verbal language competencies, thereby facilitating comprehensive linguistic development (Harianto, 2020). Among the four fundamental language skills—reading, writing, listening, and speaking—reading occupies a particularly critical position in the educational trajectory of elementary school students, as it directly correlates with overall academic achievement and learning efficacy across all content areas (Handayani, 2022).

The cultivation of reading skills at the elementary school level represents a pedagogical priority, as proficiency in reading significantly influences students' capacity to engage with curriculum content across all subject areas (Fauziah, 2024). Regular engagement in reading activities yields multifaceted benefits, including vocabulary expansion, development of imagination and creativity (Fisnawati, 2024), knowledge acquisition, articulation refinement, cultivation of critical thinking abilities (Rahma et al., 2024), and enhancement of social-emotional development (Fahyuni & Bandono, 2015). Conversely, reading difficulties present serious and potentially lifelong challenges, as students with deficiencies in reading competence experience impeded academic progress, diminished learning outcomes, and increased risk of grade retention and school dropout (Hall et al., 2023; Fauziah, 2024; Vaughn & Fletcher, 2020). Indeed, reading proficiency has been identified as one of the principal determinants of academic success, with students at the bottom tenth percentile of word reading presenting particularly challenging cases for intervention (Gustiawati, 2020; Wanzek et al., 2013).

Despite the recognized importance of reading skills, empirical evidence suggests that a substantial proportion of elementary school students continue to experience difficulties in developing adequate reading competencies. Recent assessments indicate that approximately 65% of fourth-grade students in the United States read at or below basic levels, highlighting persistent challenges in reading instruction (National Assessment of Educational Progress, 2020). This pedagogical challenge necessitates the implementation of innovative instructional approaches that can effectively address the diverse learning needs of students while maintaining engagement and motivation. Traditional teaching methodologies, characterized by conventional lecture-based instruction and passive learning environments, have demonstrated limited effectiveness in developing students' reading abilities, particularly in fostering active participation and critical engagement with textual material (Wanzek et al., 2018). This pedagogical gap has prompted educational researchers to explore alternative instructional models that emphasize interactive learning, student-centered activities, and strategic engagement with written content.

Game-based learning approaches represent promising pedagogical innovations that integrate play-based elements with systematic reading instruction. Research indicates that game-based interventions significantly improve students' motivation, engagement, and reading skills compared to traditional instructional methods (Eng et al., 2023; Ronimus & Lyytinen, 2015). Word puzzle activities, including crossword puzzles and word search games, have demonstrated effectiveness in enhancing vocabulary acquisition and reading comprehension among elementary students (Lauterbach & Zipke, 2024; Orawiwatnakul, 2013). The Word Square model, which employs a structured word-finding technique utilizing grid-based puzzles, requires students to actively search for, identify, and connect words within a matrix format. The pedagogical rationale underlying this model posits that interactive approaches promote active learning, enhance cognitive engagement, and develop both analytical and critical thinking skills while simultaneously improving reading fluency and comprehension (Azizah et al., 2022; Listini, 2021). Metalinguistic activities involving word puzzles have been shown to positively impact reading development by encouraging conscious reflection on and manipulation of language at multiple levels (Lauterbach & Zipke, 2024).

Previous research has documented the potential effectiveness of puzzle-based learning interventions in various educational contexts. Systematic reviews of word search puzzle implementation reveal consistent improvements in students' vocabulary across experimental studies,

with experimental groups demonstrating significantly higher post-test scores compared to control groups receiving conventional instruction (Orawiwatnakul, 2013). Furthermore, Ayuningtyas et al. (2020) emphasized that the Word Square learning model, as a development of enriched lecture methods, combines the ability to answer questions with the capacity to match answers in grid-based formats, thereby promoting active learning orientations. The model's pedagogical design distinguishes it from traditional crossword puzzles through its use of pre-populated answers concealed among distractor letters, requiring careful analysis and critical thinking rather than generation of answers from memory alone. Research by Adilah et al. (2024) demonstrated that interactive, game-based instructional materials incorporating letter cards significantly enhanced letter recognition and vocabulary development among elementary students, supporting the efficacy of engaging learning tools in literacy instruction.

Despite promising theoretical foundations and preliminary empirical support for game-based reading interventions, significant gaps remain in understanding the effectiveness of the Word Square model in specific educational contexts, particularly in Indonesian language instruction at the elementary school level. While meta-analyses have established general effectiveness of game-based learning with medium to large effect sizes (Bang & Siebert-Evenstone, 2025), limited systematic research has examined the quantitative impact of the Word Square model specifically on reading ability outcomes using rigorous experimental designs with appropriate control groups. Moreover, the extent to which this model can produce statistically significant improvements in reading competencies compared to conventional instructional approaches in Indonesian educational settings remains inadequately documented.

This study addresses this research gap by systematically investigating the effectiveness of the Word Square model on the reading ability of fourth-grade elementary school students in Indonesian language instruction. Employing a quasi-experimental design with nonequivalent control groups, this research aims to provide empirical evidence regarding the efficacy of the Word Square model through comprehensive pre-test and post-test assessments. The significance of this study lies in its potential to contribute evidence-based pedagogical recommendations for Indonesian language instruction, thereby informing instructional practice and curriculum development initiatives aimed at enhancing reading competencies among elementary school students. By examining the quantitative impact of this innovative instructional model within the Indonesian educational context, this research seeks to advance understanding of effective pedagogical approaches for reading instruction and provide empirical support for the integration of interactive, student-centered learning strategies in elementary education.

## METHODS

This study employed a quantitative research approach utilizing a quasi-experimental design, specifically the nonequivalent control group pretest-posttest design. This design was selected as it allows for manipulation of the independent variable while examining causality between an intervention and outcome when random assignment is not logistically feasible. The nonequivalent control group design represents a between-subjects research framework in which participants are assigned to groups based on pre-existing classroom configurations rather than random assignment, making the resulting groups potentially dissimilar in certain respects (Sugiyono, 2021). Despite the absence of randomization, this design is generally higher in internal validity than correlational studies, particularly when enhanced through pretest-posttest measurements that account for baseline differences between groups. The research focused specifically on evaluating the effectiveness of the Word Square learning model on reading ability outcomes among fourth-grade elementary school students in Indonesian language instruction.

The study was conducted at SD Inpres Kalukubula, located on Jalan Guru Tua in Kalukubula Village, Biromaru District, Sigi Regency, during the second semester of the 2024-2025 academic year. The research sample consisted of 44 fourth-grade students divided into two intact classroom groups:

an experimental group of 22 students who received instruction using the Word Square learning model, and a control group of 22 students who received conventional instruction. Both groups were measured at two time points—before and after the intervention—to assess changes in reading ability. This dual-measurement approach strengthens the internal validity of the quasi-experimental design by enabling evaluation of whether students receiving the treatment improve more substantially than those receiving conventional instruction (Gray, 2023). The selection of intact classroom groups rather than individual random assignment reflects practical constraints in educational settings where disrupting existing class compositions would be neither feasible nor ethically appropriate.

Data collection was conducted through administration of a reading ability questionnaire developed specifically for this study. Prior to implementation in the main study, the instrument underwent rigorous validation procedures to ensure both validity and reliability. Validity refers to the accuracy and meaningfulness of measurements, examining whether the instrument effectively measures what it purports to measure (Taherdoost, 2016). The questionnaire's validity was assessed using the Product Moment Correlation formula with IBM SPSS Statistics version 26, following established procedures for content validity evaluation. Content validity ensures that the instrument adequately covers the domain of reading ability being assessed. Reliability, which refers to the consistency and stability of measurement results across repeated administrations, was similarly evaluated using SPSS version 26 (Drost, 2011). An instrument demonstrating high reliability produces consistent results when applied under similar conditions, thereby enhancing confidence in the measurement process. The validation process followed standard psychometric procedures to establish that the instrument possessed adequate measurement properties before deployment in the experimental study (Budiastuti, 2018; Ahmad et al., 2020).

The research procedure involved three distinct phases. First, both experimental and control groups completed a pretest to establish baseline reading ability levels. Subsequently, the experimental group participated in instructional sessions incorporating the Word Square learning model, while the control group received conventional reading instruction without the specialized intervention. Finally, both groups completed a posttest using the same reading ability instrument administered during the pretest phase. This pretest-posttest design with nonequivalent comparison groups represents one of the most widely employed quasi-experimental approaches in educational research, as it provides evidence of treatment effectiveness by comparing improvement trajectories between groups rather than relying solely on posttest differences.

Data analysis procedures followed a systematic sequence designed to ensure appropriate statistical treatment of the quantitative data. Initially, descriptive statistics were calculated to summarize pretest and posttest performance in both groups, providing measures of central tendency and dispersion. Subsequently, the Shapiro-Wilk normality test was conducted using IBM SPSS Statistics version 23 to determine whether the data met the assumption of normal distribution, with the criterion that significance values greater than 0.05 indicate normally distributed data. Following confirmation of normality, the homogeneity of variance between groups was assessed using Levene's test, with significance values exceeding 0.05 indicating homogeneous variance. These prerequisite analyses ensured that parametric statistical procedures were appropriate for hypothesis testing. The independent samples t-test was then employed to evaluate whether statistically significant differences existed between the experimental and control groups' posttest scores, with the decision criterion that significance values less than 0.05 warrant rejection of the null hypothesis. Additionally, normalized gain (N-gain) analysis was conducted to quantify the effectiveness of the Word Square learning model intervention by calculating the magnitude of improvement from pretest to posttest in the experimental group. The N-gain coefficient provides a standardized measure of learning effectiveness that accounts for initial performance levels, with values exceeding 0.7 indicating high effectiveness. This comprehensive analytical approach enabled rigorous evaluation of the Word Square model's impact on students' reading ability development.

## RESULTS AND DISCUSSION

### Results

This section presents the findings from the quasi-experimental study investigating the effectiveness of the Word Square learning model on reading ability among fourth-grade students at SD Inpres Kalukubula. The analysis encompasses descriptive statistics, prerequisite tests for parametric analysis, hypothesis testing, and normalized gain assessment. Data were collected from 44 students divided into experimental (n=22) and control (n=22) groups through pretest and posttest measurements administered before and after the instructional intervention.

#### *Descriptive Statistics of Pretest and Posttest Performance*

The initial assessment phase involved administering a reading ability pretest to both experimental and control groups to establish baseline performance levels. Table 1 presents the descriptive statistics for pretest scores across both groups. The experimental group achieved a mean pretest score of 61.45 (SD not reported), with scores ranging from a minimum of 52 to a maximum of 72. In comparison, the control group obtained a mean pretest score of 61.86, with scores ranging from 53 to 69. The negligible difference of 0.41 points between group means indicates that both groups possessed comparable initial reading ability levels prior to intervention, suggesting that the groups were relatively equivalent at baseline despite the non-random assignment inherent in the quasi-experimental design.

**Table 1.** Descriptive Statistics for Pretest Scores

| Statistics | Control Group | Experimental Group |
|---|---|---|
| Mean | 61.45 | 61.86 |
| Minimum Score | 52 | 53 |
| Maximum Score | 72 | 69 |

Following the completion of the intervention period, during which the experimental group received instruction utilizing the Word Square learning model and the control group participated in conventional reading instruction, both groups completed a posttest using the same reading ability instrument. Table 2 presents the descriptive statistics for posttest performance. The experimental group demonstrated substantial improvement, achieving a mean posttest score of 89.64, with scores ranging from 80 to 100. In contrast, the control group obtained a mean posttest score of 70.86, with scores ranging from 62 to 78. The experimental group's mean score exceeded that of the control group by 18.78 points, representing a substantial difference in reading ability outcomes following differential instructional treatment.

**Table 2.** Descriptive Statistics for Posttest Scores

| Statistics | Control Group | Experimental Group |
|---|---|---|
| Mean | 70.86 | 89.64 |
| Minimum Score | 62 | 80 |
| Maximum Score | 78 | 100 |

The magnitude of improvement from pretest to posttest differed considerably between groups. The experimental group exhibited a mean gain of 28.19 points (from 61.45 to 89.64), whereas the control group demonstrated a mean gain of 9.00 points (from 61.86 to 70.86). This differential improvement pattern suggests that the Word Square learning model produced substantially greater enhancement in reading ability compared to conventional instructional methods.

#### *Normality Testing*

Prior to conducting parametric statistical analyses, the assumption of normal distribution was evaluated using the Shapiro-Wilk test, which is generally preferred for sample sizes below 50 due to its superior statistical power compared to the Kolmogorov-Smirnov test. Table 3 presents the normality test results for both pretest and posttest data across experimental and control groups. The decision criterion established that data would be considered normally distributed if the significance value exceeded 0.05 ($p > 0.05$).

**Table 3.** Shapiro-Wilk Normality Test Results

| Data Set | Shapiro-Wilk Statistic | df | Significance |
|---|---|---|---|
| Experimental Group Pretest | 0.977 | 22 | 0.866 |
| Experimental Group Posttest | 0.926 | 22 | 0.100 |
| Control Group Pretest | 0.946 | 22 | 0.257 |
| Control Group Posttest | 0.946 | 22 | 0.257 |

For the experimental group, the Shapiro-Wilk test yielded a significance value of 0.866 for pretest data and 0.100 for posttest data, both substantially exceeding the 0.05 threshold. Similarly, the control group demonstrated significance values of 0.257 for both pretest and posttest assessments. All significance values exceeded 0.05, thereby satisfying the normality assumption and confirming that the data distributions across all measurement points did not deviate significantly from normal distribution. This finding validated the appropriateness of employing parametric statistical procedures for subsequent hypothesis testing.

***Homogeneity of Variance Testing***

To assess whether the experimental and control groups exhibited similar variance in their reading ability scores, Levene's test for homogeneity of variance was conducted. This prerequisite test determines whether the assumption of equal variances across groups is satisfied, which is essential for valid interpretation of independent samples t-test results. Table 4 presents the homogeneity test results.

**Table 4.** Levene's Test for Homogeneity of Variance

| Test Basis | Levene Statistic | df1 | df2 | Significance |
|---|---|---|---|---|
| Based on Mean | 0.559 | 1 | 42 | 0.459 |

The Levene's test yielded a significance value of 0.459, which exceeds the 0.05 criterion. This result indicates that the variances between the experimental and control groups were statistically homogeneous, satisfying the homogeneity of variance assumption required for independent samples t-test. The fulfillment of both normality and homogeneity assumptions confirmed that parametric testing procedures were appropriate for evaluating between-group differences.

***Hypothesis Testing***

To determine whether statistically significant differences existed in reading ability outcomes between the experimental and control groups following the intervention, an independent samples t-test was conducted on posttest scores. The null hypothesis ($H_0$) posited that no significant difference existed between groups, while the alternative hypothesis ($H_a$) proposed that the Word Square learning model produced significantly different reading ability outcomes compared to conventional instruction. Table 5 presents the results of the independent samples t-test.

**Table 5.** Independent Samples T-Test Results

| Variance Assumption | t-statistic | df | One-Sided p | Two-Sided p |
|---|---|---|---|---|
| Equal variances assumed | 13.636 | 42 | 0.001 | 0.001 |

The independent samples t-test yielded a t-statistic of 13.636 with 42 degrees of freedom. Both one-sided ($p = 0.001$) and two-sided ($p = 0.001$) significance values were substantially below the 0.05 alpha level, indicating statistical significance at the $p < 0.001$ level. This result provides strong evidence to reject the null hypothesis and accept the alternative hypothesis, confirming that a statistically significant difference existed in reading ability between students who received instruction using the Word Square model and those who participated in conventional instruction. The magnitude of the t-statistic (13.636) indicates a large effect size, suggesting substantial practical significance in addition to statistical significance.

### *Normalized Gain Analysis*

To evaluate the effectiveness magnitude of the Word Square learning model intervention, normalized gain (N-gain) analysis was conducted. The N-gain coefficient quantifies the actual gain relative to the maximum possible gain, providing a standardized measure of learning effectiveness that accounts for initial performance levels. N-gain values are interpreted according to established criteria: low effectiveness (N-gain < 0.3), moderate effectiveness (0.3 ≤ N-gain ≤ 0.7), and high effectiveness (N-gain > 0.7). Table 6 presents the descriptive statistics for N-gain scores in the experimental group.

**Table 6.** Normalized Gain (N-gain) Analysis Results

| Statistic | N | Minimum | Maximum | Mean | Standard Deviation |
|---|---|---|---|---|---|
| N-gain | 22 | 0.46 | 1.00 | 0.7325 | 0.12672 |

The N-gain analysis revealed a mean value of 0.7325 (SD = 0.12672) for the experimental group, with individual N-gain scores ranging from 0.46 to 1.00. The mean N-gain value of 0.7325 exceeds the 0.7 threshold, indicating high effectiveness according to established interpretation criteria. This finding demonstrates that the Word Square learning model produced substantial learning gains that cannot be attributed solely to natural maturation or test familiarity effects. The relatively small standard deviation (0.12672) suggests consistent effectiveness across students, although the range from 0.46 to 1.00 indicates some variability in individual responses to the intervention.

### *Unexpected Findings*

An unexpected observation emerged from the maximum score analysis, where several students in the experimental group achieved perfect scores of 100 on the posttest, a result that warrants careful interpretation. While this outcome might initially suggest ceiling effects that could potentially constrain measurement sensitivity, the presence of minimum scores of 80 within the same group indicates that the assessment retained sufficient difficulty to differentiate among high-performing students. The achievement of perfect scores by multiple students may reflect genuine mastery facilitated by the Word Square model's emphasis on active word recognition and contextual learning, rather than measurement limitations. However, this finding also suggests that future research might benefit from incorporating more challenging assessment items or additional measures of reading comprehension depth to capture the full range of reading ability development, particularly among high-achieving students.

Additionally, the control group's modest improvement (mean gain of 9.00 points) relative to their baseline performance, though statistically lower than the experimental group, exceeded what might be expected from test-retest effects alone. This pattern suggests that conventional instruction did produce measurable learning, but the magnitude was substantially smaller compared to the Word Square intervention. The fact that both instructional approaches produced improvement, albeit of different magnitudes, underscores the importance of controlled comparison designs in educational intervention research.

## Discussion

The present study investigated the effectiveness of the Word Square learning model on reading ability among fourth-grade elementary school students in Indonesian language instruction. The findings provide compelling evidence that game-based puzzle interventions can significantly enhance reading outcomes compared to conventional instructional approaches. This discussion interprets the results within theoretical frameworks of interactive learning, relates findings to existing literature, examines implications for educational practice, acknowledges limitations, and synthesizes the study's conceptual contributions.

The central finding of this study—that students receiving instruction through the Word Square model demonstrated significantly superior reading ability (M = 89.64) compared to peers in

conventional instruction (M = 70.86)—directly supports the research hypothesis and aligns with contemporary theories of active learning and cognitive engagement. The substantial mean difference of 18.78 points, coupled with a highly significant t-statistic ($t = 13.636$, $p < 0.001$) and high normalized gain coefficient (N-gain = 0.7325), provides robust evidence for the Word Square model's effectiveness. These results confirm that puzzle-based instructional approaches incorporating word search and recognition activities can produce meaningful improvements in elementary students' reading competencies.

The theoretical rationale underlying this effectiveness can be understood through multiple cognitive mechanisms. First, the Word Square model operationalizes active learning principles by requiring students to engage in systematic visual scanning, pattern recognition, and strategic word identification rather than passive reception of information (Fredricks et al., 2016). This active cognitive engagement likely enhanced attention allocation and working memory activation during learning, processes known to facilitate encoding and retention (Laal & Ghodsi, 2012). Second, the puzzle-based format provides immediate feedback through self-checking mechanisms inherent in word-finding activities, enabling students to monitor their performance and adjust strategies in real time. This aligns with research demonstrating that timely, specific feedback promotes engagement and learning outcomes (Hattie & Timperley, 2007).

The high N-gain coefficient (0.7325) observed in this study suggests that the Word Square model facilitated substantial learning relative to students' growth potential, indicating effectiveness that extends beyond mere statistical significance to pedagogically meaningful impact. This magnitude of normalized gain is particularly noteworthy given that educational interventions frequently produce moderate effect sizes (0.3-0.7 range). The consistency of high effectiveness across the experimental group, as evidenced by the relatively small standard deviation (0.12672), further supports the robustness and generalizability of the intervention's impact across diverse learners within the fourth-grade population.

The present findings align closely with previous research documenting the efficacy of game-based and puzzle-based learning interventions in literacy development. The substantial improvements in reading ability observed in this study are consistent with Orawiwatnakul's (2013) research demonstrating that crossword puzzles significantly enhanced vocabulary acquisition among university students, with experimental groups showing significantly higher posttest scores compared to control groups. Although Orawiwatnakul's study focused on older learners and vocabulary specifically, the underlying mechanism—active word manipulation within structured puzzle formats—parallels the cognitive processes engaged by the Word Square model in elementary reading instruction.

Furthermore, the current results corroborate findings from meta-analytic research on game-based learning in early literacy. Bang and Siebert-Evenstone (2025) reported that educational apps incorporating game-based elements produced effect sizes comparable to tutoring interventions (ES = 0.37) and early elementary literacy programs (ES = 0.39). The high N-gain coefficient observed in the present study (0.7325) suggests that the Word Square model may produce even stronger effects than typical game-based applications, potentially due to its specific targeting of word recognition and contextual learning—fundamental components of reading proficiency. Similarly, Eng et al. (2023) and Ronimus and Lyytinen (2015) documented significant improvements in reading skills and motivation among struggling readers and first-grade students, respectively, using game-based reading programs. The present study extends these findings to Indonesian language instruction in fourth-grade students using an intact classroom quasi-experimental design, thereby enhancing ecological validity.

The motivational dimension of the Word Square model, while not directly measured in this study, likely contributed to observed learning gains. Research consistently demonstrates that game-based learning enhances student motivation and engagement (Liu & Chen, 2013; Ronimus & Lyytinen, 2015), which in turn correlates with improved academic outcomes. The puzzle-based format inherently incorporates elements of challenge, discovery, and accomplishment—psychological factors that support intrinsic motivation according to self-determination theory. Students' active search for

words within the grid likely fostered a sense of autonomy and competence, both critical components of motivated learning behavior.

Notably, the present findings diverge somewhat from studies reporting more modest effects of word puzzle interventions. For instance, some research on word search puzzles has shown mixed results, with benefits primarily limited to vocabulary recognition rather than broader reading comprehension. The stronger effects observed in this study may reflect the Word Square model's integration with structured Indonesian language instruction, systematic implementation protocols, and alignment with curriculum objectives—factors that may amplify intervention effectiveness compared to standalone puzzle activities.

The results also resonate with Adilah et al.'s (2024) study demonstrating that interactive, game-based instructional materials incorporating letter cards significantly enhanced letter recognition and vocabulary development among elementary students. Both interventions share core design principles emphasizing active manipulation of linguistic elements within engaging formats. However, the present study advances this work by employing rigorous experimental methodology with matched control conditions, thereby isolating the specific contribution of the Word Square model to reading ability development.

Theoretically, this study contributes to the growing body of evidence supporting constructivist and cognitive engagement perspectives in literacy education. The findings suggest that reading instruction benefits substantially when students actively construct knowledge through interactive problem-solving activities rather than passively receiving information. The Word Square model exemplifies how strategic integration of game elements with academic content can enhance cognitive processing depth, thereby facilitating more durable learning. These results encourage educators to reconsider traditional lecture-based approaches, particularly in foundational skills instruction where active engagement proves especially beneficial.

From a practical standpoint, the Word Square model offers a scalable, low-technology intervention that can be readily implemented in diverse educational settings, including resource-limited contexts. The model's effectiveness in producing high normalized gains (0.7325) suggests substantial return on instructional investment, making it an attractive option for elementary educators seeking evidence-based strategies to improve reading outcomes. The structured puzzle format provides clear activity frameworks that teachers can adapt to various curriculum topics, vocabulary sets, and difficulty levels, thereby enhancing instructional flexibility while maintaining pedagogical rigor.

Despite the study's robust findings, several limitations warrant acknowledgment. First, the quasi-experimental design utilizing intact classroom groups, while pragmatically necessary, introduces potential selection bias and limits causal inference strength compared to fully randomized controlled trials. Although pretest scores demonstrated group equivalence and statistical controls addressed this concern, unmeasured confounding variables related to classroom dynamics, teacher characteristics, or student motivation may have influenced outcomes. Future research should employ randomized controlled designs where feasible to strengthen causal claims.

Second, the study focused exclusively on reading ability as measured by a single researcher-developed questionnaire, which may not capture the full complexity of reading competence including comprehension depth, critical analysis, and application to varied text types. Future investigations should incorporate multiple assessment methods, including standardized reading tests, comprehension measures, and authentic reading tasks, to evaluate intervention effects across the multidimensional construct of reading literacy. Third, the six-week intervention period, while sufficient to demonstrate significant effects, precludes conclusions about long-term retention and transfer effects. Longitudinal studies examining whether Word Square-enhanced reading skills persist over time and generalize to novel contexts would substantially strengthen the evidence base.

In synthesis, this study affirms that the Word Square learning model represents an effective pedagogical innovation for enhancing elementary students' reading ability in Indonesian language

instruction. The convergence of statistically significant between-group differences, high normalized gain coefficients, and alignment with theoretical frameworks of active learning provides compelling evidence for the model's educational value. By demonstrating that structured, game-based puzzle activities can produce substantial reading improvements exceeding those achieved through conventional instruction, this research contributes empirically-grounded support for interactive, student-centered pedagogical approaches in literacy education. The findings encourage educators to embrace innovative instructional methods that strategically integrate engagement-enhancing elements with rigorous academic content, thereby creating learning environments that are simultaneously motivating and educationally effective. As educational systems increasingly prioritize evidence-based practice, interventions demonstrating robust effectiveness like the Word Square model merit serious consideration for broader implementation and continued refinement through systematic research.

## CONCLUSION

This quasi-experimental study provides robust empirical evidence that the Word Square learning model significantly enhances reading ability among fourth-grade elementary students in Indonesian language instruction. Students receiving Word Square instruction achieved substantially higher posttest scores ($M = 89.64$) compared to peers in conventional instruction ($M = 70.86$), with statistical significance ($t = 13.636$, $p < 0.001$) and high effectiveness (N-gain = 0.7325) confirming meaningful pedagogical impact. These findings affirm that game-based puzzle interventions incorporating active word recognition and contextual learning processes produce superior outcomes to traditional lecture-based approaches, thereby validating contemporary theories of cognitive engagement and constructivist learning in literacy education.

The study contributes to the evidence base supporting interactive, student-centered pedagogical innovations in elementary education, particularly within Indonesian language contexts where research on game-based literacy interventions remains limited. Theoretically, the results reinforce cognitive engagement frameworks demonstrating that active manipulation of linguistic elements enhances reading skill development. Practically, the Word Square model offers an accessible, scalable intervention requiring minimal technological resources, making it viable for diverse educational settings including resource-constrained contexts. Educators are encouraged to integrate puzzle-based activities systematically within reading curricula to enhance student engagement and learning outcomes.

However, several limitations warrant consideration. The quasi-experimental design using intact classroom groups, while pragmatically necessary, limits causal inference strength compared to fully randomized trials. The reliance on a single researcher-developed reading instrument may not capture the multidimensional nature of reading competence, and the six-week intervention period precludes conclusions about long-term retention. Future research should employ randomized controlled designs, incorporate multiple assessment methods including standardized measures and authentic reading tasks, conduct longitudinal investigations examining sustained effects and transfer to novel contexts, and explore moderating variables such as student characteristics and implementation fidelity that may influence intervention effectiveness across diverse populations and settings.

## REFERENCES

Adilah, N., Herlina, H., & Maariwuth, A. (2024). Meningkatkan Keaktifan Peserta Didik Melalui Media Kartu Huruf Dan Gambar Dengan Pendekatan Tarl Pada Mata Pelajaran Bahasa Indonesia Kelas 1a Sd Inpres 6 Lolu Palu. *Jurnal Koulutus*, *7*(2), 243-253. https://ejournal.kahuripan.ac.id/index.php/koulutus/article/view/1266

Ahmad, J., Rahmawati, D., & Anwar, R. B. (2020). Proses translasi representasi siswa dalam menyelesaikan permasalahan matematika yang berorientasi pada high order thinking skills. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *9*(3), 631-640. http://dx.doi.org/10.24127/ajpm.v9i3.2886

Ayuningtyas, N. M. D., Darsana, I. W., & Kristiantari, R. (2019). Pengaruh model pembelajaran Word Square terhadap penguasaan kompetensi pengetahuan IPA siswa kelas V. *Jurnal Pendidikan dan Pembelajaran IPA Indonesia*, *9*(1), 25-33. https://ejournal-pasca.undiksha.ac.id/index.php/jurnal_ipa/article/view/2910

Azizah, A., Awaliyah, B., & Lestari, S. (2022). Pengaruh Model Pembelajaran Word Square terhadap Hasil Belajar Tema 1 Sub Tema 1 Kelas 5 Di MI Nurul Huda. *MITRA PGMI Jurnal Kependidikan MI, 8*(1), 29–36. https://doi.org/10.46963/mpgmi.v8i1.441

Bang, H., & Siebert-Evenstone, A. (2025). The effectiveness of educational apps for early literacy learning: A systematic review and meta-analysis. *Educational Technology Research and Development*. Advance online publication.

Budiastuti, D., & Agustinus, B. (2018). *Validity and reliability study with analysis using NVIVO, SPSS and AMOS*. Mitra Media Discourse. https://files.core.ac.uk/download/pdf/187726085.pdf

Drost, E. A. (2011). Validity and reliability in social science research. *Education Research and perspectives*, *38*(1), 105-123. https://search.informit.org/doi/abs/10.3316/INFORMIT.491551710186460

Eng, C. M., Flynn, R. M., Thiessen, E. D., & Fisher, A. V. (2023). A literature review on the effects of exergames on executive function in youth. *Technology, mind, and behavior*, *4*(3), 10-1037. https://pmc.ncbi.nlm.nih.gov/articles/PMC10756422/

Fahyuni, E. F., & Bandono, A. (2015). Story media development illustrated as effort increase ability read student school base. *Halaqa, 14*(1), 75–89. http://eprints.umsida.ac.id/1123/

Fauziah, A. (2024). Pentingnya kemampuan membaca kritis di era informasi digital. *Jurnal Citra Pendidikan*, *4*(2), 1685-1689. https://doi.org/10.38048/jcp.v4i2.3527

Fisnawati, F. (2024). Membaca Sebagai Sarana Mengasa Imajinasi Dan Kreativitas Anak Sekolah Dasar. *Journal Sultra Elementary School*, *5*(1), 745-757. https://jurnal.yayasanmeisyarainsanmadani.com/index.php/JSES/article/view/404

Fredricks, J. A., Filsecker, M., & Lawson, M. A. (2016). Student engagement, context, and adjustment: Addressing definitional, measurement, and methodological issues. *Learning and Instruction, 43*, 1–4. https://doi.org/10.1016/j.learninstruc.2016.02.002

Gray, C. D. (2023). *SPSS statistics: A straightforward guide* (2nd ed.). Red Globe Press.

Gustiawati, R., Arief, D., & Zikri, A. (2020). Pengembangan bahan ajar membaca permulaan dengan menggunakan cerita fabel pada siswa sekolah dasar. *Jurnal basicedu*, *4*(2), 355-360. https://doi.org/10.31004/basicedu.v4i2.339

Hall, C., Dahl-Leonard, K., Cho, E., Solari, E. J., Capin, P., Conner, C. L., Henry, A. R., Cook, L., Hayes, L., Vargas, I., Richmond, C. L., & Kehoe, K. F. (2023). Forty years of reading intervention research for elementary students with or at risk for dyslexia: A systematic review and meta-analysis. *Reading Research Quarterly, 58*(4), 500–522. https://doi.org/10.1002/rrq.477

Handayani, L. T. (2022). *Implementasi Teknik Analisis Data Kuantitatif*. PT. Scifintech Andrew Wijaya.

Harianto, E. (2020). Keterampilan membaca dalam pembelajaran bahasa. *Didaktika: Jurnal Kependidikan*, *9*(1), 1-8. https://www.jurnaldidaktika.org/contents/article/view/2

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Laal, M., & Ghodsi, S. M. (2012). Benefits of collaborative learning. *Procedia - Social and Behavioral Sciences, 31*, 486–490. https://doi.org/10.1016/j.sbspro.2011.12.091

Lauterbach, M., & Zipke, M. (2024). Wordling with Elementary Students: Developing Discrete Literacy Skills through Puzzles and Word Games. *The Reading Teacher*, *78*(3), 195-201. https://doi.org/10.1002/trtr.2362

Listini, L., Parwanti, S., Gunawan, G., Mustofa, M., Danto, D., & Refson, R. (2022). Peningkatan Kemampuan Membaca Pemahaman Melalui Model Pembelajaran Word Square Pada Kelas

VIII. *Jurnal Bindo Sastra*, *5*(2), 47-52. http://jurnal.um-palembang.ac.id/bisastra/article/view/4277

Liu, T. Y., & Chen, Y. L. (2013). The effects of game-based learning on students' learning performance and motivation in mobile learning. *International Journal of Distance Education Technologies, 11*(2), 36–50. https://doi.org/10.1016/j.sbspro.2013.10.430

Lyon, G. R. (1998). *Overview of reading and literacy initiatives*. National Institute of Child Health and Human Development. https://www.nichd.nih.gov/publications

National Assessment of Educational Progress. (2020). *NAEP reading: National achievement level results*. National Center for Education Statistics. https://www.nationsreportcard.gov/reading/nation/achievement

Orawiwatnakul, W. (2013). Crossword puzzles as a learning tool for vocabulary development. *Electronic Journal of Research in Educational Psychology, 11*(2), 413–428. https://doi.org/10.14204/ejrep.30.12186

Rahma, S. N., Deyanti, F., & Fitriyah, M. (2024). Peran membaca dalam meningkatkan kemampuan berpikir kritis di kalangan mahasiswa. *DHARMA ACARIYA NUSANTARA: Jurnal Pendidikan, Bahasa dan Budaya*, *2*(1), 100-108. https://doi.org/10.37802/society.v2i1.174

Ronimus, M., & Lyytinen, H. (2015). Is school a better environment than home for digital game-based learning? The case of GraphoGame. *Human Technology, 11*(2), 123–147. https://doi.org/10.17011/ht/urn.201511113637

Senechal, M., & LeFevre, J. A. (2002). Parental involvement in the development of children's reading skill: A five-year longitudinal study. *Child Development, 73*(2), 445–460. https://doi.org/10.1111/1467-8624.00417

Sugiarti, U. (2012). Pentingnya pembinaan kegiatan membaca sebagai implikasi pembelajaran bahasa indonesia. *Basastra*, *1*(1), 54467. https://jurnal.unimed.ac.id/2012/index.php/basastra/article/view/201

Taherdoost, H. (2016). Validity and reliability of the research instrument; How to test the validation of a questionnaire/survey in a research. *International Journal of Academic Research in Management, 5*(3), 28–36. https://doi.org/10.2139/ssrn.3205040

Vaughn, S., & Fletcher, J. M. (2020). Identifying and teaching students with significant reading problems. In R. E. O'Connor & P. F. Vadasy (Eds.), *Handbook of reading interventions* (pp. 206–236). Guilford Press.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., Donahue, J., & Garon, T. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study. *Developmental Psychology, 33*(3), 468–479. https://doi.org/10.1037/0012-1649.33.3.468

Wanzek, J., Stevens, E. A., Williams, K. J., Scammacca, N., Vaughn, S., & Sargent, K. (2018). Current evidence on the effects of intensive early reading interventions. *Journal of Learning Disabilities, 51*(6), 612–624. https://doi.org/10.1177/0022219418775110

Wanzek, J., Wexler, J., Vaughn, S., & Ciullo, S. (2013). Reading interventions for struggling readers in the upper elementary grades: A synthesis of 20 years of research. *Reading and Writing, 23*(8), 889–912. https://doi.org/10.1007/s11145-009-9179-5