

Integration of Augmented Reality in Natural Appearance Learning to Improve The Visual-Spatial Abilities of Elementary School Students

Nanis Regina Choerunnisa, Devi Afriyuni Yonanda, Eka Nurhidayat

Universitas Majalengka, Indonesia

Email: nanis.r.choerunnisa@gmail.com, deviyonanda1990@gmail.com,
ekanurhidayat@unma.ac.id

Keywords:

Spatial Ability; Cognitive Intervention; Spatial Visualization; STEM Education; Learning Effectiveness

ABSTRACT

This research aimed to investigate the effectiveness of a structured intervention program in enhancing spatial ability among twenty-one participants using a pretest-posttest design. Spatial ability was measured across four dimensions: Spatial Visualization, Spatial Orientation, Spatial Relations, and Mental Imagery. The intervention consisted of systematic training activities designed to develop multiple facets of spatial thinking. Results demonstrated substantial improvements in spatial ability following the intervention, with mean scores increasing from forty-eight point eight-nine at pretest to seventy-seven point four-six at posttest. This improvement represents a fifty-eight percent increase and yielded an exceptionally large effect size of four point three-eight, substantially exceeding the conventional threshold for large effects. The normalized gain analysis revealed a mean value of zero point five-six-five, classified as moderate effectiveness. Notably, the distribution of performance categories shifted dramatically, with all participants successfully transitioning from low to moderate or high performance levels. The paired samples t-test confirmed statistically significant differences between pretest and posttest scores. Improvements were balanced across all four spatial ability dimensions, ranging from fifty-two point nine percent to sixty-three point two percent, indicating that the intervention comprehensively addressed multiple aspects of spatial cognition rather than focusing narrowly on specific components. These findings provide robust evidence that spatial ability is a malleable cognitive competency that can be systematically developed through targeted instruction, with important implications for educational practice, particularly in STEM fields where spatial thinking is fundamental to success.

INTRODUCTION

The development of digital technologies has fundamentally changed the global education landscape, creating transformative opportunities in the learning process in the 21st century (Zou et al, 2025; Kalyani, L. 2024). According to Alenezi et al (2023), the integration of technology in education is no longer an option but an urgent need to prepare a generation that is able to compete in the digital era. In Indonesia, the implementation of the Independent Curriculum emphasizes the importance of technology-based learning and 21st-century competency development, but the reality on the ground shows significant gaps between policy and practice, especially in the use of innovative technologies such as Augmented Reality (AR) for science learning at the elementary school level.

Natural Sciences (IPA) learning, especially natural appearance materials, faces complex pedagogical challenges in elementary schools (Andriani, 2025; Armawati et al., 2025; Rosa et al., 2026; Syahputra et al., 2022; Zakirman et al., 2022). Natural appearances are abstract concepts that include an understanding of geographical formations, geological structures, and natural phenomena that are difficult to visualize through conventional methods (Pilehvar, A., 2021; Maudidah, et al, 2025). Research conducted by Wahyuningsih and Budiningsih (2022) identified that 68% of grade IV students had difficulty understanding the concept of natural appearances due to the limitations of learning media that were able to describe three-dimensional objects realistically. Traditional learning methods that rely on two-dimensional images, textbooks, and verbal explanations have proven to be less effective in building students' spatial understanding of landscapes such as mountains, valleys, plateaus, and other geological formations. This condition is exacerbated by limited access to field learning or visits to geographical locations that are the object of learning, especially in schools with limited resources.

Visual-spatial skills are fundamental cognitive competencies that play a crucial role in learning science, mathematics, and various other disciplines. According to Gardner's (1983, in Khairah et al., 2025) theory of compound intelligence, visual-spatial abilities include the capacity to visualize objects, understand spatial relationships, manipulate mental images, and navigate three-dimensional space. Educational neuroscience research by Uttal et al. (2013) and Gagnier, K., Holochwost, S., & Fisher, K. (2021). revealed that visual-spatial skills developed from an early age are positively correlated with academic achievement in the field of STEM (Science, Technology, Engineering, and Mathematics) at the next level of education. In the context of natural appearance learning, visual-spatial abilities allow students to understand scales, proportions, orientations, and relationships between geographic elements that cannot be directly observed. However, the results of the 2019 TIMSS (Trends in International Mathematics and Science Study) study show that the visual-spatial abilities of Indonesian students are below the international average, indicating the need for more effective pedagogical interventions to develop these competencies since primary education.

Augmented Reality (AR) has emerged as an innovative learning technology that promises solutions to the challenge of visualizing abstract concepts in education. AR enables the superimposition of three-dimensional virtual objects into real environments, creating an immersive and interactive learning experience. Several empirical studies have demonstrated the effectiveness of AR in educational contexts. Cheng and Tsai (2013) found that AR-based learning increased students' motivation and concept understanding of astronomy material by 34% compared to conventional methods. Similarly, a meta-analysis study by Garzón and Acevedo (2019) of 61 studies showed that AR had a moderate to large positive effect (effect size = 0.68) on student learning outcomes at various levels of education. In the learning of geography and earth science, research by Ibanez et al. (2020) and Schmidt, R., & Stumpe, B. (2025) confirms that AR facilitates a better spatial understanding of complex geographic phenomena by providing multi-perspective representation and manipulation of virtual objects in real-time. These findings indicate the significant potential of AR as a learning medium that can overcome the limitations of traditional methods in teaching spatial concepts.

State of the art in AR research for science education shows rapid development in the last five years, with a focus on developing more intuitive, accessible, and pedagogical

applications. Marker-based AR technology has evolved into markerless AR that utilizes SLAM (Simultaneous Localization and Mapping) technology for more precise placement of virtual objects. Research by Akçayır and Akçayır (2017) identified that the latest generation of AR applications integrate gamification elements, adaptive scaffolding, and collaborative features to improve learning engagement and effectiveness. In Indonesia, several early studies have explored the application of AR in science learning, such as research by Mustaqim and Kurniawan (2017) on solar system materials and Fauzi, Y. M. (2022) on digestive system materials. However, AR development is still focused on biology and physics content, with very limited applications for learning physical geography, particularly natural appearances. Furthermore, the majority of existing studies measure the impact of AR on general cognitive learning outcomes, while its influence on specific abilities such as visual-spatial has not been explored in depth, especially in the context of elementary school students in Indonesia.

Although the potential of AR in education has been well documented, there is a significant research gap in several aspects. First, the majority of AR research in science education focuses on secondary and higher education, while its application at the elementary school level, especially grade IV, which is a critical period of cognitive skill development, is still very limited. Second, there has not been a comprehensive study that specifically develops and tests AR media for natural appearance learning with a focus on improving the visual-spatial abilities of elementary school students. Third, previous studies tend to adopt generic or commercial AR applications that are not specifically designed based on the curriculum needs and characteristics of Indonesian students. Fourth, the practicality and sustainability aspects of AR implementation in elementary schools with various limitations of technological infrastructure have not been systematically studied. These gaps create an urgent need for research that not only develops contextual and pedagogical AR media, but also evaluates its effectiveness on visual-spatial capabilities and the practicality of its implementation in basic education settings in Indonesia.

The urgency of this research is increasing considering several convergent factors that are happening today. First, the implementation of the Independent Curriculum requires a learning transformation that is more student-centered, technology-based, and competency-oriented. Second, smartphone and tablet penetration in Indonesia has reached 62.8% in 2023 (APJII, 2023), creating a technology infrastructure that enables the implementation of AR more widely. Third, the COVID-19 pandemic has accelerated the adoption of educational technology and improved the digital literacy of teachers and students, creating the right momentum to integrate innovative technologies such as AR. Fourth, Indonesia's challenges in the global STEM competition urge the development of effective pedagogical interventions to improve basic competencies such as visual-spatial skills since primary education. This research is significant because it can provide practical contributions in the form of innovative learning media that can be implemented directly in schools, as well as theoretical contributions in enriching the understanding of AR mechanisms in developing the visual-spatial abilities of elementary school age students.

The novelty or novelty of this research lies in several fundamental aspects that distinguish it from previous studies. First, this study developed an AR media specifically designed for learning the appearance of nature based on an in-depth analysis of the Basic

Competencies of the Indonesian curriculum, the cognitive characteristics of grade IV students, and the principles of instructional design. Second, this study uses a systematic development approach with the ADDIE (Analysis, Design, Development, Implementation, Evaluation) model which ensures the pedagogical and technical validity of the developed media. Third, in contrast to previous studies that measured the impact of AR on general cognitive learning outcomes, this study specifically measured the improvement of visual-spatial abilities using valid and reliable instruments, providing a deeper understanding of the cognitive mechanisms that occur. Fourth, this study evaluates not only the effectiveness but also the practical aspects of AR implementation in the real context of Indonesian elementary schools, resulting in recommendations that are applicable to the scaling up of this technology. This innovation makes this research important because it can produce a contextual, scientifically validated, and practical AR media development model to improve the quality of science learning in elementary schools.

Based on the context of the problems, research gaps, and urgency that have been presented, this study aims to develop and test Augmented Reality-based "KENAL" (Natural Appearance) learning media specifically designed to improve the visual-spatial abilities of grade IV elementary school students. This research will answer fundamental questions about how the process of developing valid and practical AR media, how effective the media is in improving visual-spatial skills, and how practical its implementation is in the context of actual learning in Indonesian elementary schools. Thus, this research is expected to make a significant contribution both theoretically and practically in the field of educational technology, especially in the development of innovative learning media for basic education.

METHOD

This study used a quantitative approach with an experimental method to test the effectiveness of Augmented Reality-based "KENAL" (Natural Appearance) learning media in improving the visual-spatial abilities of elementary school students. The experimental method was selected to examine the causal relationship between the use of AR-based learning media and students' visual-spatial abilities under controlled conditions (Creswell & Creswell, 2018).

The research design employed a quasi-experimental pretest–posttest control group design. This design was chosen due to limitations in conducting full randomization, as students were already organized into pre-existing classes within schools (Shadish et al., 2002). The quasi-experimental design still allowed adequate control of internal validity through the use of a control group as a comparison group (Campbell & Stanley, 2015).

The research design can be described as follows:

$$O_1 \rightarrow X \rightarrow O_2$$

Description:

O_1 = Pretest of visual-spatial ability before treatment

X = Treatment using Augmented Reality-based learning media "KNOW"

O_2 = Posttest of visual-spatial ability after treatment

Students received learning material on natural appearances using AR media "KENAL" for six meetings with a total duration of 18×35 minutes (630 minutes). Learning follows a

specially designed Learning Implementation Plan (RPP) by integrating interactive features of AR to facilitate the development of students' visual-spatial abilities.

The variables in this study consist of: (1) independent variables, namely Augmented Reality-based learning media "KENAL"; (2) dependent variables, namely students' visual-spatial abilities measured through standardized tests; and (3) control variables include teachers, learning materials, learning duration, and learning implementation time that are maintained consistently during the research.

Population and Sample

The population in this study is all grade IV students of Rajagaluh I State Elementary School for the 2024/2025 school year which totals 21 students. The selection of class IV is based on the consideration that: (1) the material of natural appearances is listed in the Basic Competencies of class IV for even semesters; (2) grade IV students are at the stage of concrete operational cognitive development towards formal operations (7-11 years) according to Piaget's theory, so that they are able to understand more complex spatial concepts; and (3) grade IV students have adequate basic digital literacy to operate AR applications on mobile devices.

Given the limited population, this study uses a saturated sampling technique (total sampling), which is a sample determination technique where all members of the population are used as research samples (Sugiyono, 2019). Thus, the research sample is all grade IV students of SD Negeri Rajagaluh I which totals 21 students. The saturated sampling technique was chosen because: (1) the population is relatively small (< 30); (2) the researcher wants to make generalizations with very small errors; and (3) all students in one class have relatively homogeneous characteristics so that they are representative to be used as samples.

The characteristics of the research sample are as follows:

1. Age: Students aged 9-11 years who correspond to the characteristics of cognitive development grade IV of primary school
2. Education status: Registered as an active student in grade IV of SD Negeri Rajagaluh I for the 2024/2025 school year
3. Learning experience: Have not received material on natural appearances in the current semester
4. Physical condition: Lack of significant visual impairment that may inhibit interaction with AR media
5. Technology access: Able to operate a smartphone or tablet device (for students who do not have a personal device, device provided by the school or researcher)
6. Willingness to participate: The student's parents/guardians have given informed consent for their child's participation in the study

The composition of the sample by gender was: 11 male students (52.4%) and 10 female students (47.6%). Complete demographic data on socio-economic background, initial academic ability, and experience of using technology were collected for the purposes of descriptive analysis and identification of sample characteristics.

Although the study design did not use an external control group, to reinforce the internal validity, the study applied several control strategies: (1) repeated measurements (pretest-posttest) to compare pre- and post-treatment abilities; (2) time control by carrying out pretest and posttest in the same time interval for all students; (3) standardization of learning

procedures using structured lesson plans; and (4) documentation of the learning process through observation to ensure the fidelity of the implementation of AR media

Research Instruments

The main instrument in this study is the Visual-Spatial Ability Test which was developed specifically to measure students' visual-spatial abilities in the context of natural appearance materials. The development of the instrument refers to the theory of visual-spatial ability from Lohman (1996) and is adapted to the characteristics of grade IV elementary school students

Instrument Specifications

The test is multiple-choice with 30 question items covering four dimensions of visual-spatial ability:

1. Spatial Visualization - 8 items: the ability to manipulate, rotate, or transform the mental image of a three-dimensional geographic object (e.g., imagining the shape of a mountain from various points of view).
2. Spatial Orientation - 8 items: the ability to understand the relative position and direction of an object or location in space (e.g., determining the direction of a river flow based on a topographic map).
3. Spatial Relations - 7 items: the ability to understand the relationships between elements in space and object configurations (e.g., understanding the relationship between altitudes, valleys, and mountains).
4. Mental Imagery - 7 items: the ability to form and manipulate mental images of geographical objects or phenomena (e.g., imagining the process of forming a valley or lake).

Each question item has 4 answer options (a, b, c, d) with one correct answer. The maximum score is 30 points with the scoring system: correct answer = 1 point, wrong answer = 0 points. For ease of interpretation, the raw score is converted to a scale of 0-100 using the formula:

$$\text{Value} = (\text{Acquisition Score} / \text{Maximum Score}) \times 100$$

The questions are equipped with clear and colored pictures, diagrams, and illustrations to facilitate students' understanding of the spatial context being asked. Language and sentence structure are adjusted to the level of readability of grade IV elementary school students.

Instrument Validity

The validity of the instrument is tested through three stages:

1. Content Validity: Conducted through expert judgment by 5 experts consisting of 2 science education lecturers, 2 educational psychology lecturers, and 1 experienced grade IV science teacher. Content validity is measured using Content Validity Ratio (CVR) and Content Validity Index (CVI). An item is declared valid if the CVR ≥ 0.62 (for 5 members) and the CVI ≥ 0.78 (Lawshe, 1975).
2. Construct Validity: Tested using Confirmatory Factor Analysis (CFA) with the help of AMOS 24 software to confirm whether the test items measure four dimensions of visual-spatial ability according to the theoretical construct. The model is declared fit if it meets the criteria: $\chi^2/df < 3$, CFI ≥ 0.90 , TLI ≥ 0.90 , RMSEA ≤ 0.08 , and SRMR ≤ 0.08 (Hu & Bentler, 1999).

3. Validity of Question Items (Item Validity): Tested through instrument trials on 35 grade IV students from other elementary schools who have similar characteristics to the research sample. The validity of the item was calculated using point-biserial correlation with the criterion of $r_{pbis} \geq 0.30$ and significant at $p < 0.05$. Items that don't meet the criteria are revised or eliminated.

Instrument Reliability

The reliability of the instrument was tested using the Kuder-Richardson 20 method (KR-20) which is suitable for multiple-choice tests with dichotomy assessments. The instrument is declared reliable if the KR-20 coefficient is ≥ 0.70 (Fraenkel et al., 2012). Based on the results of the test on 35 students, a reliability coefficient of 0.84 was obtained which showed that the instrument had good internal consistency (high category).

In addition, a quality analysis of question items was carried out including:

- Difficulty Index: Calculated using the formula $P = B/N$, with a criterion of $0.30 \leq P \leq 0.70$ for questions with medium difficulty
- Discriminating Power (Discrimination Index): Calculated using the formula $D = (BA/JA) - (BB/JB)$, with criteria $D \geq 0.30$ to ensure that items can distinguish between high-ability and low-ability students

Of the 35 question items tested, 30 items met the criteria of validity, reliability, difficulty, and good differentiation, so that they were used as the final research instrument.

Supporting Instruments

In addition to the main instruments, the study also used:

- Learning Observation Sheet: To document the learning process and students' interaction with AR media
- Documentation Sheet: To record learning activities through photos and videos
- Field Notes: To record important events during the implementation of AR media

Research Procedure

The research was carried out through six systematic stages during the period February - April 2025:

- Phase 1: Preparation (2 weeks) - Processing of research permits, coordination with schools, socialization to parents/guardians of students, teacher training on the use of AR media for 2 days (8 hours), preparation of lesson plans for 6 meetings, and technical trials of AR applications in the classroom.
- Stage 2: Pretest (1 week) - Implementation of visual-spatial ability pretest for all 21 students with a working time of 60 minutes. Supervision is carried out by classroom teachers accompanied by researchers to ensure data integrity.
- Stage 3: Analysis of Pretest Results - Entry and cleaning data, descriptive analysis to determine the student's initial ability profile, and documentation of initial ability characteristics as a baseline.
- Stage 4: Treatment (3 weeks, 6 meetings) - Learning of natural appearance materials using AR media "KENAL" with a duration of 3×35 minutes each (105 minutes per meeting). Students work in pairs (2 students per device) to encourage collaboration. The material includes: (1) Meeting 1: Introduction to the concept of the appearance of the terrestrial nature (mountains and mountains); (2) Meeting 2: Highlands and lowlands; (3)

Meeting 3: Natural appearance of waters (rivers); (4) Meeting 4: Lakes, beaches, and seas; (5) Meeting 5: Analysis of the relationship between natural appearances; (6) Meeting 6: Implementation and evaluation of learning.

- Stage 5: Posttest (1 week) - Posttest of visual-spatial ability 3 days after the last meeting using an instrument identical to the pretest. The implementation time, instructions, and conditions of supervision are made the same as the pretest.
- Stage 6: Data Analysis and Reporting (2 weeks) - Entry and cleaning of posttest data results, descriptive and inferential statistical analysis using IBM SPSS Statistics version 26, interpretation of results, and preparation of research reports.

Data Analysis Techniques

Data analysis in this study used descriptive statistics and inferential statistics to test the effectiveness of AR media "KNOW" on improving students' visual-spatial skills.

Descriptive Statistical Analysis

Descriptive statistics were used to describe the characteristics of the pretest and posttest data, including the calculation of the centering size (mean, median, mode), the size of the spread (standard deviation, variance, range, minimum and maximum values), the distribution of frequencies by ability category (low: 0-59, medium: 60-79, high: 80-100), and data visualization using histograms, boxplots, and bar charts.

Analysis Prerequisites Test

Before conducting the hypothesis test, a normality test was carried out using the Shapiro-Wilk test because the sample size was small ($n = 21 < 50$). The data is considered to be normally distributed if the significance value (p -value) > 0.05 . If the data is not normally distributed, data transformation can be performed or use the non-parametric Wilcoxon Signed-Rank Test as an alternative.

Inferential Statistical Analysis

Hypothesis testing was performed at a significance level of $\alpha = 0.05$ using the Paired Samples t-test (if the data is normally distributed) to test the difference in the mean of pretest and posttest scores. Hypothesis: $H_0: \mu_{\text{posttest}} \leq \mu_{\text{pretest}}$ (no increase); $H_1: \mu_{\text{posttest}} > \mu_{\text{pretest}}$ (there is a significant improvement). Formula: $t = (\bar{D} - \mu_0) / (SD / \sqrt{n})$, where \bar{D} is the mean of the difference (gain score), $\mu_0 = 0$, SD is the standard deviation of the difference, and $n = 21$. Criteria: if the p -value < 0.05 and $t_{\text{hitung}} > t_{\text{tabel}}$, then H_0 is rejected (there is a significant increase).

Effect Size and N-Gain Analysis

To measure the magnitude of the effect of the treatment, Cohen's d is calculated with the formula: $d = (M_{\text{posttest}} - M_{\text{pretest}}) / SD_{\text{pooled}}$. Interpretation: $d < 0.2$ (very small effect); $0.2 \leq d < 0.5$ (small effect); $0.5 \leq d < 0.8$ (moderate effect); $d \geq 0.8$ (large effect). The effectiveness of the increase was measured using normalized gain (N-gain) with the formula of Hake (1999): $N\text{-gain} = (\text{Posttest Score} - \text{Pretest Score}) / (\text{Maximum Score} - \text{Pretest Score})$. Interpretation: $N\text{-gain} < 0.30$ (low); $0.30 \leq N\text{-gain} < 0.70$ (moderate); $N\text{-gain} \geq 0.70$ (high).

To understand the dimensions of the visual-spatial ability that were most affected, separate analyses were performed for all four dimensions (spatial visualization, spatial orientation, spatial relations, and mental imagery) using the same procedure. All statistical analysis was conducted using IBM SPSS Statistics version 26.

AR learning media "KENAL" is declared effective if it meets the following criteria: (1) statistical significance with a p-value of < 0.05 ; (2) Cohen's effect size $d \geq 0.5$; (3) average N-gain ≥ 0.30 ; and (4) at least 70% of students experience an increase in scores.

RESULT AND DISCUSSION

This study investigated the effectiveness of a spatial ability intervention program involving 21 participants. The study used a pretest-posttest design to measure changes in spatial abilities in four dimensions: Spatial Visualization (VS), Spatial Orientation (OS), Spatial Relations (RS), and Mental Imagery (MI).

Descriptive Statistics

Table 1 presents descriptive statistics for pretest, posttest, and normalized gain (N-Gain) scores. The mean pretest score was 48.89 (SD = 6.29), indicating that participants generally had below-average spatial abilities before the intervention. After the intervention, the average posttest score increased substantially to 77.46 (SD = 6.74), which represented a significant improvement in spatial ability performance.

The normalized gain score, which takes into account the potential ceiling effect and provides a standard measure of increase, yields an average of 0.565 (SD = 0.101). These N-Gain values fall into the moderate category ($0.3 < \text{N-Gain} < 0.7$), indicating that the intervention results in a significant improvement in spatial ability. The pretest (33.33-60.00) and posttest (63.33-90.00) score ranges indicated variability in baseline and post-intervention ability levels, although all participants showed improvement.

Table 1. Descriptive Statistics of Spatial Ability Score

Statistics	Pretest	Posttest	N-Gain
Red	48,89	77,46	0,565
Median	50,00	76,67	0,563
Std Dev	6,29	6,74	0,101
Min	33,33	63,33	0,375
Max	60,00	90,00	0,750

Performance Category Distribution

The distribution of participants in the performance category shows a dramatic shift from pretest to posttest (Table 2). In the pretest, the majority of participants (95.2%, $n=20$) were in the low category (0-59), with only one participant (4.8%) achieving a moderate score (60-79). None of the participants reached the high category (80-100) at baseline.

After the intervention, this distribution changed markedly. In the posttest, there were no participants left in the low category. In contrast, 57.1% ($n=12$) achieved a moderate score and 42.9% ($n=9$) achieved the high category. This category redistribution showed that the intervention not only improved the average score but also fundamentally changed the distribution of performance levels, moving all participants out of the lowest performance levels.

Table 2. Performance Category Distribution

Categories	Pretest(s)	Pretest (%)	Posttest(s)	Post-test (%)
Height (80-100)	0	0,0%	9	42,9%
Medium (60-79)	1	4,8%	12	57,1%
Low (0-59)	20	95,2%	0	0,0%

Analysis Per Dimension of Spatial Ability

Table 3 shows improvements in each of the four dimensions of spatial ability. All dimensions showed substantial improvements, with percentage increases ranging from 52.9% to 63.2%. Spatial Orientation showed the highest relative increase (63.2%), with the average score increasing from 3.62 to 5.90 from a maximum of 8 points, representing a gain of 2.29 points. This is followed by Spatial Relations (59.7% increase) and Mental Imagery (58.9% increase).

Spatial Visualization, despite showing the lowest relative percentage improvement (52.9%), still showed a meaningful gain, with the score increasing from 4.14 to 6.33 (2.19 points). Similar magnitudes of improvement across all dimensions (ranging from 2.05 to 2.29 points) suggest that the intervention effectively addresses multiple facets of spatial ability rather than narrowly focusing on a single dimension.

Table 3. Increases per dimension of spatial ability

Dimensions	Max Score	Red Pretest	Mean Posttest	Gain	Gain (%)
Spatial Visualization	8	4,14	6,33	2,19	52,9%
Spatial Orientation	8	3,62	5,90	2,29	63,2%
Spatial Relations	7	3,43	5,48	2,05	59,7%
Mental Imagery	7	3,48	5,52	2,05	58,9%

Hypothesis Test

Before performing the main analysis, a normality test is performed using the Shapiro-Wilk test to verify the parametric test assumptions. Pretest scores ($W = 0.940$, $p = 0.217$) and posttest ($W = 0.965$, $p = 0.616$) showed normal distribution, as indicated by p values exceeding the significance threshold of 0.05. These results confirm that the data meet the assumptions required to perform paired samples t-test.

Paired samples t-test showed a statistically significant difference between pretest and posttest scores ($t = 27.39$, $p < 0.001$). This very small p-value provides strong evidence against the null hypothesis of no difference, indicating that the observed increase in spatial ability scores is not due to chance. The effect size, measured by Cohen's d, is 4.38, which substantially exceeds Cohen's conventional threshold of 0.8 for large effects. This extraordinarily large effect size suggests that the intervention has a profound practical impact on participants' spatial abilities.

Normalized gain analysis corroborates these findings, with an average N-Gain of 0.565 which falls into the category of moderate effectiveness. According to the Hake classification, an N-Gain value between 0.3 and 0.7 indicates a moderate learning gain. Although it did not reach a high gain threshold ($N\text{-Gain} > 0.7$), the moderate classification still represented a substantial and educationally meaningful improvement, especially given that all participants showed positive gains.

The findings of this study provide strong evidence about the effectiveness of spatial ability intervention programs. The substantial improvement observed on all dimensions of spatial ability measured, combined with a very large effect size (Cohen's $d = 4.38$), suggests that the intervention successfully improved the participants' spatial thinking capabilities. These results contribute to a growing body of literature demonstrating that spatial abilities are malleable and can be systematically developed through targeted instruction.

Interpretation of Key Findings

The average increase of 28.57 points (from 48.89 to 77.46) represents more than half of the standard deviation of the shift in the score distribution. More importantly, the shift in category in performance levels demonstrates not only statistical significance but also practical significance. The elimination of the low-performance category in the posttest indicated that the intervention successfully brought all participants to at least moderate levels of performance, with almost half achieving high performance. Uniform improvements across the sample showed that the intervention was accessible and effective for participants with varying baseline abilities.

A balanced increase in all four dimensions of spatial ability is noteworthy. Although Spatial Orientation shows the highest relative gain (63.2%), the difference between dimensions is relatively modest, with all improvements falling within the range of 11 percentage points. This pattern suggests that interventions address spatial abilities comprehensively rather than emphasizing certain dimensions at the expense of others. Integrated development in the dimensions of visualization, orientation, relationships, and mental imagery can contribute to more robust and transferable spatial thinking skills.

The classification of moderate N-Gain (mean = 0.565) requires careful interpretation. While this is below the high gain threshold (0.7), it represents a substantial improvement given the baseline performance level. The N-Gain metric takes into account the ceiling effect taking into account the maximum possible increase for each participant, making it a conservative estimate of effectiveness. In addition, the consistency of improvement in participants (SD = 0.101) suggests that average moderate gain reflects systematic learning rather than extreme scores of a subset of participants.

Implications for Practice

The results of the study have several important implications for educational practice. First, the demonstration of the malleability of spatial ability challenges the idea that spatial thinking is a fixed trait. Educators should view spatial abilities as developable competencies that can be systematically improved through appropriate instruction. This perspective is particularly important given the documented relationship between spatial ability and success in STEM fields, suggesting that spatial training can play a role in improving STEM educational outcomes and potentially reducing achievement gaps.

Second, a comprehensive improvement in multiple dimensions of spatial ability suggests that intervention programs should adopt a multifaceted approach rather than focusing narrowly on a single aspect of spatial thinking. The balanced gains observed in this study indicate that handling visualization, orientation, relationships, and imagery together can result in more robust improvements than isolated training on individual components. These

findings support the development of an integrated curriculum that exposes students to a variety of spatial challenges and tasks.

Third, the universal improvement in all participants, moving everyone out of the low-performance category, demonstrates that spatial interventions can be effective at different levels of baseline ability. This suggests that such programs can be implemented widely without the need for extensive screening or capability-based tracking, potentially making spatial training more accessible and equitable.

Limitations and Future Research

Some limitations should be considered when interpreting these findings. First, the absence of a control group limits causal inferences about the effects of the intervention. Although the large effect size and statistical significance provide strong evidence of change, alternative explanations such as practice effect, maturation, or regression to mean cannot be definitively ruled out. Future research should use experimental designs with control groups, ideally by random assignment, to reinforce causal claims about the effectiveness of interventions.

Second, the study did not address the observed increased durability. Posttests were administered immediately after the intervention, leaving questions about whether the gains persisted over time with no answers. Long-term follow-up assessments will provide valuable information about the lasting impact of spatial training and whether periodic refresher sessions may be necessary to sustain improvement.

Third, although the study demonstrated an improvement in spatial ability test performance, the study did not examine whether these gains transferred to real-world tasks or academic performance in domains that rely on spatial thinking. Future research should investigate whether spatial training leads to measurable improvement in areas such as math achievement, science learning, or engineering problem-solving. Such transfer studies would provide stronger evidence for the practical value of spatial ability interventions.

Fourth, the relatively small sample size ($n = 21$) and the potential lack of diversity within the participant pool may limit the generalizability of the findings. Larger-scale studies with more diverse populations will help determine whether similar improvements can be achieved across different age groups, educational backgrounds, and cultural contexts. In addition, studies that examine potential moderators of intervention effectiveness—such as baseline ability levels, motivation, or learning preferences—may inform the development of more targeted and personalized training approaches.

CONCLUSION

Spatial ability is a fundamental cognitive competency that plays an important role in various domains of life, from everyday navigation to complex problem-solving in STEM fields. This research shows that these abilities can be significantly developed through structured interventions, opening up opportunities to enhance individual cognitive potential and reduce disparities in STEM career preparation.

With a very large effect size (4.38) and a universal improvement in all participants, these findings provide optimism that we can systematically improve the spatial abilities of the population through well-designed education. However, further research with more rigorous

methodologies and investigations on durability and transfer effects is needed to fully understand the potential and limitations of spatial capability interventions. This research is an important step in understanding the development of spatial abilities, but there are still many questions that need to be answered. Continued research effort, combined with practical implementation in educational settings, will help us optimize approaches to developing these critical cognitive competencies and empower more individuals to succeed in a world that increasingly requires strong spatial thinking skills

REFERENCES

- Akçayır, M., & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review*, 20, 1–11. <https://doi.org/10.1016/j.edurev.2016.11.002>.
- Alenezi, M., Wardat, S., & Akour, M. (2023). The Need of Integrating Digital Education in Higher Education: Challenges and Opportunities. *Sustainability*. <https://doi.org/10.3390/su15064782>.
- Andriani, N. (2025). Activities, resources, and learning environment: a case study of science learning at sekolah alam. *Jurnal Penelitian Pendidikan IPA*, 11(1), 1075–1091.
- APJII. (2023). Indonesian internet survey report 2023. Indonesian Internet Service Providers Association. <https://survei.apjii.or.id/>
- Armawati, A., Vebrianto, R., Habibi, M., & Radeswandri, R. (2025). Overcoming Difficult Science Topics through Alternative Media in Elementary Education. *Journal of Islamic Education Students (JIES)*, 5(1), 266–276.
- Campbell, D. T., & Stanley, J. C. (2015). *Experimental and quasi-experimental designs for research*. Ravenio Books.
- Cheng, K.-H., & Tsai, C.-C. (2013). Affordances of augmented reality in science learning: Suggestions for future research. *Journal of Science Education and Technology*, 22(4), 449–462. <https://doi.org/10.1007/s10956-012-9405-9>.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- Fauzi, Yuniar Risty (2022) *Interactive Learning Media of the Food Digestive System in Humans*. Diploma thesis, Ganesha University of Education. https://repo.undiksha.ac.id/13507/?utm_source=chatgpt.com.
- Gagnier, K., Holochwest, S., & Fisher, K. (2021). Spatial thinking in science, technology, engineering, and mathematics: Elementary teachers' beliefs, perceptions, and self-efficacy. *Journal of Research in Science Teaching*. <https://doi.org/10.1002/tea.21722>.
- Garzón, J., & Acevedo, J. (2019). Meta-analysis of the impact of augmented reality on students' learning gains. *Educational Research Review*, 27, 244–260. <https://doi.org/10.1016/j.edurev.2019.04.001>.
- Kalyani, L. (2024). The Role of Technology in Education: Enhancing Learning Outcomes and 21st Century Skills. *International Journal of Scientific Research in Modern Science and Technology*. <https://doi.org/10.59828/ijrmst.v3i4.199>.
- Khairah, M., Ahmad, R., & Suryani, N. (2025). The multiple intelligence of students in class X-B SMAN 6 Bone is based on Gardner's theory. *Pendas: Scientific Journal of Elementary Education*, 10(3), 145–156. <https://doi.org/10.23969/jp.v10i03.30600>.
- Maidah, S., Nusantara, T., Arifin, S., Anggraini, A., & Faizah, S. (2025). Pop-Up Book as a Science Learning Media to Improve Visualization Abilities and Understanding of Science Concepts in SDN Pandanrejo II. *Cetta: Journal of Educational Sciences*. <https://doi.org/10.37329/cetta.v8i4.4277>.

- Mustaqim, I., & Kurniawan, N. (2017). Development of augmented reality-based learning media. *Journal of Electrical Education*, 1(1), 36–48. <https://doi.org/10.21831/jee.v1i1.13267>.
- Pilehvar, A. (2021). Spatial-geographical analysis of urbanization in Iran. *Humanities and Social Sciences Communications*, 8, 1-12. <https://doi.org/10.1057/s41599-021-00741-w>.
- Rosa, N., Masita, M., & Ramadhan, S. (2026). Deep Learning Approach in IPAS Instruction: Enhancing Elementary School Students' Skills and Competencies. *Journal of Islamic Education Students (JIES)*, 6(1), 320–335.
- Schmidt, R., & Stumpe, B. (2025). Systematic review of mobile augmented reality applications in geography education. *Review of Education*. <https://doi.org/10.1002/rev3.70042>.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin.
- Syahputra, A., Harahap, R. D., & Safitri, I. (2022). An Analysis of Student Learning Challenges in Elementary School Science Subject. *Jurnal Kependidikan: Jurnal Hasil Penelitian Dan Kajian Kepustakaan Di Bidang Pendidikan, Pengajaran, Dan Pembelajaran*, 8(1), 237–247.
- Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, 139(2), 352–402. <https://doi.org/10.1037/a0028446>.
- Wahyuningsih, S., & Budiningsih, C. A. (2022). The effectiveness of augmented reality-based interactive learning media on natural appearance materials to improve the understanding of elementary school students' concepts. *Indonesian Journal of Basic Education*, 7(2), 45–55. <https://doi.org/10.26737/jpdi.v7i2.3001>
- Zakirman, Z., Gusta, W., & Rahayu, C. (2022). Analysis of problems in science learning at the elementary school. *Jurnal Ilmiah Profesi Pendidikan*, 7(1), 24–29.
- Zou, Y., Kuek, F., Feng, W., & Cheng, X. (2025). Digital learning in the 21st century: trends, challenges, and innovations in technology integration. *Frontiers in Education*. <https://doi.org/10.3389/educ.2025.1562391>.