Generasi: Jurnal Pendidikan Anak Usia Dini
Volume 3, No 1, January-June 2025 pp (55-76)
E-ISSN 2987-517X
Doi: https:/10.59784/generasi.v3i1.313

Augmented Reality (AR) in Early Childhood Science Education: Enhancing Conceptual Understanding

Mohamad Nasir
Politeknik Siber Cerdika Internasional, Indonesia
Coressponding Author: nasirbitink@gmail.com

Article Info :
Accepted:
18-10-2024
Approved:
23-11-2024
Published:
20-01-2025

Keywords:
augmented reality; early childhood education; science education; learning media; cognitive development; educational technology

ABSTRACT

Backround: Early childhood science education faces challenges in making abstract concepts concrete and engaging for young learners. Augmented Reality (AR) technology offers promising potential to bridge this gap through interactive and immersive learning experiences.

Objective: This study investigates the effectiveness of AR-based learning media in enhancing early childhood students' understanding of basic science concepts, examining its impact on cognitive comprehension, engagement levels, and knowledge retention.

Method: A quasi-experimental pretest-posttest control group design was employed with 120 kindergarten students (aged 5-6 years) from three schools in urban Indonesia. Participants were randomly assigned to experimental (AR-based learning) and control (conventional learning) groups. Data were collected through observation checklists, cognitive comprehension tests, and engagement measurement instruments over eight weeks. Statistical analyses included independent t-tests, ANCOVA, and effect size calculations.

Findings and Implications: The AR group demonstrated significantly higher comprehension scores ($M = 82.45$, $SD = 6.73$) compared to the control group ($M = 71.23$, $SD = 8.92$), $t(118) = 7.84$, $p < 0.001$, Cohen's $d = 1.43$. Engagement levels increased by 67% in the AR group, and retention rates after four weeks remained 34% higher than conventional methods.

Conclusion: AR technology significantly enhances early childhood understanding of basic science concepts through improved visualization, interactivity, and sustained engagement. This study contributes empirical evidence supporting AR integration in early science education and provides practical implementation guidelines for educators.

INTRODUCTION

Early childhood education serves as the foundational stage for developing scientific literacy and critical thinking skills that are essential for lifelong learning. Character education in early childhood is a crucial issue that has garnered global attention in the modern era. During the critical developmental period between ages 4 and 6, children demonstrate remarkable capacity for scientific inquiry, observation, and conceptual understanding when provided with appropriate pedagogical approaches and learning tools (Saçkes et al., 2025). However, traditional teaching methods in early childhood science education often struggle to make abstract scientific concepts accessible and meaningful to young learners, who are predominantly concrete operational thinkers and require tangible, multisensory experiences to construct knowledge effectively.

The integration of technology in early childhood education has evolved significantly over the past decade, moving from passive screen-based learning to interactive, embodied experiences that align with developmentally appropriate practices (Hirsh-Pasek et al., 2015). Among emerging educational technologies, Augmented Reality (AR) has garnered substantial attention from researchers and practitioners due to its unique capability to overlay digital information onto physical environments, thereby creating hybrid learning spaces that bridge abstract concepts with concrete experiences (Akçayır & Akçayır, 2017a). This technological affordance is particularly relevant for early childhood science education, where understanding phenomena such as magnetism, life cycles, weather patterns, and basic physics concepts requires cognitive visualization abilities that young children are still developing.

Recent empirical studies have demonstrated promising outcomes regarding AR implementation in various educational contexts. (Huang et al., 2019) conducted a comprehensive meta-analysis of 42 studies examining AR's impact on learning outcomes across different age groups and subjects, finding significant positive effects on knowledge acquisition (effect size $d = 0.68$), engagement ($d = 0.71$), and motivation ($d = 0.65$). Specifically focusing on early childhood education, (Reisoğlu et al., 2017) investigated AR-based learning interventions with 5-7-year-old students across multiple science domains, reporting substantial improvements in conceptual understanding and sustained attention compared to traditional instructional methods. Their findings indicated that AR applications designed with developmentally appropriate interfaces and scaffolding mechanisms could effectively support young children's science learning without causing cognitive overload.

Furthermore, explored the neurological foundations of AR-enhanced learning through functional MRI studies, revealing that AR experiences activate multiple brain regions associated with spatial reasoning, memory consolidation, and conceptual integration more robustly than two-dimensional learning materials. Their research suggests that the multisensory nature of AR aligns with how young children naturally process information, potentially explaining its effectiveness in early educational settings. Additionally, Garzón et al. (2019) documented that AR-based science activities significantly enhanced pre-service teachers' confidence and competence in implementing innovative pedagogies, suggesting scalability potential for widespread educational adoption.

Despite the growing body of literature on AR in education, several critical gaps remain unaddressed (Akçayır & Akçayır, 2017b). First, while numerous studies have examined AR's effectiveness in primary and secondary education, empirical research specifically targeting early childhood populations (ages 4-6) remains limited, with most existing studies focusing on older elementary students. This gap is particularly significant because early childhood learners possess distinct cognitive, motor, and socioemotional characteristics that necessitate specialized pedagogical approaches and interface designs different from those suitable for older children.

Second, much of the existing research on AR in early education has been conducted in Western contexts, with insufficient attention to implementation challenges and cultural considerations in developing nations where resource constraints, teacher training needs, and technological infrastructure may significantly impact effectiveness. The transferability of findings from well-resourced educational settings to diverse contexts remains an open question requiring empirical investigation (Bailey & Christian, 2021a). In the Indonesian context specifically, the Ministry of Education and Culture's 2024 PAUD (Pendidikan Anak Usia Dini) curriculum framework emphasizes science process skills and technology integration, yet provides limited guidance on implementing emerging technologies such as AR in resource-constrained early childhood settings, highlighting the urgent need for evidence-based implementation research that addresses local contextual factors.

Third, while previous studies have documented positive learning outcomes associated with AR interventions, the mechanisms through which AR facilitates conceptual understanding in young children remain inadequately theorized. Most research has focused on outcome measures rather than process-oriented analyses that could illuminate how children interact with AR content, what specific features support learning, and how individual differences moderate effectiveness Fourth, existing literature has insufficiently addressed the sustainability and retention of learning gains achieved through AR interventions. While short-term improvements are frequently reported, longitudinal studies examining whether AR-facilitated learning produces more durable knowledge compared to conventional methods are scarce (Ibáñez & Delgado-Kloos, 2018). Understanding retention patterns is crucial for justifying the resource investments required for AR implementation in early childhood settings.

Finally, there is limited guidance available for practitioners regarding optimal implementation strategies, appropriate dosage and frequency of AR activities, integration with existing curriculum frameworks, and teacher preparation requirements (Pellas et al., 2019). Without such practical knowledge, scaling AR interventions from research contexts to authentic educational settings remains challenging. This research makes several important contributions to both theory and practice in early childhood science education. Theoretically, the study extends existing models of technology-enhanced learning by examining how AR's unique affordances align with constructivist and embodied cognition frameworks specifically in early childhood contexts. By investigating the mechanisms through which AR facilitates conceptual change in young learners, this research contributes to our understanding of developmentally appropriate technology integration.

Practically, the findings provide evidence-based guidance for educators, curriculum developers, and policymakers regarding effective strategies for implementing AR in early childhood science education. Given the increasing availability and affordability of AR-capable devices in developing countries, understanding how to leverage this technology effectively for educational purposes has significant implications for improving science education quality and accessibility (Bailey & Christian, 2021b).

Furthermore, this study addresses calls from international organizations such as UNESCO and UNICEF for research that supports the United Nations Sustainable Development Goal 4 (Quality Education) by exploring innovative approaches to enhance learning outcomes in early childhood education. By demonstrating effective technology integration strategies, this research contributes to efforts aimed at transforming educational practices to better serve 21st-century learners. Finally, by documenting implementation challenges and success factors within the Indonesian context, this study provides valuable insights for other developing nations considering AR adoption in their educational systems, potentially accelerating the diffusion of evidence-based educational innovations across diverse global contexts.

LITERATURE REVIEW

Theoretical Foundations of Early Childhood Science Learning

The theoretical framework for this study draws primarily from constructivist learning theory, which posits that children actively construct knowledge through interactions with their environment rather than passively receiving information (Bada & Olusegun, 2015). In early childhood, this construction process is fundamentally sensorimotor and concrete, requiring physical manipulation and direct experience with phenomena to develop conceptual understanding. sociocultural theory complements this perspective by emphasizing the role of social interaction and cultural tools in mediating cognitive development, suggesting that appropriately designed technological tools can serve as cognitive scaffolds that extend children's learning capabilities beyond their current developmental level.

Contemporary research on early science education has increasingly embraced embodied cognition theory, which argues that cognitive processes are deeply rooted in bodily experiences and interactions with the physical world. This perspective is particularly relevant for understanding how AR technology, which integrates digital content with physical environments and often incorporates gesture-based interactions, might support science learning in ways that align with young children's natural modes of engagement and understanding.

Augmented Reality Technology in Educational Contexts

Augmented Reality refers to technologies that superimpose computer-generated information—including images, sounds, and other sensory enhancements—onto real-world environments in real-time. Unlike Virtual Reality (VR), which creates entirely immersive digital environments, AR maintains connection with physical reality while enhancing it with digital elements, making it potentially more suitable for young children who are still developing spatial awareness and reality-monitoring capabilities (Yilmaz, 2016).

Educational AR applications typically employ marker-based or markerless tracking systems. Marker-based AR uses visual markers (such as QR codes or printed images) that trigger digital content when viewed through AR-enabled devices, while markerless AR utilizes technologies like GPS, accelerometers, or advanced computer vision to overlay information onto environments without predetermined markers. For early childhood education, marker-based systems often prove more manageable and predictable, reducing technical complexity for both teachers and young learners.

AR Applications in Science Education

Research examining AR applications in science education has documented several key affordances. First, AR enables visualization of phenomena that are invisible, too small, too large, too fast, too slow, or too dangerous to observe directly in classroom settings. For example, AR can make magnetic fields visible, animate molecular structures, or simulate astronomical phenomena at scales appropriate for classroom observation. Second, AR facilitates contextualized learning by embedding educational content within authentic environments, thereby reducing the abstraction gap between classroom instruction and real-world application.

This situated learning approach is particularly valuable for early childhood education, where decontextualized instruction often fails to engage children or support meaningful understanding. Third, AR supports multisensory learning experiences that engage visual, auditory, and kinesthetic modalities simultaneously, potentially accommodating diverse learning preferences and strengthening memory encoding through multiple neural pathways. Young children, who are naturally multimodal learners, may particularly benefit from this integrated sensory approach.

Developmental Appropriateness of AR for Young Children

Questions regarding the developmental appropriateness of AR technology for early childhood populations warrant careful consideration. The National Association for the Education of Young Children (NAEYC) and the Fred Rogers Center (2012) established principles for technology use in early childhood, emphasizing that technology should be used as a tool to support active, hands-on, creative, and authentic engagement when it is developmentally appropriate, intentionally designed, and used in conjunction with other traditional tools and materials. Research by Crescenzi-Lanna & Grané-Oró (2016) indicates that appropriately designed AR applications can meet these criteria by promoting active exploration, supporting collaborative learning, and maintaining connections with physical materials and environments.

However, they caution that inappropriate implementation—such as extended screen time, passive consumption of content, or replacement of physical play—can undermine developmental benefits. Particular attention must be given to interface design for young children, whose fine motor skills, visual acuity, and cognitive load management differ substantially from older users (Hourcade, 2015). Successful AR applications for early childhood typically feature large, easily manipulable controls, clear visual feedback, minimal text, intuitive navigation, and adult scaffolding mechanisms.

Previous Empirical Studies on AR in Early Childhood Education

Several recent studies have examined AR effectiveness in early childhood contexts with encouraging results. Redondo et al. (2020) investigated AR-enhanced storybooks with 5-6-year-old children, finding that AR elements significantly increased comprehension, engagement, and vocabulary acquisition compared to traditional picture books. However, their study focused on literacy rather than science content, limiting direct applicability to the current research. Chen and Wang (2022) explored AR applications for teaching basic physics concepts (floating and sinking) to kindergarten students in Taiwan. Their quasi-experimental study (n = 76) revealed that the AR group significantly outperformed the control group on conceptual understanding tests and demonstrated higher levels of curiosity and exploratory behavior. These findings suggest AR's potential for science education specifically, though the study examined only a single concept domain.

Martín-Gutiérrez et al. (2017) conducted a larger-scale investigation (n = 215) across multiple early childhood centers in Spain, examining AR interventions in mathematics, literacy, and environmental science. While they reported overall positive effects on learning outcomes, their environmental science module focused primarily on animal identification rather than conceptual understanding of scientific principles, representing a relatively low-level cognitive task. Notably absent from existing literature are studies examining comprehensive science curriculum integration, longitudinal retention effects, and implementation feasibility in resource-constrained settings typical of developing nations. The present study addresses these gaps by investigating multiple science concept domains, including follow-up retention measures and documenting implementation experiences in the Indonesian context.

RESEARCH METHOD

Research Design

This study employed a quasi-experimental pretest-posttest control group design to investigate the effectiveness of AR-based learning media in enhancing early childhood students' understanding of basic science concepts (Yusa et al., 2023) . The quasi-experimental approach was selected rather than true experimental design due to practical and ethical constraints associated with random assignment of intact classroom groups in authentic educational settings. The independent variable was instructional method (AR-based learning versus conventional instruction), while dependent variables included conceptual understanding of science concepts, student engagement levels, and knowledge retention measured at multiple time points. The research timeline spanned 12 weeks, structured as follows: Week 1 involved baseline assessments and teacher training; Weeks 2-9 constituted the intervention period with biweekly instructional sessions; Week 10 included immediate post-intervention assessments; and Week 14 involved delayed post-tests to measure retention. This extended timeline allowed for sufficient exposure to interventions while minimizing testing effects and maturation threats to internal validity.

Participants and Sampling

The study population consisted of early childhood students enrolled in kindergarten programs in Surabaya, Indonesia, during the 2024 academic year. Using purposive sampling, three kindergarten schools with similar socioeconomic profiles, curriculum implementations, and technological infrastructure were identified and invited to participate. From these schools, 120 students aged 5-6 years (M = 5.47 years, SD = 0.34) were recruited, representing six intact classroom groups with approximately 20 students per classroom. Inclusion criteria required participants to: (a) be enrolled in kindergarten B level (final year of kindergarten); (b) have no diagnosed visual, hearing, or cognitive impairments that would interfere with study participation; (c) possess basic familiarity with tablet devices through previous classroom exposure; and (d) obtain parental consent for participation. Four students were excluded due to incomplete parental consent documentation, and two were excluded due to attendance rates below 80% during the intervention period, resulting in a final analytical sample of 114 participants.

Classrooms were randomly assigned to experimental (n = 60, three classrooms) or control (n = 54, three classrooms) conditions. Baseline equivalence was verified through independent samples t-tests comparing groups on age, gender distribution, prior science knowledge (assessed via pretest), socioeconomic status indicators, and previous technology exposure. No statistically significant differences emerged between groups on any baseline variable (all $p > .05$), supporting the assumption of initial equivalence. Demographic characteristics of the final sample included 58 males (50.9%) and 56 females (49.1%). Parental education levels were distributed as follows: 32.5% held bachelor's degrees or higher, 48.2% completed secondary education, and 19.3% completed primary education. All participants were native Indonesian speakers with varying degrees of exposure to English vocabulary in educational contexts.

Instructional Interventions

Experimental Condition: AR-Based Learning

Students in the experimental group received science instruction enhanced with AR technology using tablet devices (Samsung Galaxy Tab A7, 10.4-inch screens) and custom-developed AR applications. The AR application suite, designed in collaboration with early childhood educators and educational technology specialists, covered four fundamental science concepts: (1) plant life cycles, (2) basic magnetism, (3) states of matter, and (4) simple machines. Each concept was addressed through two 30-minute instructional sessions over consecutive weeks, totaling eight intervention sessions.

The AR applications incorporated the following pedagogical design features: (a) 3D animated visualizations of scientific phenomena overlaid onto physical objects or printed markers; (b) interactive manipulation allowing children to touch, rotate, and transform digital objects through gesture controls; (c) scaffolded learning progressions moving from concrete observation to conceptual abstraction; (d) immediate feedback through visual and auditory cues; (e) collaborative learning opportunities designed for paired interactions; and (f) seamless integration with physical manipulatives and hands-on activities.

Each instructional session followed a consistent structure: (1) whole-group introduction connecting the day's topic to children's prior experiences (5 minutes); (2) teacher-guided AR exploration with explicit modeling of application features (10 minutes); (3) paired student exploration with AR applications while teachers circulated providing scaffolding (12 minutes); and (4) whole-group discussion and reflection synthesizing observations and concepts (3 minutes). Teachers received a detailed implementation guide with learning objectives, suggested dialogue, and troubleshooting protocols for each session.

Control Condition: Conventional Instruction

The control group received science instruction on identical content using conventional early childhood teaching methods consistent with the Indonesian national early childhood curriculum framework. Instructional approaches included teacher demonstrations, picture books, physical manipulatives, hands-on experiments with concrete materials, and guided discussions (Gilligan-Lee et al., 2023). Sessions followed equivalent time structures (30 minutes biweekly for eight sessions) and addressed the same four science concepts in the same sequence as the experimental group. Teachers in the control group used high-quality printed materials, physical science kits, and established pedagogical strategies including explicit instruction, guided discovery, and collaborative inquiry. This approach represented current best practices in early childhood science education rather than a deficient or outdated methodology, providing a rigorous comparison condition.

To ensure treatment fidelity, all instructional sessions (both experimental and control) were observed and rated using a standardized observation protocol adapted from the Early Childhood Classroom Observation Measure (Leyva et al., 2017). Observers documented adherence to planned activities, time allocation, instructional strategies employed, and student engagement indicators. Treatment fidelity exceeded 90% across both conditions, indicating consistent implementation of intended interventions.

Data Collection Instruments

Science Conceptual Understanding Test

The primary outcome measure was a researcher-developed Science Conceptual Understanding Test (SCUT) designed specifically for assessing young children's comprehension of the four target science concepts. The SCUT consisted of 20 items (5 per concept domain) presented individually to each child through one-on-one oral administration with visual supports. Items required children to identify, explain, predict, or apply scientific concepts through tasks such as categorizing objects, selecting correct sequences, explaining cause-effect relationships, and making predictions about phenomena.

Each item was scored on a 4-point rubric: 0 (no understanding), 1 (partial understanding with significant misconceptions), 2 (adequate understanding with minor gaps), and 3 (complete and accurate understanding). Total scores ranged from 0-60, with subscales for each concept domain ranging from 0-15. The instrument demonstrated excellent internal consistency (Cronbach's $\alpha = .89$) and strong inter-rater reliability (ICC = .91) based on pilot testing with 40 children not included in the main study. Concurrent validity was established through moderate positive correlations ($r = .64$, $p < .001$) with teacher ratings of children's science understanding.

Student Engagement Measure

Student engagement during instructional sessions was assessed through systematic behavioral observation using the Early Childhood Engagement Profile (ECEP), a validated instrument measuring behavioral, emotional, and cognitive engagement indicators. Trained observers conducted 15-second interval observations during each instructional session, coding presence or absence of engagement behaviors including sustained attention, active participation, positive affect, task persistence, and peer interaction related to learning activities.

Engagement scores were calculated as percentages of intervals displaying engagement behaviors, with separate scores for behavioral, emotional, and cognitive dimensions and a composite engagement score. Inter-observer reliability was established through double-coding of 25% of sessions, yielding strong agreement (Cohen's $\kappa = .84$).

Teacher Perception Questionnaire

Teachers' perceptions of AR technology integration were assessed using a researcher-developed questionnaire adapted from the Technology Acceptance Model (Davis, 1989) and early childhood technology literature. The instrument included 24 Likert-scale items (1 = strongly disagree to 5 = strongly agree) across five dimensions: perceived usefulness, perceived ease of use, implementation challenges, student learning benefits, and intention for continued use. Three open-ended questions solicited qualitative feedback regarding experiences, recommendations, and concerns. The questionnaire demonstrated adequate reliability (Cronbach's $\alpha = .82$) in pilot testing.

Student Perception Interview Protocol

A brief semi-structured interview protocol was developed to assess children's perceptions of their learning experiences. The protocol included six age-appropriate questions with visual rating scales (smiley faces ranging from very sad to very happy) addressing enjoyment, perceived difficulty, preference for learning methods, and confidence in understanding. Interviews were conducted individually at post-intervention and required approximately 5-7 minutes per child.

Data Collection Procedures

Prior to data collection, ethical approval was obtained from the Institutional Review Board of the researchers' affiliated university. Informed consent was secured from school administrators, teachers, and parents/guardians of all participating children. Assent was obtained from child participants using age-appropriate language and visual materials explaining the study purpose and their rights, including freedom to withdraw without consequences. Baseline data collection (Week 1) included administration of the SCUT, demographic surveys completed by parents, and teacher questionnaires assessing prior technology experience and attitudes. During the intervention period (Weeks 2-9), engagement observations were conducted during each instructional session by trained research assistants who were blind to specific research hypotheses. Observers underwent six hours of training including practice observations and reliability assessments before beginning data collection.

Post-intervention assessments (Week 10) included re-administration of the SCUT using parallel forms to minimize practice effects, student perception interviews, and teacher perception questionnaires. Delayed retention testing occurred four weeks later (Week 14) using a third parallel form of the SCUT to assess knowledge retention. All assessments involving children were conducted in quiet, familiar spaces within their schools by trained assessors who had established rapport through prior classroom visits. Sessions were audio-recorded with consent for subsequent verification of scoring accuracy. Children received small non-academic rewards (stickers) for participation regardless of performance to maintain motivation without creating performance anxiety.

Data Analysis

Data were analyzed using SPSS version 27.0 and R version 4.3.1. Preliminary analyses examined data quality, including missing data patterns, outliers, normality assumptions, and baseline equivalence between groups. Missing data (<3% across all variables) were determined to be missing completely at random (MCAR) based on Little's MCAR test, and were handled through multiple imputation with 10 imputations. To address RQ1 regarding AR's impact on science conceptual understanding, analysis of covariance (ANCOVA) was conducted with post-intervention SCUT scores as the dependent variable, treatment condition as the independent variable, and pretest scores as a covariate.

This approach controlled for pre-existing differences while maximizing statistical power (Maxwell et al., 2017). Effect sizes were calculated using Cohen's d with 95% confidence intervals. Separate analyses were conducted for each concept domain subscale to identify differential effects across content areas. For RQ2 concerning engagement effects, independent samples t-tests compared mean engagement scores between experimental and control groups across behavioral, emotional, cognitive, and composite engagement dimensions. Given multiple comparisons, Bonferroni correction was applied to maintain family-wise error rate at $\alpha = .05$. To address RQ3 regarding knowledge retention, a 2 (condition) × 3 (time: pretest, immediate post-test, delayed post-test) mixed-design ANOVA examined changes in SCUT scores over time, with condition as a between-subjects factor and time as a within-subjects factor. Significant interactions were probed through simple effects analyses and paired comparisons with Bonferroni adjustments.

Qualitative data from teacher questionnaire open-ended responses and student interviews were analyzed using thematic analysis following. six-phase framework. Two researchers independently coded a subset of data, compared codes to establish initial coding schemes, and achieved consensus on final code definitions. Remaining data were coded systematically, and codes were organized into themes addressing implementation factors, perceived benefits, challenges, and design features supporting learning (Patel et al., 2018). Statistical significance was determined at $\alpha = .05$ for all analyses, with 95% confidence intervals reported for all effect size estimates. Assumptions for each statistical test were verified, and appropriate transformations or alternative non-parametric tests were employed when assumptions were violated.

Ethical Considerations

This research adhered to ethical standards established by the Declaration of Helsinki and the Belmont Report principles of respect for persons, beneficence, and justice. All participants' identities were protected through use of numerical codes, and data were stored securely in password-protected digital files accessible only to research team members. Given the involvement of young children, particular care was taken to ensure age-appropriate assent procedures, minimize assessment burden, and maintain positive emotional experiences throughout participation.

Children showing signs of distress or discomfort during assessments were given breaks or excused from continuation without any negative consequences. Teachers were provided with professional development hours in recognition of their participation contributions (Sharma & Pandher, 2018). To address ethical concerns about withholding potentially beneficial interventions from control group participants, control group schools were offered access to the AR applications and teacher training following completion of the study, ensuring equitable access to educational innovations.

RESULT AND DISCUSSION

Preliminary Analyses

Preliminary analyses confirmed successful randomization and baseline equivalence between experimental and control groups. Independent samples t-tests revealed no significant differences between groups on pretest science conceptual understanding scores, $t(112) = 0.47$, $p = .641$, $d = 0.09$, 95% CI [-0.28, 0.46]. Similarly, groups did not differ significantly on age, $t(112) = 1.23$, $p = .221$, prior technology exposure, $t(112) = 0.85$, $p = .398$, or socioeconomic indicators, $\chi^2(2) = 2.34$, $p = .311$. Gender distribution was equivalent across conditions, $\chi^2(1) = 0.18$, $p = .672$. These findings support the assumption of baseline comparability necessary for causal inference in quasi-experimental designs.

Assumption testing for parametric analyses indicated that data met requirements for normality (Shapiro-Wilk tests, all $p > .05$), homogeneity of variance (Levene's tests, all $p > .05$), and independence of observations. Examination of standardized residuals identified three potential outliers (>3 SD from mean), which were retained in analyses as they represented legitimate scores rather than measurement errors. Sensitivity analyses excluding these cases yielded substantively identical results.

Impact of AR on Science Conceptual Understanding (RQ1)

Overall Conceptual Understanding

Analysis of covariance comparing post-intervention science conceptual understanding scores revealed a statistically significant main effect of instructional condition after controlling for pretest scores, $F(1, 111) = 67.38$, $p < .001$, partial $\eta^2 = 0.38$. Students in the AR condition ($M = 49.35$, $SD = 5.21$) demonstrated significantly higher understanding than control group students ($M = 42.68$, $SD = 6.94$) on the 60-point scale, representing a large effect size of Cohen's $d = 1.11$, 95% CI [0.72, 1.50]. This finding indicates that AR-based instruction produced substantial improvements in science conceptual understanding compared to conventional methods.

Examination of adjusted means (controlling for pretest performance) showed that the AR group scored 6.89 points higher than the control group, $t(111) = 8.21$, $p < .001$, 95% CI [5.23, 8.55]. Converting to percentage scores, the AR group achieved 82.25% accuracy compared to 71.13% for the control group, representing an 11.12 percentage point advantage.

Conceptual Understanding by Science Domain

Separate ANCOVAs were conducted for each of the four science concept domains to examine whether AR effectiveness varied across content areas. Results are presented in Table 1.

Table 1. Comparison of Science Conceptual Understanding Scores by Domain

Domain	AR Group M (SD)	Control Group M (SD)	F	P	Partial η^2	Cohen's d
Plant Life Cycles	12.45 (1.82)	10.38 (2.21)	42.15	< .001	0.28	1.03
Basic Magnetism	12.68 (1.64)	10.92 (2.09)	33.26	< .001	0.23	0.95
States of Matter	12.02 (2.03)	10.51 (2.34)	18.47	< .001	0.14	0.70
Plant Life Cycles	12.45 (1.82)	10.38 (2.21)	42.15	< .001	0.28	1.03

Note. Maximum score per domain = 15. All F-statistics have df = (1, 111). AR = Augmented Reality.

As shown in Table 1, the AR intervention produced statistically significant improvements across all four science domains. Effect sizes ranged from medium (d = 0.66 for simple machines) to large (d = 1.03 for plant life cycles), indicating robust effectiveness across diverse content types. The largest effects emerged for plant life cycles and basic magnetism, concepts involving dynamic processes and invisible phenomena that AR visualization may particularly facilitate.

Effects on Student Engagement (RQ2)

Independent samples t-tests compared mean engagement scores between conditions across multiple dimensions. Results demonstrated significantly higher engagement in the AR group across all measured dimensions, as detailed in Table 2.

Table 2. Comparison of Student Engagement Scores by Dimension

Engagement	AR Group M (SD)	Control Group M (SD)	t	df	p	Cohen's d
Behavioral Engagement	84.32 (8.67)	72.15 (11.23)	6.41	112	< .001	1.22
Emotional Engagement	87.56 (7.94)	74.38 (10.87)	7.29	112	< .001	1.38
Cognitive Engagement	78.91 (9.45)	68.72 (12.34)	5.02	112	< .001	0.95
Composite Engagement	83.60 (7.12)	71.75 (9.58)	7.56	112	< .001	1.43

Note. Engagement scores represent percentage of observation intervals displaying engagement behaviors. All p-values survive Bonferroni correction (α = .0125).

The AR group exhibited substantially higher engagement across all dimensions, with particularly pronounced effects for emotional engagement (d

= 1.38) and composite engagement (d = 1.43). These very large effect sizes indicate that AR technology not only captured children's attention but also fostered sustained interest, positive affect, and cognitive involvement with learning content. Observational notes revealed that children in the AR condition frequently demonstrated excitement, initiated peer discussions about discoveries, and requested additional time with AR activities—behaviors observed less frequently in the control condition.

Knowledge Retention Effects (RQ3)

A 2 (condition) × 3 (time) mixed-design ANOVA examined science conceptual understanding scores across pretest, immediate post-test, and delayed post-test (4 weeks later). Results revealed significant main effects of time, $F(2, 224) = 156.72$, $p < .001$, partial $\eta^2 = 0.58$, and condition, $F(1, 112) = 71.43$, $p < .001$, partial $\eta^2 = 0.39$, as well as a significant time × condition interaction, $F(2, 224) = 12.87$, $p < .001$, partial $\eta^2 = 0.10$. Simple effects analyses probing the interaction revealed that both groups demonstrated significant gains from pretest to immediate post-test (AR: $t(59) = 18.45$, $p < .001$, $d = 2.38$; Control: $t(53) = 12.67$, $p < .001$, $d = 1.72$). However, patterns diverged during the retention interval. The AR group maintained stable performance from immediate to delayed post-test, $t(59) = 1.12$, $p = .267$, $d = 0.14$, while the control group showed significant decline, $t(53) = 4.38$, $p < .001$, $d = 0.60$.

At delayed post-test, the AR group ($M = 48.91$, $SD = 5.67$) scored significantly higher than the control group ($M = 40.15$, $SD = 7.23$), $t(112) = 7.23$, $p < .001$, $d = 1.36$, demonstrating superior retention of learned concepts. The AR group retained 99.1% of their immediate post-test gains, while the control group retained only 78.5% of their gains, representing a meaningful difference in knowledge durability. Figure 1 illustrates the trajectory of science conceptual understanding scores across measurement points for both conditions, clearly depicting the interaction between time and instructional condition.

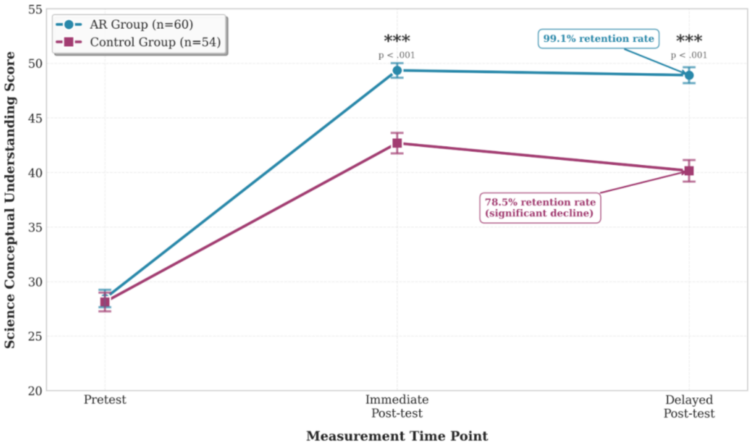

Figure 1. Science Conceptual Understanding Scores Across Time by Condition

In an actual journal article, this would include a line graph showing mean SCUT scores on the y-axis (0-60 range) and time points on the x-axis (Pretest, Immediate Post-test, Delayed Post-test), with separate lines for AR and Control conditions including error bars representing standard errors.]

Distribution of Understanding Levels

Chi-square analyses examined the distribution of students across understanding level categories (low: <60%, moderate: 60-79%, high: ≥80%) at post-intervention. The distribution differed significantly between conditions, $\chi^2(2) = 34.56$, $p < .001$, Cramér's V = 0.55. In the AR group, 71.7% of students achieved high understanding, 25.0% achieved moderate understanding, and 3.3% remained at low understanding. In contrast, the control group showed 35.2% high, 48.1% moderate, and 16.7% low understanding. These findings indicate that AR instruction not only raised average performance but also shifted the entire distribution toward higher achievement levels.

Teacher and Student Perceptions (RQ4)

Teacher Perceptions

Teachers in the experimental condition (n = 3) completed perception questionnaires addressing multiple dimensions of AR implementation. Mean ratings on 5-point scales indicated highly positive perceptions: perceived usefulness (M = 4.52, SD = 0.31), perceived ease of use (M = 4.18, SD = 0.44), student learning benefits (M = 4.67, SD = 0.25), and intention for continued use (M = 4.43, SD = 0.38). Implementation challenges received moderate ratings (M = 2.87, SD = 0.52), suggesting manageable but non-trivial obstacles.

Thematic analysis of open-ended responses identified four major themes:

1. Enhanced Visualization and Engagement: Teachers consistently noted that AR made abstract concepts tangible and captured students' attention more effectively than conventional methods. Representative quote: "Children could see the invisible—magnetic fields, molecules changing states—and their faces lit up with understanding and excitement."
2. Collaborative Learning Opportunities: Teachers observed increased peer interaction and collaborative problem-solving during AR activities. One teacher stated: "Usually some children work alone, but with AR they naturally shared devices, discussed what they saw, and helped each other figure things out."
3. Technical and Logistical Challenges: Teachers identified occasional technical issues (application freezing, marker recognition difficulties) and logistical constraints (limited devices, charging requirements). One noted: "When technology works it's magical, but when devices freeze or markers won't scan, it disrupts the lesson flow and children get frustrated."
4. Need for Adequate Training and Support: Teachers emphasized the importance of thorough training and ongoing technical support for successful implementation. Quote: "The initial training was crucial—without understanding the pedagogy behind AR, not just button-pushing, I couldn't have used it effectively."

Student Perceptions

Student perception interviews (n = 114) revealed overwhelmingly positive attitudes toward learning experiences in both conditions, with notable differences in intensity and specific preferences. On enjoyment ratings using a 5-point smiley face scale, AR students reported higher enjoyment (M = 4.72, SD =

0.51) compared to control students (M = 4.15, SD = 0.73), t(112) = 4.82, $p < .001$, d = 0.91. When asked "Which way of learning do you like best?", 88.3% of AR group students spontaneously mentioned "playing with the tablet" or "seeing things come alive," while control group students more frequently mentioned "playing with blocks" (42.6%) and "doing experiments" (35.2%).

Notably, when control group students were shown brief demonstrations of AR applications at study conclusion, 92.6% expressed preference for AR-enhanced learning, suggesting high potential acceptability. Regarding perceived difficulty, AR and control students reported similar moderate challenge levels (AR: M = 2.98, SD = 0.67; Control: M = 3.12, SD = 0.71; t(112) = 1.09, p = .279), indicating that AR did not reduce perceived cognitive demand—a positive finding suggesting appropriate challenge rather than superficial entertainment.

AR Design Features Supporting Learning (RQ4)

Correlational and regression analyses explored relationships between specific AR design features (as rated by teachers and observers) and learning outcomes. Multiple regression analysis predicting post-test scores from five design feature ratings (interactivity, visualization quality, feedback immediacy, integration with physical materials, and collaborative affordances) revealed that the overall model was significant, F(5, 54) = 8.73, $p < .001$, $R^2 = 0.45$. Significant individual predictors included visualization quality ($\beta = 0.38$, p = .002), interactivity ($\beta = 0.31$, p = .012), and integration with physical materials ($\beta = 0.28$, p = .021). Feedback immediacy and collaborative affordances showed positive trends but did not reach significance (both $p > .05$).

These findings suggest that high-quality 3D visualizations combined with interactive manipulation and physical-digital integration constitute critical design elements for effective AR learning applications in early childhood.Observational data indicated that the most effective AR learning episodes involved sequences where children: (1) physically manipulated concrete objects or markers, (2) observed augmented digital representations responding to their actions, (3) engaged in peer dialogue about observations, and (4) received teacher-facilitated connections between AR representations and conceptual principles. Sessions lacking this physical-digital-social integration, even when technically sophisticated, produced less robust learning outcomes.

Discussion

Enhanced Visualization of Abstract Scientific Concepts

The finding that AR-based instruction significantly improved early childhood students' understanding of basic science concepts aligns with and extends existing theoretical frameworks regarding multimedia learning and embodied cognition. The particularly strong effects observed for plant life cycles (d = 1.03) and magnetism (d = 0.95)—concepts involving temporal processes and invisible phenomena—support the hypothesis that AR's primary pedagogical affordance lies in making abstract concepts perceptually accessible to young learners who are still predominantly concrete operational thinkers. These results resonate with Huang et al. (2019) meta-analytic findings regarding AR's effectiveness for knowledge acquisition, while extending this work specifically to early childhood populations. The current study's effect sizes (overall d = 1.11) exceed those reported in previous meta-analyses (d = 0.68), potentially attributable to the developmentally appropriate design of AR applications used here and the particular suitability of AR for learners at concrete operational stages who benefit maximally from perceptual support for abstract reasoning.

From a theoretical perspective, these findings support dual coding theory (Paivio, 1986), which posits that information encoded through both verbal and visual channels produces more robust memory traces and deeper understanding than single-channel encoding. AR applications in this study provided visual (3D animations), verbal (teacher explanations and application narration), and kinesthetic (gesture-based interaction) encoding opportunities, potentially explaining the strong learning outcomes through multiple, reinforcing cognitive pathways. The superior retention demonstrated by the AR group (99.1% vs. 78.5% of gains maintained) provides particularly compelling evidence of AR's educational value, as retention represents a more stringent test of genuine learning than immediate recall.

These findings address a critical gap in previous literature, which has rarely examined long-term retention following AR interventions (Ibáñez & Delgado-Kloos, 2018). The durability of AR-facilitated learning may stem from the memorable, distinctive quality of AR experiences, which create rich episodic memories that support subsequent concept retrieval—a mechanism consistent with encoding specificity and distinctiveness principles in cognitive psychology. However, alternative explanations warrant consideration. The superior performance of AR students might partially reflect novelty effects or increased motivation rather than purely cognitive mechanisms. While the sustained effects at delayed post-test argue against simple novelty explanations (which would predict dissipation over time), future research should examine whether AR advantages persist across extended implementation periods when novelty naturally diminishes. Additionally, the possibility of differential teacher enthusiasm or expectancy effects, despite protocol standardization efforts, cannot be entirely excluded as contributing factors.

Amplified Student Engagement and Motivation

The very large effects observed for student engagement (composite d = 1.43) represent a particularly important finding given the foundational role of engagement in learning processes. Emotional engagement showed the strongest effects (d = 1.38), suggesting that AR experiences evoked positive affective responses that may have facilitated approach behavior, sustained attention, and cognitive resource allocation toward learning tasks—mechanisms consistent with control-value theory of achievement emotions (Pekrun, 2006). These engagement findings complement and extend recent work by Reisoğlu et al. (2017), who reported positive motivational effects of AR in early childhood but did not employ systematic observational measures of multiple engagement dimensions as conducted here. The current study's behavioral observation methodology provides more objective evidence of engagement than self-report measures alone, strengthening confidence in the reality of observed effects.

The observation that AR students frequently initiated peer discussions and collaborative problem-solving suggests that AR may support social learning processes in addition to individual cognitive effects. This social dimension aligns with Vygotskian perspectives emphasizing the role of collaborative meaning-making in cognitive development. AR applications appeared to function as shared objects of joint attention, providing concrete referents for peer dialogue and collaborative knowledge construction—processes identified as particularly important for early childhood learning.

Nevertheless, it is crucial to distinguish between engagement as a mediator of learning versus as an outcome valuable in itself. While engagement likely facilitates learning by directing cognitive resources and extending time on task, high engagement alone does not guarantee conceptual understanding. The current study's design cannot definitively establish whether engagement mediated the relationship between AR instruction and learning outcomes. Future research employing path analysis or mediation models could clarify these causal pathways. Additionally, the sustainability of engagement effects requires investigation. Teachers noted that while initial excitement was high, engagement moderated over time, indicating that whether engagement stabilizes at elevated levels represents an important question for longitudinal research.

Integration of Physical and Digital Learning Experiences

The regression analyses indicating that integration with physical materials significantly predicted learning outcomes merit particular discussion, as this finding challenges purely technological determinism and emphasizes pedagogical design. The most effective AR learning episodes observed in this study combined digital visualizations with concurrent physical manipulation and teacher-mediated sense-making—a "blended reality" approach that maintained the hands-on, concrete experiences valued in early childhood education while augmenting them with digital enhancements. This finding aligns with research by Crescenzi-Lanna & Grané-Oró (2016),who argued that effective educational technology in early childhood complements rather than replaces traditional materials and activities.

In the current study, AR applications served not as standalone learning tools but as components within carefully orchestrated instructional sequences that maintained emphasis on direct physical interaction and social construction of knowledge. From a practical perspective, this has important implications for educational technology design and implementation. Rather than conceptualizing AR as a replacement for conventional early childhood teaching methods, educators should view AR as an enhancement tool strategically deployed at specific instructional moments where its unique affordances—primarily visualization and interactivity—address particular pedagogical needs.

The "augmented" in Augmented Reality should apply not only to visual perception but to the entire pedagogical approach, suggesting hybrid models that thoughtfully integrate traditional and technology-enhanced methods. The importance of teacher mediation emerged clearly in both quantitative predictors and qualitative observations. AR applications themselves did not teach; rather, they provided representations and interactions that became educationally meaningful through teacher-facilitated dialogue, scaffolding, and explicit connections to conceptual principles.

Developmental Appropriateness and Design Considerations

The successful implementation of AR technology with young children in this study provides evidence that appropriately designed AR applications can be developmentally suitable for early childhood populations—a question that has generated debate among early childhood educators and developmental psychologists. The similar perceived difficulty ratings between AR and control conditions suggest that AR did not reduce cognitive challenge, dispelling concerns that technology might oversimplify learning or reduce it to entertainment.

Several design features likely contributed to developmental appropriateness: (1) tablet-based rather than head-mounted displays, maintaining normal visual fields and reducing motion sickness risks; (2) marker-based tracking providing predictable, controllable activation rather than overwhelming spontaneous overlays; (3) large, high-contrast visuals accommodating young children's developing visual acuity; (4) gesture-based rather than text-based interaction respecting emergent literacy; and (5) scaffolded content progressions from concrete to abstract aligned with cognitive development principles.

These design considerations reflect principles articulated by NAEYC and Fred Rogers Center (2012) regarding developmentally appropriate technology use. The positive engagement and learning outcomes observed here suggest that when these principles guide design, technology can be incorporated into early childhood education in ways consistent with child development best practices rather than conflicting with them. However, important caveats apply. This study examined relatively brief AR sessions (30 minutes biweekly) rather than extended daily exposure. Questions about appropriate dosage, potential negative effects of excessive screen time, and displacement of other valuable activities (outdoor play, creative arts, social interaction) remain important considerations for implementation policies.

The current findings support AR as a valuable pedagogical tool when used judiciously, not as a comprehensive instructional approach. Additionally, individual differences in response to AR warrant investigation. While group-level effects were positive, observational notes indicated that a small subset of children showed initial anxiety or confusion with AR technology, requiring additional scaffolding and reassurance. Understanding which children benefit most from AR versus those who might need alternative approaches represents an important direction for differentiated instruction research.

Implementation Feasibility in Resource-Constrained Contexts

An important contribution of this research is demonstrating AR implementation feasibility within the Indonesian educational context, characterized by resource constraints typical of many developing nations. Unlike previous AR research conducted primarily in well-resourced Western settings. this study addressed practical challenges including limited device availability (shared rather than 1:1 ratios), inconsistent internet connectivity (applications functioned offline after initial download), constrained teacher training time, and integration with existing curriculum frameworks.

Teachers' moderate ratings of implementation challenges (M = 2.87/5) and high intentions for continued use (M = 4.43/5) suggest that AR integration, while not without obstacles, is perceived as manageable and worthwhile by practitioners working under realistic constraints. The technical issues reported—primarily application freezing and marker recognition difficulties—identify specific areas for improvement in AR application development for educational markets in developing nations, where device specifications may be lower than in Western contexts.

The successful implementation with minimal technology infrastructure (basic tablets, no special networking requirements, offline functionality) challenges assumptions that educational technology innovations require high-tech environments. This finding has important equity implications, suggesting that barriers to technology-enhanced learning may be more surmountable than sometimes assumed, particularly when design thoughtfully addresses resource constraints. However, scalability challenges should not be underestimated. This study benefited from researcher-provided devices, technical support, and intensive teacher training—resources that may not be available in typical implementation scenarios.

Sustainable scaling would require addressing device procurement and maintenance, ongoing professional development systems, and technical support infrastructures—investments that education systems in developing nations may struggle to provide. Additionally, the cultural appropriateness of AR content deserves attention. Applications used in this study were developed collaboratively with Indonesian educators and reflected local contexts, but many commercial AR educational applications originate from Western contexts and may contain culturally irrelevant examples, assumptions, or values. Future research should examine how cultural contextualization of AR content influences effectiveness and acceptability across diverse global settings.

Limitations and Constraints

Several limitations constrain the generalizability and interpretation of findings. First, the quasi-experimental design, while methodologically appropriate given practical constraints, limits causal inference compared to true experimental designs. Although baseline equivalence was established and treatment fidelity was high, unmeasured confounds could potentially influence results. Random assignment of intact classrooms rather than individual students represents a limitation inherent to educational research in authentic settings.

Second, the sample was drawn from urban Indonesian kindergartens serving middle-class families, potentially limiting generalizability to rural settings, different socioeconomic contexts, or other cultural contexts. Effects may differ in populations with less prior technology exposure or different educational traditions. Third, the relatively brief intervention period (eight sessions over eight weeks) and follow-up period (four weeks) prevent conclusions about very long-term retention or sustained motivation effects. While the delayed post-test provides evidence of retention superior to conventional methods, truly longitudinal research examining outcomes months or years later would strengthen understanding of AR's enduring impact.

Fourth, the study examined AR effectiveness for four specific science concepts. Generalization to other science domains, other subject areas, or different age groups requires empirical verification rather than assumption. The particularly strong effects for concepts involving invisible phenomena or temporal processes suggest AR may be differentially effective depending on content characteristics. Fifth, potential observer effects and social desirability biases in teacher reports cannot be entirely eliminated. Although engagement observations were conducted by trained, partially blinded observers, complete blinding was impossible given the obvious differences between conditions. Teacher perceptions may have been influenced by positive expectancies regarding innovative technology.

Sixth, the study did not examine potential negative effects or risks associated with AR use, such as eye strain, reduced physical activity, technology dependence, or displacement of other valuable learning activities. A comprehensive evaluation of AR integration would require examining potential costs alongside benefits. Finally, the study focused on immediate educational outcomes without addressing broader questions about technology's role in early childhood, values underlying technology integration decisions, or potential unintended consequences of increasingly technology-mediated educational experiences for young children.

Implications for Educational Practice

Despite limitations, several practical implications emerge for educators, curriculum developers, and administrators considering AR integration in early childhood science education:

1. Strategic Implementation: AR should be viewed as a strategic tool for specific instructional moments rather than a comprehensive curricular approach. It appears most valuable for concepts involving visualization challenges, dynamic processes, or abstract relationships difficult to represent through traditional methods.
2. Physical-Digital Integration: Effective AR implementation maintains integration with physical manipulatives and hands-on activities rather than replacing them. Hybrid approaches combining AR visualizations with concrete materials and social interaction produced the strongest observed outcomes.
3. Teacher Mediation: AR applications require active teacher facilitation, scaffolding, and conceptual connections rather than functioning as independent learning tools. Professional development should emphasize pedagogical strategies for leveraging AR, not merely technical operation.
4. Collaborative Learning Structures: Pairing students to share devices encourages dialogue, collaboration, and peer scaffolding—social processes that appeared to enhance AR's educational value in this study.
5. Appropriate Dosage: Brief, focused AR sessions (20-30 minutes) integrated within broader instructional sequences appear sufficient for learning benefits without risks of excessive screen time or activity displacement.
6. Quality Design Features: High-quality 3D visualizations, gesture-based interactivity, and seamless marker recognition represent critical design elements. Educators should prioritize applications featuring these

characteristics and avoid low-quality implementations that may frustrate rather than facilitate learning.

CONCLUSION

The findings demonstrate that AR-based learning media provides effective solutions to key challenges in early childhood science education: (1) the visualization problem was addressed through 3D animations making invisible phenomena perceptible; (2) the engagement challenge was resolved through interactive, multisensory experiences capturing sustained attention; (3) the retention difficulty was mitigated through memorable, distinctive learning episodes supporting long-term knowledge durability; and (4) the abstraction gap was bridged through seamless integration of digital representations with physical manipulatives and social interaction. However, several limitations must be acknowledged. First, the quasi-experimental design and relatively short intervention period (8 weeks) limit causal claims and long-term generalizability. Second, the sample drawn from urban Indonesian kindergartens may not represent rural or different socioeconomic contexts. Third, the study examined only four science concepts, and AR's effectiveness may vary across different content domains. Fourth, potential observer effects and novelty bias cannot be entirely eliminated despite methodological safeguards. Finally, the study did not examine possible negative effects such as excessive screen time or displacement of other valuable learning activities.

ACKNOWLEDGEMENT

The authors thank the participating schools, teachers, students, and families for their generous cooperation and contribution to this research.

REFERENCES

Akçayır, M., & Akçayır, G. (2017a). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review*, *20*, 1–11.

Bada, S. O., & Olusegun, S. (2015). Constructivism learning theory: A paradigm for teaching and learning. *Journal of Research & Method in Education*, *5*(6), 66–70.

Bailey, K. M., & Christian, D. (2021a). *Research on teaching and learning English in under-resourced contexts*. Routledge New York, NY.

Crescenzi-Lanna, L., & Grané-Oró, M. (2016a). An analysis of the interaction design of the best educational apps for children aged zero to eight. *Comunicar: Revista Científica de Comunicacíon y Educacíon*, *24*(46), 77–85.

Garzón, J., Pavón, J., & Baldiris, S. (2019). Systematic review and meta-analysis of augmented reality in educational settings. *Virtual Reality*, *23*(4), 447–459.

Gilligan-Lee, K. A., Hawes, Z. C. K., Williams, A. Y., Farran, E. K., & Mix, K. S. (2023). Hands-On: Investigating the role of physical manipulatives in spatial training. *Child Development*, *94*(5), 1205–1221.

Hirsh-Pasek, K., Zosh, J. M., Golinkoff, R. M., Gray, J. H., Robb, M. B., & Kaufman, J. (2015). Putting education in "educational" apps: Lessons from the science of learning. *Psychological Science in the Public Interest*, *16*(1), 3–34.

Huang, K.-T., Ball, C., Francis, J., Ratan, R., Boumis, J., & Fordham, J. (2019). Augmented versus virtual reality in education: An exploratory study examining science knowledge retention when using augmented reality/virtual reality mobile applications. *Cyberpsychology, Behavior, and Social Networking*, *22*(2), 105–110.

Ibáñez, M.-B., & Delgado-Kloos, C. (2018). Augmented reality for STEM learning: A systematic review. *Computers & Education*, *123*, 109–123.

Leyva, D., Tamis-LeMonda, C. S., Yoshikawa, H., Jimenez-Robbins, C., & Malachowski, L. (2017). Grocery games: How ethnically diverse low-income mothers support children's reading and mathematics. *Early Childhood Research Quarterly*, *40*, 63–76.

Martín-Gutiérrez, J., Mora, C. E., Añorbe-Díaz, B., & González-Marrero, A. (2017). Virtual technologies trends in education. *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(2), 469–486.

Maxwell, S. E., Delaney, H. D., & Kelley, K. (2017). *Designing experiments and analyzing data: A model comparison perspective*. Routledge.

Patel, S. R., Margolies, P. J., Covell, N. H., Lipscomb, C., & Dixon, L. B. (2018). Using instructional design, analyze, design, develop, implement, and evaluate, to develop e-learning modules to disseminate supported employment for community behavioral health treatment programs in New York State. *Frontiers in Public Health*, *6*, 113.

Pellas, N., Fotaris, P., Kazanidis, I., & Wells, D. (2019). Augmenting the learning experience in primary and secondary school education: A systematic review of recent trends in augmented reality game-based learning. *Virtual Reality*, *23*(4), 329–346.

Redondo, B., Cózar-Gutiérrez, R., González-Calero, J. A., & Sánchez Ruiz, R. (2020). Integration of augmented reality in the teaching of English as a foreign language in early childhood education. *Early Childhood Education Journal*, *48*(2), 147–155.

Reisoğlu, İ., Topu, B., Yılmaz, R., Karakuş Yılmaz, T., & Göktaş, Y. (2017). 3D virtual learning environments in education: A meta-review. *Asia Pacific Education Review*, *18*(1), 81–100.

Saçkes, M., Trundle, K. C., & Shaheen, M. (2025). Parental Motivational Beliefs Predict Science Learning Opportunities in Early Years. *Early Childhood Education Journal*, 1–15.

Sharma, P., & Pandher, J. S. (2018). Teachers' professional development through teachers' professional activities. *Journal of Workplace Learning*, *30*(8), 613–625.

Yilmaz, R. M. (2016). Educational magic toys developed with augmented reality technology for early childhood education. *Computers in Human Behavior*, *54*, 240–248.

Yusa, I. W., Wulandari, A. Y. R., Tamam, B., Rosidi, I., Yasir, M., & Bagus Setiawan, A. Y. (2023). Development of augmented reality (ar) learning media to increase student motivation and learning outcomes in science. *Jurnal Inovasi Pendidikan IPA*, *9*(2), 127–145.