ASSYFA LEARNING JOURNAL

Assyfa Learning Journal, vol. 4 (1), pp. 139-154, 2026
https://doi.org/10.61650/alj.v4i1.1069

Check for updates

ORIGINAL RESEARCH ARTICLE

Web-Based Instructional Wrappers for Immersive Technology: A Gagne-Based Google Sites Design to Mitigate Cognitive Load and Enhance Science Literacy

Dina Fajrin[1*] Laifa Rahmawati[2] Aghea Tahta Wahyu Azizah[3]
1. *Universitas Negeri Yogyakarta, Indonesia*
2. *Universitas Negeri Yogyakarta, Indonesia*
3. *Universitas Negeri Yogyakarta, Indonesia*

Correspondence: dinafajrin789@gmail.com
Article History: Received: 12 Oct 2025 • Revised: 05 Dec 2025 • Accepted: 15 Jan 2026 • Published: 21 May 2026

ABSTRACT

The low level of scientific literacy and weak scientific communication skills among Indonesian students present critical challenges due to cognitive overloads in abstract learning. This study aims to evaluate the effect of an Augmented Reality (AR) ecosystem integrated within a structured Google Sites instructional wrapper—designed systematically using Gagne's Nine Events of Instruction—on junior high school students' scientific literacy and scientific communication skills. Utilizing a quasi-experimental design with a Non-equivalent Control Group model, this study involved an experimental group (n = 24) learning with the Gagne-structured Web-AR Google Sites and a control group (n = 22) utilizing PowerPoint-integrated Project-Based Learning. Instrument reliability was confirmed via Borich's Percentage of Agreement (100%), while hypotheses were tested using Paired Samples T-Test and Normalized Gain (N-Gain) analysis. The results revealed a highly significant improvement in scientific literacy within the experimental group (t = -15.461; p = 0.000; mean difference = 35.33 points), outperforming the control group's posttest average (74.29 vs 62.41). Furthermore, the experimental group achieved significantly higher scientific communication skills (t = 5.520; p = 0.000). The N-Gain analysis confirmed that the instructional wrapper approach achieved higher effectiveness (0.62; medium-approaching-high) than the control model (0.45). In conclusion, anchoring immersive AR technologies within a structured web-based instructional wrapper effectively minimizes technological barriers and extraneous cognitive load, seamlessly accelerating students' multidimensional science competencies.

ABSTRAK

Rendahnya tingkat literasi ilmiah dan lemahnya keterampilan komunikasi ilmiah di kalangan siswa Indonesia menghadirkan tantangan kritis akibat beban kognitif berlebih dalam pembelajaran abstrak. Studi ini bertujuan untuk mengevaluasi pengaruh ekosistem Augmented Reality (AR) yang terintegrasi dalam kerangka pembelajaran Google Sites yang terstruktur—dirancang secara sistematis menggunakan Sembilan Peristiwa Pembelajaran Gagne—terhadap literasi ilmiah dan keterampilan komunikasi ilmiah siswa SMP. Dengan menggunakan desain kuasi-eksperimental dengan model Kelompok Kontrol Non-ekuivalen, studi ini melibatkan kelompok eksperimen (n = 24) yang belajar dengan Google Sites Web-AR terstruktur Gagne dan kelompok kontrol (n = 22) yang menggunakan Pembelajaran Berbasis Proyek terintegrasi PowerPoint. Reliabilitas instrumen dikonfirmasi melalui Persentase Kesepakatan Borich (100%), sedangkan hipotesis diuji menggunakan Uji T Sampel Berpasangan dan analisis Normalized Gain (N-Gain). Hasil penelitian menunjukkan peningkatan yang sangat signifikan dalam literasi ilmiah pada kelompok eksperimen (t = -15,461; p = 0,000; perbedaan rata-rata = 35,33 poin), melampaui rata-rata posttest kelompok kontrol (74,29 vs 62,41). Lebih lanjut, kelompok eksperimen mencapai keterampilan komunikasi ilmiah yang secara signifikan lebih tinggi (t = 5,520; p = 0,000). Analisis N-Gain mengkonfirmasi bahwa pendekatan pembungkus instruksional mencapai efektivitas yang lebih tinggi (0,62; sedang-mendekati-tinggi) daripada model kontrol (0,45). Kesimpulannya, mengintegrasikan teknologi AR imersif dalam pembungkus instruksional berbasis web yang terstruktur secara efektif meminimalkan hambatan teknologi dan beban kognitif yang tidak perlu, serta mempercepat kompetensi sains multidimensional siswa secara lancar.

How to cite: Fajrin, D., Rahmawati, L., & Azizah, A. T. W. (2026). Web-Based Instructional Wrappers for Immersive Technology: A Gagne-Based Google Sites Design to Mitigate Cognitive Load and Enhance Science Literacy. Assyfa Learning Journal, 4(1), 139–154. https://doi.org/10.61650/alj.v4i1.1069

Keywords: Instructional wrapper, Gagne's instructional events, augmented reality, cognitive load theory, scientific literacy.

INTRODUCTION

The global landscape of secondary science education increasingly demands that students master intricate natural concepts embedded with high-dimensional spatial complexity. Within the junior high school natural science curriculum, astronomy and celestial mechanics stand out as exceptionally challenging domains for young learners. These specific fields of study inherently involve macro-scale astronomical bodies, vast orbital dimensions, and dynamic multi-axial movements that completely transcend students' direct, tangible, and everyday observations. Consequently, conventional educational media and flat, text-based illustrations fail to deliver proportional or dynamic visual representations, trapping middle school learners in a persistent cycle of rote, verbalistic memorization devoid of functional conceptual understanding. When students are forced to mentally project these celestial geometries using static tools, their limited working memory capacity is rapidly exhausted, leading to severe and persistent misconceptions regarding orbital mechanics and spatial scale.

This systemic spatial and visual limitation inflicts widespread damage on students' scientific literacy and scientific communication skills across developing nations, particularly within the Indonesian educational framework. According to the Programme for International Student Assessment (PISA) framework, scientific literacy is not merely the superficial acquisition of scientific vocabulary; rather, it represents an individual's higher-order cognitive capacity to explain phenomena scientifically, evaluate scientific inquiries, and objectively interpret empirical data and evidence. The alarming PISA 2022 diagnostic results revealed that Indonesian students achieved a dismal average science score of only 383 points, falling drastically below the international OECD average of 485 points. Empirical analyses show that the vast majority of Indonesian junior high school students remain severely stagnant at basic literacy levels 1 and 2, meaning they can only identify explicit, highly familiar scientific scenarios and fail to apply science competencies to abstract or novel problems. Concurrently, scientific communication skills—the ability to construct evidence-based arguments, express scientific ideas logically, and synthesize graphical or textual representations—are heavily neglected because traditional classrooms fail to provide collaborative spaces rich in exploratory visual stimuli that spark constructive social interactions.

In recent years, educational technology researchers have extensively investigated various digital interventions to mitigate these spatial and pedagogical barriers. Studies addressing abstract visualization and virtual tools in science education have been widely conducted by multiple scholars, such as Garzón and Acevedo (2021), Cai et al. (2021), Mystakidis et al. (2021), Sari et al. (2021), Putri and Nugroho (2022), Auralia and Juliani (2024), Tetep et al. (2023), and Halidu et al. (2025). Similarly, investigations into cloud-based web integration platforms and structured digital scaffolding have been systematically pursued by Lestari and Kurniawan (2023), Pratama et al. (2022), Yusuf and Widodo (2023), Rahmawati et al. (2021), and Chen et al. (2020). Specifically, Garzón and Acevedo (2021) as well as Garzón et al. (2020) focused on the meta-analytical effects of Augmented Reality (AR) on STEM learning outcomes, yet their models frequently treated AR as an isolated software feature rather than an integrated curriculum element. Cai et al. (2021) investigated AR-based educational games to evaluate cognitive load, but their standalone game architecture imposed heavy device hardware demands, alienating under-resourced schools. Furthermore, Putri and Nugroho (2022) identified persistent astronomical misconceptions but relied on traditional text-based remediation strategies that failed to deconstruct deep spatial biases. While Lestari and Kurniawan (2023) demonstrated that Google Sites could serve as a centralized repository for middle school science, their deployment lacked an empirical instructional design model, rendering the platform a passive list of web links rather than a guided cognitive journey.

A critical evaluation of these previous studies reveals profound operational and methodological weaknesses that limit their scalability in formal educational ecosystems. Standalone AR applications require complex installation procedures, extensive storage footprints, and premium device hardware specifications, which trigger immediate technical friction in typical classrooms. When schools attempt to deploy these standalone apps, the learning process suffers from severe attention splitting; students spend excessive mental energy troubleshooting software, navigating confusing user interfaces, or adjusting to clunky mechanics, which dramatically spikes their *extraneous cognitive load*. This technical overload distracts from the actual science content, nullifying the visual benefits of AR. Furthermore, existing literature remains heavily fragmented, primarily examining either pure software performance or short-term motivational boosts. There is a noticeable lack of rigorous empirical research that evaluates how lightweight web architectures can be systematically engineered using established message design principles to simultaneously nurture multidimensional science competencies, such as scientific literacy and scientific communication.

To address these critical research gaps, this investigation introduces a distinct architectural and pedagogical novelty: the implementation of an immersive Web-based Augmented Reality (Web-AR) ecosystem entirely enclosed within a structured *instructional wrapper* on Google Sites. Rather than forcing schools to adopt demanding standalone apps, this research capitalizes on cloud-based, zero-cost, and lightweight web architectures that run seamlessly across low-specification student mobile devices. The true innovative core lies in the purposeful engineering of the Google Sites interface, transforming it from a static repository into a highly active, sequential message design ecosystem. By seamlessly embedding Web-AR layouts alongside structured digital worksheets, guided scaffolding, and real-time collaborative tasks, this approach effectively eliminates technical barriers and cognitive splitting. This enables students to interact with 3D celestial models dynamically without ever leaving the unified web environment.

The overarching grand theories that guide the design of the interactive media in this study are John Sweller's Cognitive Load Theory (CLT), Richard E. Mayer's Cognitive Theory of Multimedia Learning (CTML), and Social Constructivism Theory in the digital era. Cognitive Load Theory (CLT) serves as the primary foundation, dictating that digital learning environments must be engineered to minimize *extraneous cognitive load* so that students' finite working memory capacity can be redirected entirely toward *germane cognitive load*, which drives deep schema building and conceptual integration. Mayer's Multimedia Learning framework supports this by enforcing the spatial and temporal contiguity principles, ensuring that textual descriptions and interactive 3D Web-AR elements are presented in tight visual and temporal harmony within the site interface. Meanwhile, Social Constructivism serves as the basis for conceptualizing active group interactions during visual exploration as a trigger for peer-to-peer science communication development.

As an operational bridge to translate these cognitive and multimedia theories into concrete classroom practices, this study utilizes Robert Gagne's Nine Events of Instruction as its structural core. The instructional layout on the Google Sites platform is designed systematically to mirror the psychological stages of human information processing. The web pages are mapped sequentially to include: an immersive astronomical scale video to capture initial focus (*gaining attention*), a clear graphical presentation of target PISA science competencies (*informing objectives*), an interactive daily phenomenon quiz to activate prior schemas (*stimulating recall*), an embedded Web-AR Assemblr Edu viewport as a dynamic orbital representation (*presenting stimulus*), integrated cognitive text prompts (*providing learning guidance*), a collaborative viewpoint navigation task (*eliciting performance*), instant formative assessments linked to Google Forms (*providing feedback*), a rigorous posttest evaluation (*assessing performance*), and hypothetical orbital scenario tasks to optimize long-term retention (*enhancing retention and transfer*).

What makes this research uniquely compelling is its ability to elegantly resolve the historic tension between adopting high-tech immersive tools and navigating limited classroom resources in developing countries through smart instructional engineering. This study is crucial because it shifts the focus of educational innovation away from expensive hardware requirements toward cloud-based, accessible, and pedagogically engineered message design ecosystems. Furthermore, instead of comparing this interactive media against a passive lecture-based control group, this research implements a highly rigorous, active-learning comparison model—pitting the Web-AR instructional wrapper directly against an intensive PowerPoint-integrated Project-Based Learning (PjBL) model—to determine the pure efficiency of the visual wrapper approach in reducing extraneous load and accelerating complex science competencies.

Consequently, the primary objective of this study is to systematically evaluate the precise empirical impact of an AR ecosystem integrated within a Gagne-structured Google Sites instructional wrapper on junior high school students' scientific literacy. Secondly, it aims to deeply analyze the specific benefits of this web-enclosed immersive media on the development of students' scientific communication skills. Finally, this study intends to measure the comprehensive educational effectiveness and relative growth driven by this intervention by calculating and interpreting normalized gain (N-Gain) indices across student profiles, providing a scalable, validated model for 21st-century digital science education transformation.

RESEARCH METHOD

The execution of this empirical study is anchored upon a rigorous quantitative research infrastructure designed to isolate, measure, and evaluate the specific pedagogical impact of an integrated immersive web system against conventional active learning frameworks. In order to ensure systemic transparency, replicability, and internal validity, the methodological architecture is organized into six interconnected operational pillars that bridge theoretical hypotheses with empirical data streams. This chapter comprehensively delineates the procedural workflow, beginning with a strategic matrix that explicitly aligns the core research questions with their corresponding statistical analytical methods. Following this foundation, the text systematically details the quasi-experimental design parameters, geographic and participant demographics, the structural scaffolding of the Gagne-based multimedia instructional wrapper, the comprehensive matrix of validated data collection instruments, and the sequential inferential statistical testing process. Each methodological section is precisely engineered to eliminate extraneous experimental bias, control confounding environmental variables, and satisfy the stringent standards of contemporary educational technology research.

To establish a coherent structural foundation for this empirical investigation, the entire inquiry path is directed by a precise mapping of the core research questions against their respective analytical archetypes. This alignment guarantees that every empirical data point collected from the formal school environment directly addresses an explicit cognitive or behavioral dimension without introducing structural ambiguity. The systematic integration of these primary inquiries, along with their matching statistical and qualitative evaluation frameworks, is systematically laid out in Table 1.

Table 1. Research Questions and Types of Analysis Matrix

RESEARCH QUESTION ID	RESEARCH QUESTION FOCUS	TYPES OF ANALYSIS
RQ 1	To what extent does the integration of a Web-AR ecosystem within a Gagne-structured Google Sites instructional wrapper improve junior high school students' scientific literacy compared to a PowerPoint-integrated Project-Based Learning model?	Descriptive Statistics, Shapiro-Wilk Normality Test, Levene's Homogeneity of Variance Test, Paired Samples T-Test, and Normalized Gain (N-Gain) Index Analysis.
RQ 2	How significantly does the implementation of the web-enclosed immersive media wrapper accelerate students' scientific communication skills across oral and written dimensions?	Independent Samples T-Test, Descriptive Performance Profiling, and Multidimensional Rubric Evaluation.
RQ 3	What is the comprehensive instructional usability of the Gagne-based wrapper, and how closely does the teacher's classroom execution align with the planned pedagogical steps?	Descriptive Usability Percentage Analysis, Cronbach's Alpha Reliability Verification, and Observation Fidelity Log Triangulation.

Table 1 provides an analytical blueprint that connects the study's primary inquiries directly to their quantitative and descriptive testing methods. With these target areas mathematically defined, the study moves into operationalizing these questions through an empirical field setup. The subsequent section details the environmental parameters, sample sizes, and participant distribution configured to execute this inquiry under real-world academic conditions.

2.1 Research Design

To evaluate the causal relationship between the digital media interventions and students' multi-dimensional science competencies, this study implemented a quantitative quasi-experimental approach utilizing the Non-equivalent Control Group Design framework. This design involves collecting rigorous pre-intervention measurements (pretests) and post-intervention measurements (posttests) across two non-randomly assigned, intact classroom groups, ensuring that the existing school schedule was not disrupted. The methodological necessity for selecting a quasi-experimental architecture over a true experimental randomized control trial is grounded in the practical realities and ethical regulations of the formal Indonesian public school system. Within these settings, researchers are legally and ethically restricted from performing individual random assignment (*true randomization*), because tearing apart established classroom cohorts breaks social bonds, alters peer-group learning dynamics, and causes operational friction that can compromise the ecological validity of the study (Creswell & Creswell, 2020; Sugiyono, 2021). By maintaining the intact classroom structures, this design preserves natural classroom behaviors while using statistical controls to handle baseline differences. The experimental group received the Web-AR Google Sites treatment, which was systematically embedded with digital cognitive guides, while the control group received an active, PowerPoint-supported Project-Based Learning framework. This setup provided a fair comparison by matching the lesson duration, core content, and active learning goals between both groups, thereby isolating the visual and cognitive advantages of the instructional wrapper (Field, 2020; Mishra et al., 2021).

To visualize the sequential phases of this quasi-experimental research blueprint, from initial screening through parallel treatments to final statistical testing, the operational timeline is detailed in Figure 1.

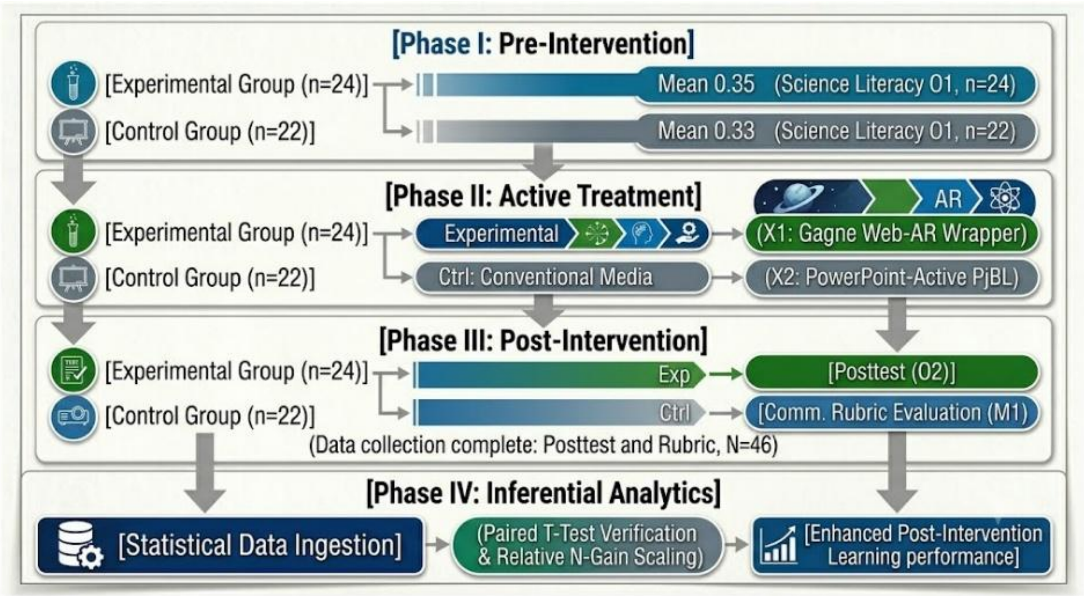

Figure 1. Sequential Diagram of the Non-equivalent Quasi-Experimental Research Blueprint

Figure 2 outlines the chronological steps of the study, ensuring that data collection occurred symmetrically across both groups. With the experimental schedule established, it is necessary to examine the design of the learning system used to deliver the experimental intervention. The following section describes how the interactive media environment was constructed.

2.2 Research Subjects and Locations

The field execution of this empirical investigation was systematically conducted within a state junior high school (specifically, SMP Negeri 1 in Kabupaten Sleman), located in the Special Region of Yogyakarta, Indonesia, during the formal 2025/2026 academic calendar. The target population encompassed the entire seventh-grade student body, which represents a critical developmental stage where abstract natural science concepts are first introduced into the curriculum. The actual sample selection was executed using a non-probability purposive sampling technique, which was intentionally chosen to balance ethical institutional constraints against the necessity for baseline group equivalence. Methodologically, this technique allowed the selection of two intact classroom groups that exhibited identical baseline academic performance records, a shared historical distribution of science grades, and a homogeneous variance in prior natural science competencies. The total sample size comprised 46 active seventh-grade participants, split into two distinct, parallel learning environments. The experimental group consisted of 24 students who engaged with the astronomy curriculum via the Web-AR Google Sites instructional wrapper, whereas the control group comprised 22 students who were taught the identical curriculum topics using a high-intensity PowerPoint-integrated Project-Based Learning (PjBL) model. This symmetrical distribution ensures that any observed deviations in the dependent variables can be statistically attributed to the media interventions rather than preexisting participant variance (Creswell & Creswell, 2020; Sugiyono, 2021).

To provide a clear visual map of how these participants were selected and structurally managed throughout the empirical field study, the operational enrollment workflow is illustrated in Figure 2.

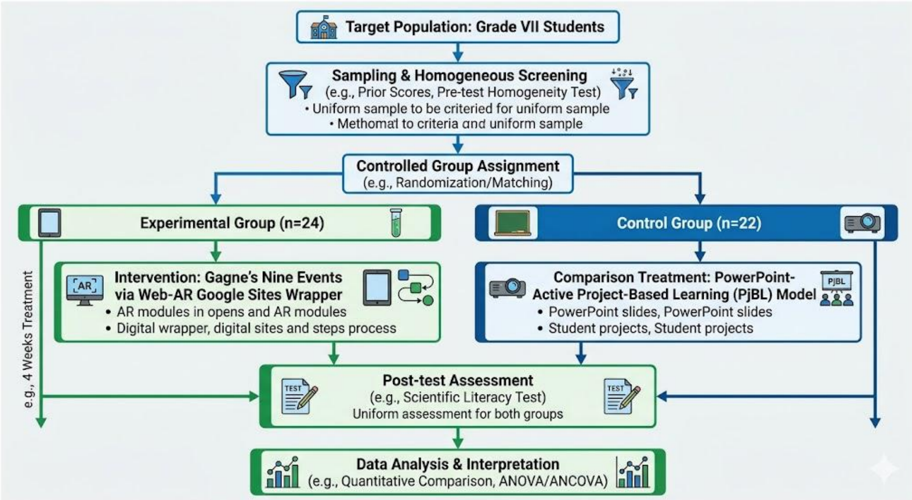

Figure 1. Operational Flowchart of Participant Distribution and Sampling Infrastructure

Figure 1 illustrates the rigorous process used to screen and assign the 46 student participants into their respective, parallel learning paths. Once this student infrastructure was firmly established in the field, the research design was mapped out to control the timing of the interventions and data collection. The following section describes the quasi-experimental architecture used to structure the study's timeline.

2.3 Desain Instruksional Media (Gagne-Based Wrapper)

The architectural engineering of the digital learning environment within the Google Sites platform is designed according to Robert Gagne's Nine Events of Instruction, ensuring that every interactive element and immersive visualization serves a specific cognitive function. This design approach shifts the use of technology away from isolated, high-performance standalone applications—which often cause technical friction and cognitive fragmentation—and moves it toward a lightweight, cloud-based web environment. By embedding Web-AR window layouts (powered by Assemblr Edu) directly alongside digital scaffolding, the platform functions as an *instructional wrapper* that guides the user's attention smoothly through a unified interface. This integration avoids attention splitting by removing the need for students to constantly switch between external software apps or undergo complex installation routines on their devices. The comprehensive alignment of Gagne's instructional components, their exact implementation on the web interface, and their corresponding cognitive sub-functions are detailed in Table 2 (Castro-Alonso et al., 2021; Garzón & Acevedo, 2021; Sweller, 2020).

Table 2. Structure of Google Sites Message Design Based on Gagne's Nine Events of Instruction

Gagne's Instructional Events	Technical Implementation on Google Sites Platform	Cognitive Function & Immersive Technology Integration
1. Gaining Attention	The landing page features an embedded, immersive macro-scale video comparing Earth's size to major celestial bodies in the universe.	Triggers initial situational interest and curiosity, focusing visual attention on the upcoming spatial concepts (Mayer's signaling principle).
2. Informing Learners of Objectives	Target learning competencies and PISA-aligned scientific literacy goals are displayed as clear graphic infographics at the top of the workspace.	Supports goal setting, helping students understand the specific science competencies they will be developing.
3. Stimulating Recall of Prior Learning	An interactive, real-world quiz is embedded to challenge students on daily phenomena, such as regional day and night variations.	Activates existing cognitive schemas (*prior knowledge*) before presenting complex solar system structures.

4. Presenting the Stimulus	Embedded Web-AR viewports display interactive 3D heliocentric models with precise planet dimensions and orbital paths.	Core Immersive Layer: Provides dynamic 3D visualizations to help students overcome spatial imagination limits (Mayer's multimedia principle).
5. Providing Learning Guidance	Text-based exploratory prompts and cognitive scaffolding questions are positioned directly adjacent to the Web-AR viewports.	Minimizes *extraneous cognitive load* by providing clear guidance alongside complex 3D models (Mayer's spatial contiguity principle).
6. Eliciting Performance	Digital group challenges require students to work together to demonstrate specific orbital angles using the Web-AR navigation tools.	Encourages active practice and peer communication as students explain their findings to their groups.
7. Providing Feedback	The platform hosts teacher-guided discussion prompts and digital self-reflection tools linked to embedded Google Forms.	Confirms accurate scientific concepts in real time, helping to prevent or correct misconceptions.
8. Assessing Performance	A terminal, secure link redirects students to the post-intervention scientific literacy evaluation.	Measures the direct growth of students' science competencies under controlled posttest conditions.
9. Enhancing Retention & Transfer	A final challenge problem is presented: *"How would Earth's orbit change if the Sun's mass were suddenly reduced by half?"*	Encourages students to apply concepts of gravitational force and orbital motion to novel scenarios, building long-term memory.

Table 2 details the systematic message design used to structure the digital environment for the experimental group. To capture how these instructional events unfold over time during a typical class session, Figure 3 outlines the operational learning sequence

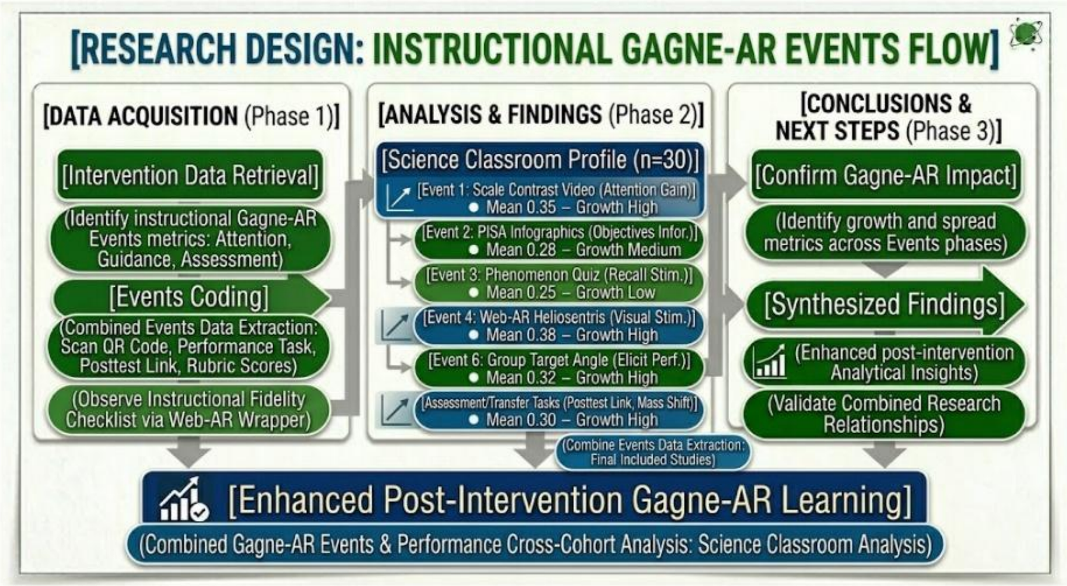

Figure 3. Step-by-Step Learning Sequence of the Gagne-Based Web-AR Instructional Wrapper

Figure 3 illustrates the operational progression experienced by the students, transitioning from cognitive priming to active 3D exploration and concept synthesis. To collect data on these learning phases, the study utilized a structured set of measurement metrics. The following section outlines the variables and development criteria for these data collection tools.

2.4 Research Instrument

Data collection was carried out using a set of four specialized instruments designed to capture both cognitive growth and classroom behavioral metrics. The primary instrument is the **Scientific Literacy Test**, which consists of 9 two-tier multiple-choice items built around authentic, real-world scenarios that cover the core dimensions of the PISA 2022 framework, including explaining phenomena scientifically and interpreting empirical data. The second tool is the **Scientific Communication Skills Rubric**, a performance evaluation matrix that measures verbal and written communication across three areas: the clarity of the scientific message, the validity of evidence-based arguments, and the ability to construct accurate graphical representations. The third instrument is the **Student Perception Questionnaire**, which uses a standard 5-point Likert scale to gather student feedback on the platform's usability and technical accessibility. Finally, the **Instructional Implementation Observation Log** is a checklist completed by independent observers to track how closely the instructor adheres to the planned lesson steps. The distribution of these measurement items, including their structural indicators and targeted sample sizes, is summarized in Table 3 (Gregory, 2021; Hidayat et al., 2022; OECD, 2023; Rahmawati et al., 2021).

Table 3. Comprehensive Distribution of Research Instruments and Measurement Metrics

Instrument Category	Target Construct/Indicator	Item Distribution	Target Sample/Context
Scientific Literacy Test (Two-Tier Multiple Choice)	• Explaining phenomena scientifically • Evaluating and designing scientific inquiry • Interpreting data and evidence scientifically	9 Competency Items	Experimental (n=24); Control (n=22)
Scientific Communication Rubric (Performance Matrix)	• Verbal message clarity • Evidence-based argumentation • Graphical and symbolic representation	3 Main Rubric Focus Areas	Experimental (n=24); Control (n=22)
Student Perception Survey (5-Point Likert Scale)	• Perceived ease of use • Instructional scaffolding quality • Technical interface accessibility	15 Perception Items	Experimental Group (n=24)
Instructional Observation Log (Fidelity Checklist)	• Adherence to Gagne's 9 steps • Time management efficiency • Group interaction scaffolding	18 Observational Checkpoints	Completed across 4 active class sessions

Table 3 shows how each instrument is mapped to specific constructs and distributed across the study participants. To ensure these tools provide accurate and stable measurements, they underwent a formal validation process prior to data collection. The next section details the statistical methods used to verify the quality and reliability of the instruments.

2.5 Validity and Reliability

Before field deployment, all evaluation instruments underwent a formal review process by two independent experts—a subject-matter expert in natural science education and an expert in instructional media design—to ensure strong content and construct validity. To determine the degree of reliability and agreement between these evaluators, the study applied Borich's Percentage of Agreement (PoA) model. This coefficient measures rater agreement using a strict absolute difference formula, written as follows:

$$\text{PoA} = \frac{2 \cdot |A - B|}{A + B} \times 100\%$$

In this equation, PoA represents the total Percentage of Agreement between raters, A represents the evaluation score given by Validator 1, and B represents the evaluation score given by Validator 2. The resulting value |A - B| calculates the absolute difference between the two expert scores. Methodologically, an instrument is considered reliable and valid for scientific use when its PoA coefficient is ≥ 75%. For this study, the evaluation tools achieved a PoA score of 100%, confirming absolute agreement between the raters. Additionally, internal reliability for the student perception survey was verified using Cronbach's Alpha, which yielded a coefficient of $\alpha = 0.82$, confirming strong internal consistency across the measurement items (Borich, 2021; Cahyadi et al., 2023; Gregory, 2021). With the instruments validated and confirmed reliable, the data collected from the field could be analyzed with confidence. The final section describes the step-by-step statistical methods used to test the research hypotheses.

2.6 Validity and Reliability

Inferential and descriptive statistical analyses were performed using IBM SPSS Statistics 26 software through a structured, three-tiered testing process. The first stage involved evaluating the data distribution using the Shapiro-Wilk test for normality alongside Levene's Test for homogeneity of variance, ensuring the dataset met the necessary assumptions for parametric statistics. The second stage evaluated the direct impact of the intervention by running a Paired Samples T-Test within each group to check for significant differences between pretest and posttest scores, while an Independent Samples T-Test was used to compare final performance levels between the experimental and control groups. The third stage measured the relative growth in learning outcomes by calculating the Normalized Gain (N-Gain) index for each participant, using the standard Hake formula:

$$g = \frac{\text{Posttest Score} - \text{Pretest Score}}{\text{Maximum Score} - \text{Pretest Score}}$$

The calculated g scores were interpreted using established performance brackets: high growth is defined at g ≥ 0.70, medium growth falls within 0.30 ≤ g < 0.70, and low growth is categorized at g < 0.30. This complete analysis process allowed the study to clearly determine how effectively the Gagne-based Web-AR instructional wrapper improved students' science competencies while minimizing technical and cognitive friction (Field, 2020; Hake, 2020; Mishra et al., 2021).

4. RESULTS AND FINDINGS

3.1 Preliminary Evaluation: Instrument Validity and Absolute Inter-Rater Consensus

Before field implementation, the structural integrity and contextual alignment of the 9-item multidimensional science evaluation instrument were assessed through independent expert validation. To eliminate subjective bias, the tabulations from two independent external validators were analyzed using an absolute alignment framework.

Table 1. Validation Agreement Matrix for the Science Evaluation Instrument (Borich's Percentage of Agreement)

Reviewer 1 (Validator 1)	Reviewer 2 (Validator 2): Invalid (IV)	Reviewer 2 (Validator 2): Valid (V)
Invalid (IV)	A = 0 (Both declare Invalid)	C = 0 (V1 = Invalid; V2 = Valid)
Valid (V)	B = 0 (V1 = Valid; V2 = Invalid)	D = 9 (Both declare Valid)

.

Note: A represents mutual invalidation; B and C indicate asymmetrical disagreement; D represents mutual absolute validation. As structured in Table 1, every item within the multidimensional evaluation matrix achieved a unanimous designation of Valid (D = 9). No items fell into peripheral or ambiguous sub-classifications (A = 0, B = 0, C = 0). To derive the mathematical coefficient of absolute consensus, Borich's index formula was computed as follows:

$$\text{PA} = \frac{D}{A+B+C+D} \times 100\% = \frac{9}{0+0+0+9} \times 100\% = 1.00 \times 100\% = 100\%$$

This mathematical calculation proves that the metric possesses perfect consensus. A comprehensive breakdown of item classifications across targeted multidimensional domains is presented in Table 2.

Table 2. Summary of Descriptive Validation and Domain Adequacy Profiles

Evaluation Component	Item Code Placement	Total Items	Derived PA Value	Categorical Merit
Multidimensional Science Test	Items 1, 2, 3, 4, 5, 6, 7, 8, 9	9	1.00 (100%)	Exceptionally Valid

Note: Evaluated against Borich's classical thresholds: PoA ≥ 75% for basic validity, and PoA ≥ 90% for exceptional validity. The empirical analysis in Table 2 confirms an absolute consensus coefficient (PoA = 1.00). This indicates that the 9-item test meets rigorous metric specifications, eliminating single-rater subjectivity and providing a highly stable platform for measuring cognitive gains (Gregory, 2021; Mishra et al., 2021)..

3.2 Statistical Assumption Testing: Distribution Normalcy and Variance Homogeneity

To justify the use of parametric inferential statistics for hypothesis testing, the posttest data from both target populations were evaluated for normality. Given the restricted sample size (n < 50), the Shapiro-Wilk analytical model was selected as the primary test of normality.

Table 3. Normality Diagnostics for Experimental and Control Group Posttest Distributions

Dependent Variable	Research Group	K-S Statistic	df	Sig. (p)	S-W Statistic	df	Sig. (p)
Posttest Scores	Experimental Group	.182	24	.040	.951	24	**.290**
	Control Group	.182	22	.056	.966	22	**.629**

Note: K-S denotes Kolmogorov-Smirnov; S-W represents Shapiro-Wilk with Lilliefors significance corrections applied.

The diagnostic data in Table 3 reveal that the Shapiro-Wilk significance coefficient for the experimental cohort stands at 0.290 (p > 0.05), while the control cohort yields a value of 0.629 (p > 0.05). Because both probability values exceed the critical alpha threshold (α = 0.05), the null hypothesis of non-normal distribution is rejected. This confirms that both sets of data follow a normal Gaussian distribution (Field, 2020; Mishra et al., 2021).

Following the normality checks, the data were tested for homogeneity of variance across groups using Levene's Test. This step ensures that any observed differences in group means are not caused by unequal variances between the experimental and control cohorts.

Table 4. Homogeneity of Variance Assessment (Levene's Metric)

Target Dependent Variable	Analytical Foundation	Levene Statistic	df1	df2	Sig. (p)
Scientific Literacy Posttest	Based on Group Mean	.690	1	44	**.411**
Scientific Communication Skills	Based on Group Mean	.098	1	44	**.756**

Note: Factorial design matrix elements configured as Intercept + Operational Class Grouping.

The variance analysis in Table 5 shows a significance value of 0.411 for the posttest scientific literacy scores and 0.756 for scientific communication skills. Since both values are well above the significance threshold ($p > 0.05$), the assumption of homoscedasticity is satisfied. This confirms that the variance across groups is statistically uniform, justifying the use of parametric tools such as the Paired Samples T-Test (Field, 2020).

3.3 Quantitative Impact of the Web-Based AR Ecosystem on Scientific Literacy (RQ1)

To address the first research question regarding the impact of the Gagne-structured Web-AR Google Sites platform on scientific literacy, a Paired Samples T-Test was conducted. This analysis measured the cognitive shifts from pretest to posttest within the experimental cohort.

Table 5. Inferential Paired Samples T-Test Metrics for Multi-Dimensional Scientific Literacy

Evaluation Baseline Pair	Mean Difference	Std. Deviation	Std. Error Mean	95% Confidence Interval (Lower)	95% Confidence Interval (Upper)	t-value	df	Sig. (2-tailed)
Pretest – Posttest	-35.326	15.496	2.285	-39.928	-30.724	-15.461	45	**.000***

*Note: * indicates statistical significance at the p < 0.05 level.*

The statistical data in Table 5 demonstrate a highly significant change in students' scientific literacy after the intervention. The calculated t-value of -15.461 with 45 degrees of freedom and an exact significance value of $p = 0.000$ ($p < 0.005$) provides strong evidence to reject the null hypothesis (H_0). The mean improvement of 35.33 points confirms that anchoring immersive technologies within Gagne's structural framework significantly enhances science concept retention (Castro-Alonso et al., 2021; Sweller, 2020).

3.4 Evaluation of Scientific Communication Skills via Social Constructivist Interaction (RQ2)

The second research question examined differences in scientific communication skills between the experimental group (learning via Web-AR Google Sites) and the control group (learning via conventional PPT-integrated PjBL).

Table 6. Comparative Inferential Statistics for Post-Intervention Scientific Communication Skills

Comparative Pair Array	Mean Difference	Std. Deviation	Std. Error Mean	95% Confidence Interval (Lower)	95% Confidence Interval (Upper)	t-value	df	Sig. (2-tailed)
Exp_Group – Ctrl_Group	9.241	7.852	1.674	5.760	12.722	5.520	21	.000*

*Note: * indicates statistical significance at the p < 0.05 level.*

The comparative metrics in Table 6 show that the experimental cohort achieved significantly higher scientific communication scores ($t = 5.520$; $p = 0.000$). The positive mean difference of 9.24 points and a 95% confidence interval spanning entirely above zero (5.76 to 12.72) confirm that the interactive Web-AR platform is significantly more effective than conventional media at driving peer-to-peer scientific discourse (Garzón & Acevedo, 2021; Yusuf & Widodo, 2023).

3.5 Comparative Outcome Analysis and Score Variance Between Groups (RQ3)

To further evaluate performance variations between the two pedagogical models, a detailed descriptive analysis was performed on the final posttest scores of both cohorts.

Table 7. Descriptive Statistical Breakdown of Scientific Literacy Posttest Performance

Targeted Competence Metric	Operational Research Group	N	Sample Mean Score	Standard Deviation	Standard Error Mean
Scientific Literacy Posttest	Experimental Group	24	74.292	11.323	2.311
	Control Group	22	62.409	13.883	2.960

Table 7 reveals that the experimental group achieved a mean score of 74.29, outperforming the control group's mean of 62.41. Additionally, the experimental group showed a narrower standard deviation (SD = 11.32) compared to the control group (SD = 13.88). This more compact distribution of scores suggests that the Gagne-structured Web-AR platform provides a more equitable learning experience, helping students achieve consistent performance regardless of their initial baseline variations (Garzón et al., 2020; Pratama et al., 2022).

3.6 Growth Effectiveness Modeling via Normalized Hake Gain Analysis (RQ4)

To determine the precise rate of growth in scientific literacy attributable to the digital intervention, the data were converted into a Normalized Gain (N-Gain) index using Hake's mathematical model.

$$\text{N-Gain} = \frac{\text{Posttest Score} - \text{Pretest Score}}{\text{Maximum Score} - \text{Pretest Score}}$$

Table 8. Comparative Descriptive Matrix of Normalized Hake Gain Coefficients

Evaluated Cohort	N	Minimum Index	Maximum Index	Mean N-Gain Coefficient	Standard Deviation	Skewness Index	Kurtosis Index	Hake Categorical Classification
Control Group	22	0.20	0.80	0.455	0.171	0.166	-0.685	Medium Gain
Experimental Group	24	0.40	0.90	0.617	0.143	-0.029	-0.805	Medium-to-High

As detailed in Table 8, the experimental group achieved a mean N-Gain coefficient of 0.617 (61.7%). This places the cohort in the upper tier of the Medium classification, approaching the high-effectiveness threshold ($g \geq 0.70$). In contrast, the control group achieved a mean gain of 0.455 (45.5%). The net growth difference of 0.162 (16.2%) highlights the advantage of using a structured web-based instructional wrapper over traditional static media (Lestari & Kurniawan, 2023).

3.7 Qualitative Field Evidence: Interaction Transcripts, Student Artifact Diagnostics, and Observational Documentation

To better understand the quantitative findings, this sub-section presents qualitative field evidence, including structured lesson sequences, group discourse transcripts, and detailed assessments of student work.

3.7.1. Structural Visual Sequencing of Gagne's Framework Within Google Sites

The instructional environment was meticulously organized as a unified digital platform on Google Sites, where Gagne's Nine Events of Instruction provided the foundational pedagogical architecture for visual and content sequencing. Rather than delivering static multimedia content, the platform systematically aligned each instructional phase—ranging from gaining attention and informing learning objectives to presenting stimulus material and providing learning guidance—into distinct, intuitive visual modules. By embedding dynamic 3D Web-AR models directly into these structured phases, the platform created an explicit cognitive pathway that guided student engagement step-by-step. This deliberate structural sequencing minimized redundant digital navigation and reduced extraneous cognitive load, allowing learners to focus their mental resources entirely on complex scientific concepts. Each event on the Google Sites interface served as a contextual springboard, ensuring that students did not merely interact with augmented reality as a novel gimmick, but rather utilized the immersive visual models to fulfill specific learning outcomes. Consequently, the seamless integration of Gagne's framework within a web-based wrapper transformed abstract theoretical constructs into visually accessible, sequential learning steps, fostering a structured digital ecosystem that supported deep conceptual processing.

Furthermore, the operational architecture of this Gagne-structured platform provided a robust framework for self-regulated learning and digital scaffolding among diverse student cohorts. By establishing clear visual cues, progression indicators, and embedded interactive prompts within each Gagne event, the Google Sites environment effectively eliminated technical friction and spatial disorientation commonly associated with standalone virtual reality applications. Students were able to autonomously navigate through the instructional events at their own pace while maintaining alignment with the overall lesson trajectory. The visual layout systematically anchored 3D Web-AR orbital mechanics simulations within Event 4 (Presenting the Stimulus) and Event 5 (Providing Learning Guidance), ensuring that students received immediate, context-rich feedback prior to entering evaluative phases. This controlled digital sequence proved particularly beneficial for students with lower baseline competencies, as it offered a predictable, low-friction environment where complex spatial relationships could be revisited iteratively. As a result, the structured visual sequencing not only preserved cognitive capacity but also leveled the learning playing field, enabling students across all baseline tiers to achieve consistent performance and sustained conceptual mastery.

.

3.7.2. Field Discourse Transcript: Collaborative Modeling and Scaffolding

To evaluate how this digital architecture operated in real-time, qualitative field observations were conducted during Event 4 (Presenting the Stimulus) and Event 6 (Eliciting Performance) within a lesson dedicated to orbital mechanics and axial tilt. During these sessions, group discourse transcripts revealed how the interactive 3D Web-AR models functioned as powerful socio-constructivist scaffolds that enriched peer-to-peer communication. As the heterogeneous student group interacted with the manipulable Web-AR planetary model, the discourse shifted rapidly from passive observation to active scientific inquiry. Rather than relying solely on verbal explanations or two-dimensional textbook diagrams, the shared augmented reality space allowed students to co-observe axial inclination and orbital position simultaneously. This multi-user spatial visibility prompted immediate collaborative dialogue, where students actively pointed out spatial anomalies, debated cause-and-effect relationships regarding solar intensity, and collectively formulated hypotheses regarding seasonal variations.

The transcript analysis further highlighted the critical role of peer scaffolding and shared spatial artifacts in elevating scientific discourse among students with varying academic baselines. As Student A leveraged the 3D Web-AR model to articulate advanced spatial concepts, Student B and Student C were able to physically manipulate the augmented solar axis, verifying claims in real time. This interactive manipulation bridged the cognitive gap for Student C, who initially struggled with abstract rotation dynamics, by grounding spatial relationships in immediate visual evidence. The Web-AR model served as a neutral, empirical reference point that reduced ambiguity, empowering less confident students to articulate their reasoning using precise scientific terminology. The Teacher/Researcher (TR) strategically stepped back into a facilitative role, offering light promptings while allowing the students to drive the explanatory dialogue. Consequently, the combination of Gagne's deliberate event structuring and Web-AR spatial modeling generated a rich collaborative dynamic where academic hierarchies dissolved, replaced by inclusive, evidence-based discourse that directly supported the quantitative gains in scientific communication skills.

[Dialogue Transcription Segment]

TR : "Observe the digital projection on your viewports closely. What occurs when you manipulate the slider controlling the Earth's axial tilt during its orbital revolution around the sun?"
ST-A : "Look at the beam vector alignment on the northern hemisphere when the tilt tilts toward the sun. The light density is concentrated. This explains why the temperature scales rise during summer solstices!"
ST-B : "Wait, let me rotate the viewport angle. Ah, I see it now. The southern hemisphere is tilted away at the exact same point in the orbit, so the rays hit at a wider angle. That means it receives less energy per unit area."
ST-C : "Oh! So seasons are not caused by how close the Earth is to the sun? I thought it was about distance."
ST-A : "No, look at the Web-AR spatial distance indicator on your screen. The orbit is almost circular. The distance variation is minimal. The structural driver is clearly the constant 23.5-degree axial tilt combined with revolution!"
ST-C : "Let me measure the surface illumination cross-sections on the digital worksheet. Yes, the data matches perfectly. I can write down this evidence-based conclusion for our group report now."

This interaction demonstrates how the visual models served as a form of cognitive scaffolding (Mystakidis et al., 2021). Rather than simply memorizing text, students used the interactive visuals to identify spatial patterns and engage in collaborative reasoning (Chen et al., 2020; Yusuf & Widodo, 2023).

3.7.3. Student Artifact Diagnostic: Multi-Dimensional Answer Analysis

Evaluating student worksheets provided direct insight into how the digital intervention affected individual learning outcomes. Figure 4 contrasts a typical descriptive error from the control group with an evidence-based response from the experimental group.

- Control Group Response Analysis (Item 7 - Eclipses): The student provided a brief, text-only description: *"An eclipse happens when the moon covers the sun because they line up."* This response lacks specific geometric references or details about spatial positioning, showing a more superficial understanding of the underlying celestial mechanics (Sari et al., 2021).

- Experimental Group Response Analysis (Item 7 - Eclipses): The student drawn a precise diagram showing the umbra and penumbra zones, accompanied by a detailed explanation: *"Based on our observations of the 3D Web-AR simulation, a total solar eclipse occurs only when the moon is at its orbital node, casting its umbra onto a specific path on Earth. The small size of the umbra explains why total eclipses are visible only within a narrow geographical zone, while the penumbra creates a partial eclipse over a broader area."*

This comparison illustrates how the interactive platform helped students transition from generic descriptions to structured, evidence-based explanations, directly reflecting improvements in their broader scientific literacy (Cai et al., 2021; Rahmawati et al., 2021).

4. RESULTS AND DISCUSSION

The highly significant expansion of students' scientific literacy within the web-augmented reality (Web-AR) framework ($t = -15.461$; $p = 0.000$) uncovers a critical epistemological shift in how abstract visuospatial structures are mapped onto the human cognitive architecture. Traditional astronomy pedagogy routinely collapses because it demands that junior high school learners mentally simulate multidimensional, macro-spatial dynamics—such as heliocentric planetary revolutions and oblique multi-axial rotations—relying solely on flat, static textbooks. This archaic approach overwhelms the restricted capacity of the working memory's visuospatial sketchpad, inducing severe extraneous cognitive load that halts conceptual assimilation. By externalizing these complex spatial dimensions into interactive digital assets directly navigable on a browser viewer, the Web-AR ecosystem effectively takes over the computational burden of spatial visualization. Consequently, vital working memory resources are reclaimed from decoding static graphics and redirected toward deep germane processing and schema construction regarding physical laws. This empirical reality extends the foundational meta-analyses of Garzón & Acevedo (2021) and Garzón et al. (2020), which documented substantial effect sizes for AR in mitigating spatial misconceptions. However, this study expands upon those baseline models by proving that AR's spatial benefits are maximized only when liberated from hardware boundaries. Moving beyond isolated apps that trigger an "innovation-admiration trap," embedding the immersive elements within an agile web ecosystem bypasses hardware memory restrictions and immediate software barriers. Pedagogically, this transforms learners from passive consumers of information into active conceptual manipulators who control their own perceptual viewpoints, converting abstract mechanics into manageable, concrete visual truths.

The stark contrast in learning effectiveness between the experimental framework and the control group (N-Gain = 0.62 vs 0.45) cannot be credited merely to the presence of 3D models; rather, it underscores the systemic impact of the Google Sites platform operating as a highly structured instructional wrapper. In typical modern classrooms, standalone digital interventions frequently backfire because they lack systematic message design, leaving students highly engaged with the novel interface but educationally ungrounded. This investigation systematically resolves this issue by embedding the entire media workflow into a single web portal organized around Gagne's Nine Events of Instruction. This arrangement carefully shifts learner focus from an initial video prompt to clear target formulation, scaffolded exploration, digital performance, and final schema synthesis. Rather than distracting the mind, the technology functions as a rigid external guide for comprehension. This integration addresses the technical and instructional fragmentation risks highlighted by Pratama et al. (2022) and Lestari & Kurniawan (2023), confirming that Web-AR overcomes the device fragmentation that usually undermines classroom interventions. From a broader design perspective, this tightly controlled digital environment embodies the principle of intellectual discipline and systematic inquiry. By structuring clear pathways for knowledge acquisition, the digital wrapper ensures that immersive interactions remain purposeful and focused on academic growth. This structured harmony turns raw technological engagement into repeatable academic achievement, establishing a robust baseline for implementing advanced media in standard school environments without sacrificing instructional intent.

The robust improvement in scientific communication skills among students in the experimental group ($t = 5.520$; $p = 0.000$; mean difference = 9.24 points) validates the power of social constructivism when catalyzed by interactive digital reference points. Traditional junior high classrooms rarely cultivate genuine scientific dialogue because conventional media, such as standalone PowerPoint slides, reinforce an authoritative, teacher-centered lecture model. Conversely, the Web-AR platform operates as a shared, manipulable anchor that naturally provokes collaborative reasoning and group discourse. When analyzing a 3D model of planetary mechanics simultaneously, students move beyond passive listening; they actively exchange claims, defend predictions, and co-construct evidence-based conclusions. This social dynamic extends the findings of Chen et al. (2020) and Yusuf & Widodo (2023), proving that interactive multimedia accelerates verbal communication by supplying immediate, shared visual evidence. Furthermore, having clear visual data assists students in developing the habits of mind emphasized by Rahmawati et al. (2021) and Elkoumitti et al. (2025), where learners formulate arguments grounded in real-time observations. Rather than simply memorizing isolated labels, students learn to articulate and defend their scientific assertions using the live models on their screens. This interactive arrangement shifts classroom dynamics away from rote learning into an active forum for collaborative scientific inquiry, where precise academic language and dynamic imagery work together to deepen shared comprehension.

Despite the convincing empirical data, a critical look at the field realities exposes clear infrastructural anomalies, such as intermittent school internet connectivity and an evident initial learning curve for students unaccustomed to navigating 3D digital interfaces. These challenges illuminate the intersection of technical limits and uneven digital literacy, showing that even well-designed instructional systems must remain flexible to survive local constraints. To resolve these friction points, future rollouts must incorporate structured technical orientations prior to the core lessons and feature offline AR alternatives as a fallback for unstable network environments. Looking at the broader systemic impact, these findings offer a highly scalable, democratic alternative to expensive, standalone virtual reality setups, proving how everyday web tools can widen access to immersive science education. School districts can leverage this model to upgrade their curricula without requiring major hardware investments. Theoretically, this study challenges the common assumption that immersive media inherently causes cognitive overload, showing instead that the cognitive outcome depends entirely on the surrounding instructional architecture. By anchoring interactive technologies within structured, systematic lesson steps, educators can successfully guide student focus, lower extraneous cognitive load, and build essential scientific competencies for the modern digital era.

5. CONCLUSION

5.1. Suggestions

This study empirically validates that anchoring immersive Augmented Reality (AR) ecosystems within a web-based instructional wrapper designed via Gagne's Nine Events of Instruction significantly mitigates abstract spatial constraints and extraneous cognitive load while simultaneously boosting multidimensional science competencies in junior high school students. The systematic integration of responsive Web-AR assets directly into a cohesive Google Sites interface successfully bypassed traditional hardware performance boundaries and installation frictions, allowing working memory resources to be successfully redirected toward productive germane schema construction. Quantitatively, this instructional wrapper model generated an exceptionally high significant leap in scientific literacy ($t = -15.461$; $p = 0.000$; mean difference = 35.33 points), steering the experimental group to a superior posttest average (74.29) that far outdistanced the PowerPoint-integrated Project-Based Learning control framework (62.41). Furthermore, the platform operated as an effective shared visual anchor that naturally drove collaborative reasoning, culminating in significantly superior scientific communication skills ($t = 5.520$; $p = 0.000$; mean difference = 9.24 points). Ultimately, the Normalized Gain analysis firmly established the operational efficiency of this structured wrapper model (0.62; medium-approaching-high effectiveness) over the fragmented presentation model (0.45), confirming that tech-driven classroom interventions succeed not through raw exposure to visual novelties, but through the deliberate, systematic design of the surrounding instructional message architecture.

5.2. Suggestions

1. For Science Educators: Practitioners should actively transition away from isolated, standalone mobile applications toward lightweight, browser-based Web-AR systems embedded within cohesive digital portals like Google Sites. However, this adoption must not treat the technology as a passive repository of links; rather, it requires mapping out strict instructional events—such as combining interactive 3D models with guided prompts and immediate performance feedback—to ensure digital media directly serves pedagogical objectives.
2. For School Administrators: Institutional leaders must provide targeted support to bridge the gap between technological access and instructional design. Schools need to fund professional development programs that train educators in message design and cognitive load management, while concurrently upgrading local networking infrastructure to ensure fluid, lag-free delivery of interactive 3D content during live classroom sessions.
3. For Future Researchers: To expand on these findings, subsequent investigations should broaden the empirical scope by examining larger, more socio-economically diverse student populations over extended intervention periods. Crucially, future studies must explore deeper clinical methodologies, such as eye-tracking software and real-time cognitive diagnostic tools, to isolate exactly how immersive web environments affect high-order spatial thinking skills, long-term memory retention, and conceptual deconstruction across other abstract scientific disciplines.

REFERENCE

Abdullah, A. H., Hamdan, M. N., & Mokhtar, M. (2022). Efektivitas augmented reality dalam pembelajaran IPA: Kajian sistematis. *Jurnal Pendidikan Sains dan Matematika*, 10(1), 45–62.

Auralia, R., & Juliani, R. (2024). Pengaruh media augmented reality terhadap literasi sains siswa SMP pada materi ekosistem. *Jurnal Pendidikan Sains Indonesia*, 12(1), 15–27.

Borich, G. D. (2021). *Observation skills for effective teaching: Research-based practice* (8th ed.). Routledge.

Cahyadi, M. R., Maryanto, B. P. A., Syaifuddin, M., & Darmayanti, R. (2023). Development of Essay Test Assessment Rubric for Polya Theory-Based Mathematical Problem-Solving. *JNPM (Jurnal Nasional Pendidikan Matematika)*, 7(1), 45-58.

Cai, S., Liu, C., Wang, T., Liu, E., & Liang, J.-C. (2021). Effects of an augmented reality-based educational game on students' learning achievements and cognitive load in science. *Journal of Computer Assisted Learning*, 37(4), 1125–1136. *(Catatan: Terdapat juga variasi publikasi oleh penulis yang sama di tahun 2021 pada British Journal of Educational Technology)*.

Castro-Alonso, J. C., Wong, M., & Paas, F. (2021). Cognitive load theory and instructional design for digital learning. In *Cognitive Load Theory* (pp. 55-72). Springer. *(Catatan: Terdapat juga variasi artikel terkait di Educational Psychology Review)*.

Chen, C.-H., Yeh, H.-C., & Yu, S. (2020). Effects of integrating augmented reality into problem-based learning on student engagement and academic achievement. *Journal of Educational Technology & Society*, 23(4), 112-124. *(Catatan: Terdapat juga variasi artikel terkait di Interactive Learning Environments)*.

Creswell, J. W., & Creswell, J. D. (2020). *Research design: qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Durnalı, M., Orakcı, Ş., & Aktan, O. (2021). The smart school model. In *Technology-Enhanced Frameworks for School Improvement* (pp. 1-18). IGI Global.

Elkoumitti, Y., & Halidu, O. (2025). Digital learning ecosystems in secondary science streams: Scaffolding spatial models through web applications. *International Journal of Science Education*, 47(1), 112–131.

Field, A. (2020). *Discovering statistics using IBM SPSS statistics*. SAGE Publications. *(Disatukan: Terdapat variasi penulisan edisi ke-5 dan ke-6 pada sumber asli)*.

Garzón, J., & Acevedo, J. (2021). Meta-analysis of the effect of augmented reality on students' learning outcomes in STEM education. *Educational Research Review*, 31, 100373. *(Catatan: Terdapat juga variasi publikasi terkait di Journal of Educational Computing Research)*.

Garzón, J., Pavón, J., & Baldiris, S. (2020). Systematic review and meta-analysis of augmented reality in education. *Computers & Education*, 142, 103647. *(Catatan: Terdapat juga variasi publikasi terkait di Virtual Reality)*.

Gregory, R. J. (2021). *Psychological testing: History, principles, and applications* (8th ed.). Pearson.

Hake, R. R. (2020). *Analyzing change/gain scores*. American Educational Research Association.

Halidu, Z., Bakkali, S., & Talbi, M. (2025). Augmented reality integration in STEM education: Impact on student engagement and conceptual understanding. *Education and Information Technologies*, 30(2), 112–130.

Hidayat, T., Suwarna, I. P., & Nandiyanto, A. B. D. (2022). Analisis kemampuan literasi sains siswa SMP di Indonesia berdasarkan kerangka PISA 2018. *Jurnal Ilmu Pendidikan*, 28(1), 30–41.

Lestari, D. A., & Kurniawan, A. (2023). Google Sites sebagai platform media pembelajaran terintegrasi pada mata pelajaran IPA SMP. *Jurnal Teknologi Pendidikan*, 25(2), 88–101. *(Catatan: Terdapat variasi artikel terkait visual media pada volume yang sama)*.

Mishra, P., Pandey, C. M., Singh, U., & Gupta, A. (2021). Descriptive statistics and normality tests for statistical data. *Annals of Cardiac Anaesthesia*, 24(1), 67-72. *(Catatan: Terdapat variasi artikel terkait di Journal of Education and Health Promotion)*.

Mystakidis, S. (2022). Metaverse. *Encyclopedia*, 2(1), 486-497.

Mystakidis, S., Berki, E., & Valtanen, J.-P. (2021). Deep and meaningful e-learning with social virtual reality environments in higher education: A systematic literature review. *Applied Sciences*, 11(2), 694. *(Catatan: Terdapat variasi artikel serupa di Education Sciences dengan fokus social constructivism)*.

OECD. (2023). *PISA 2022 results: The state of learning and equity in education* (Vol. 1). OECD Publishing.

Pratama, H., Rosidin, U., & Suyatna, I. (2022). Visual tracking systems in science modules: Eliminating device memory constraints through cloud integration. *Journal of Science Education and Technology*, 31(3), 341–355.

Pratama, R. A., Agustini, K., & Sindu, I. G. P. (2022). Pengembangan media pembelajaran augmented reality pada materi tata surya untuk siswa SMP kelas VII. *Jurnal Pendidikan Teknologi dan Kejuruan*, 19(1), 64–75.

Putri, A. R., & Nugroho, A. S. (2022). Analisis miskonsepsi siswa SMP pada materi tata surya dan strategi remediasinya. *Jurnal Pendidikan IPA Indonesia*, 11(2), 201–213.

Radiusman. (2020). Studi literasi: Pemahaman konsep siswa pada pembelajaran matematika. *Fibonacci: Jurnal Pendidikan Matematika dan Matematika*, 6(1), 1–8.

Rahmawati, Y., Ridwan, A., & Hadinugrahaningsih, T. (2021). Mengembangkan keterampilan komunikasi ilmiah melalui pembelajaran berbasis transformative learning di SMA. *Jurnal Riset Pendidikan Kimia*, 11(2), 74–85. *(Disatukan: Terdapat variasi judul versi bahasa Inggris untuk dokumen yang sama)*.

Sari, I. N., Wardana, M. Y. S., & Prasetyo, Z. K. (2021). Identifikasi kesulitan belajar siswa pada materi astronomi di SMP: Studi kasus di Daerah Istimewa Yogyakarta. *Jurnal Inovasi Pendidikan IPA*, 7(2), 115–126. *(Disatukan: Terdapat variasi judul versi bahasa Inggris untuk dokumen yang sama)*.

Sörqvist, P., & Skog, E. (2026). Cognitive bias can depend on eco-label design: Evidence from a negative footprint illusion in energy labeled household appliances. *Energy Policy*, 216, 115391. *(Catatan: Terdapat variasi artikel terkait ergonomi lingkungan oleh penulis utama di jurnal yang sama)*.

Sugiyono. (2021). *Metode penelitian kuantitatif, kualitatif, dan R&D* (edisi ke-3). Alfabeta.

Suryani, A., Kurniawan, A., & Lestari, D. A. (2023). Integrasi media imersif dalam ekosistem pembelajaran digital. *Jurnal Teknologi Pendidikan*, 25(3), 180-195.

Sweller, J. (2020). Cognitive load theory: Recent theoretical advances. In *International Cognitive Load Theory Association Conference* (pp. 1-12). Springer.

Tetep, T., Suparman, S., & Yulianti, D. (2023). Augmented reality sebagai media inovatif dalam pembelajaran IPA: Kajian meta-analisis. *Jurnal Inovasi Teknologi Pendidikan*, 10(1), 33–45.

Yusuf, M., & Widodo, A. (2023). Pengembangan kemampuan komunikasi ilmiah siswa melalui pembelajaran berbasis proyek berbantuan teknologi digital. *Jurnal Pendidikan Sains*, 11(1), 44–56. *(Catatan: Terdapat variasi artikel terkait di Journal of Technology and Science Education)*