Kognisi: Jurnal Ilmu Keguruan
ISSN (online): 2987-0240
Volume 4 Issue 1 Year 2026
https://afeksi.id/journal3/index.php/kognisi/index

Augmented Reality Supported by Deep-Learning-Oriented Pedagogy for Solar System Learning: Development and Preliminary Evaluation of Junior High School Students' Creative Thinking

Bagas Brilian Ramadhan[1], Meida Wulan Sari[2], Sri Yamtinah[3], Bayu Antrakusuma[4]
Science Education Study Program, Universitas Sebelas Maret, Surakarta, Indonesia
Corresponding author: meidawulan@staff.uns.ac.id
DOI: https://doi.org/10.59698/kognisi.v4i1.625

Abstract
Creative thinking is increasingly recognized as a core learning outcome in science education, yet abstract astronomical concepts can be difficult for junior high school students to explore through conventional two-dimensional media. This study developed and preliminarily evaluated an augmented reality (AR) learning medium integrated with a deep-learning-oriented pedagogy for the Solar System topic. In this article, deep learning refers to an educational approach that emphasizes active knowledge construction, meaningful connections, reflection, and sustained cognitive engagement rather than artificial-intelligence deep learning. The study employed a Research and Development design using the ADDIE model. Two expert validators reviewed the product, and the implementation involved 32 Grade VII students selected through cluster random sampling from a population of 296 students at SMP Negeri 3 Karanganyar, Indonesia; a small-scale trial involved 10 students, and two science teachers provided practicality responses. Data were collected through observation, interviews, expert-validation questionnaires, student and teacher response questionnaires, and a 10-item pretest-posttest instrument adapted from the Torrance Test of Creative Thinking dimensions of fluency, flexibility, originality, and elaboration. Expert-validation coefficients were .89 for AR media, .82 for content and language, .87 for the deep-learning lesson plan, and .80 for test content. Practicality ratings reached 83% in the small-scale trial, 88% among students in the limited-scale implementation, and 91% among teachers. A paired-samples comparison indicated a significant pre-post difference ($p = .001$), while the mean normalized gain was .4460, representing a moderate improvement. Fluency showed the largest gain (.4683) and elaboration the smallest (.3495). The findings support the feasibility and educational promise of combining interactive AR visualization with cognitively engaging pedagogy. Because the implementation used a one-group pretest-posttest design without a control group, the results should be interpreted as preliminary evidence rather than a causal estimate of effectiveness.
Keywords: *Augmented Reality, Creative Thinking, Deep Learning Pedagogy, Science Education, Solar System, ADDIE*

INTRODUCTION

Creative thinking has become a central educational priority because contemporary learners are expected not only to reproduce established knowledge but also to generate, evaluate, refine, and communicate ideas in unfamiliar situations. In science education, this capacity is particularly important because scientific understanding develops through questioning, model building, explanation, evidence evaluation, and the revision of ideas. Divergent thinking is not identical to creativity, but fluency, flexibility, originality, and elaboration provide useful indicators of creative potential and can be cultivated through appropriately designed learning environments (Runco & Acar, 2012). Recent syntheses likewise suggest that digital technologies can support creativity when they are used to expand opportunities for exploration, production, collaboration, and iterative problem solving rather than functioning merely as presentation devices (Tang et al., 2022; Wang et al., 2024). Consequently, technology-enhanced science learning should be judged not only by whether it attracts attention but also by whether its design prompts students to think more actively and productively.

The need to strengthen creative thinking is especially relevant in Indonesia. Results from the Programme for International Student Assessment (PISA) 2022 creative-thinking assessment show that Indonesian 15-year-old students obtained an average score of 19 on a 60-point scale, compared with an OECD average of 33; only 31% of Indonesian students reached at least the baseline proficiency level (OECD, 2024). These data do not describe every Indonesian classroom, yet they highlight a systemic challenge: students need more frequent opportunities to generate varied ideas, examine alternatives, and improve their responses. Such opportunities are not automatically produced by adding digital media. The quality of learning depends on the relation among representations, tasks, teacher guidance, student activity, and assessment. Active-learning research demonstrates that learning improves when students are required to engage cognitively with content rather than remain passive recipients (Freeman et al., 2014), while the ICAP framework explains why constructive and interactive engagement generally elicit deeper learning than passive or merely active behaviors (Chi & Wylie, 2014).

The Solar System offers a compelling context for this pedagogical problem. Planetary size, distance, orbital motion, and the spatial relationships among celestial objects are difficult to experience directly at school. Textbooks, static images, and videos can provide important information, but they compress three-dimensional and dynamic phenomena into fixed or screen-bounded representations. Research in science education has therefore emphasized the value of representational tools that help learners coordinate abstract scientific models with perceptually accessible experiences. Augmented reality (AR) is relevant because it can overlay digital objects on the physical environment, allowing learners to inspect virtual three-dimensional representations while remaining situated in the real world (Azuma, 1997). In science learning, this combination may make invisible or inaccessible phenomena easier to explore, although the educational value of AR depends on task design and instructional support rather than on immersion alone (Cheng & Tsai, 2013; Radu, 2014).

A substantial research base now documents AR applications in education. Reviews have reported recurring benefits such as improved motivation, visualization, interaction, and learning achievement, alongside challenges involving usability, cognitive load, technical constraints, and inadequate pedagogical integration (Akçayır & Akçayır, 2017; Bacca et al., 2014; Wu et al., 2013). Early classroom and immersive-participatory studies similarly emphasized both the instructional potential and the orchestration demands of AR (Billinghurst & Duenser, 2012; Bower et al., 2014; Dunleavy et al., 2009), while mapping of science-education research shows sustained growth in this field (Arici et al., 2019). A systematic review focused on STEM learning concluded that AR is particularly useful for making complex spatial and scientific content more accessible, while also noting the need for stronger theory-driven designs and more rigorous evaluations (Ibáñez & Delgado-Kloos, 2018). Later meta-analyses found generally positive learning effects but also considerable variation across implementation conditions (Chang et al., 2022; Garzón & Acevedo, 2019; Garzón et al., 2019). Importantly, Garzón et al. (2020) showed that pedagogical approach moderates AR outcomes, reinforcing the argument that technological affordances and instructional strategy should be designed as an integrated system.

Evidence from individual science studies is consistent with this interpretation. AR-supported inquiry has been associated with stronger achievement and motivation in natural-science activities (Chiang et al., 2014), and AR integrated into problem-based physics learning has produced favorable achievement and attitude outcomes (Fidan & Tuncel, 2019). In biology, AR has likewise been linked to improved achievement and motivation (Erbas & Demirer, 2019). Studies that require learners to argue, explain, or solve problems with AR are especially relevant to higher-order thinking because they move the technology beyond three-dimensional display. For example, AR-based argumentation activities can support critical thinking and argumentation when students use visualized evidence as a resource for reasoning (Demircioglu et al., 2023). These findings suggest that an AR environment designed to foster creativity should make students do something intellectually consequential with the representation: compare, question, generate, justify, connect, and elaborate.

This study addresses that need through what the source project terms a Deep Learning approach. To avoid confusion with machine-learning terminology, the present article uses the phrase deep-learning-oriented pedagogy. The term refers here to an educational orientation toward meaningful understanding, mindful engagement, active knowledge construction, connections between prior and new knowledge, collaboration, and reflective application. This orientation is compatible with established learning-science accounts. Mayer (2002) distinguishes meaningful learning from rote retention by emphasizing knowledge that can be transferred and used, Hattie and Donoghue (2016) describe learning strategies that shift as learners move from acquiring surface knowledge toward deeper understanding and transfer, and the ICAP framework predicts stronger outcomes when learners construct and interact around knowledge (Chi & Wylie, 2014). Thus, the intended pedagogical function of AR in this study is not to produce novelty for its own sake but to create a manipulable representational space around which deeper cognitive activity can be organized.

AR may be particularly suited to creative-thinking tasks when it provides multiple viewpoints and manipulable objects. Spatially registered representations can help learners coordinate abstract and perceptual information (Bujak et al., 2013), and embodied mixed-reality interaction can support engagement when movement is aligned with the learning goal (Lindgren et al., 2016). However, visually rich technology does not guarantee learning; greater immersion can divert cognitive resources from essential processing (Makransky et al., 2019; Parong & Mayer, 2018). The current development therefore combined exploration with instructional prompts, navigation, explanatory panels, discussion, and reflection.

The connection between digital environments and creativity is promising but variable across tasks, duration, age groups, domains, and measurement approaches (Stolaki &

Economides, 2018; Tang et al., 2022; Wang et al., 2024). AR reviews indicate that visualization can facilitate exploration, but explicit creativity scaffolds remain necessary (Hidajat, 2024). Creative or narrative gains are more plausible when AR is embedded in structured generative activities rather than offered as unguided visualization (AlAli et al., 2025; Yilmaz & Goktas, 2017; Yousef, 2021). This literature supports the design hypothesis that AR is most likely to foster creative thinking when visual affordances are paired with cognitively generative tasks and pedagogical guidance.

The source study was motivated by classroom needs at SMP Negeri 3 Karanganyar. Observation and interviews indicated that Solar System instruction relied mainly on PowerPoint and conventional video and that students struggled to imagine inaccessible celestial objects. The development therefore created an Assemblr EDU-based AR medium organized from broad astronomical structures to specific Solar System objects and embedded it in independent preparation, classroom exploration, discussion, and reflection. The design addressed both representational difficulty and limited opportunities for creative engagement identified in the needs analysis.

Accordingly, this study had three objectives: (1) to develop an AR learning medium integrated with deep-learning-oriented pedagogy for junior high school Solar System learning; (2) to examine the content and media validity and the perceived practicality of the developed product; and (3) to examine preliminary changes in students' creative-thinking performance after implementation. The third objective is intentionally framed as preliminary evaluation. Because the implementation used a single group measured before and after instruction, it can establish whether scores changed during the intervention period but cannot isolate the AR-pedagogy package as the unique cause of that change. This distinction is important for publication-quality interpretation and informs both the analysis and the limitations discussed below.

METHODS

Research Design and Context

The study used a Research and Development (R&D) design organized through the ADDIE model: Analysis, Design, Development, Implementation, and Evaluation. ADDIE was selected because it provides a systematic sequence for linking needs, instructional design decisions, product construction, field use, and revision while allowing iterative refinement. The development was conducted for Grade VII science learning on the Solar System at SMP Negeri 3 Karanganyar, Indonesia. The product was designed in Assemblr EDU and was intended for use with mobile-device-based three-dimensional and augmented-reality interaction. The study therefore combined a product-development component with a limited field implementation to examine validity, practicality, and preliminary learning outcomes.

Participants and Sampling

Two experts participated in the validation stage: one lecturer from the Science Education Study Program at Universitas Sebelas Maret and one science teacher. The implementation stage involved 32 students from class VII-H. The class was selected using cluster random sampling from a population of 296 Grade VII students distributed across nine classes. Before the limited-scale implementation, a small-scale trial was conducted with 10 students to identify usability and response issues. Two science teachers also completed the teacher-response questionnaire during the limited-scale stage. The source report does not provide additional demographic characteristics such as student age distribution or gender; therefore, those characteristics are not inferred in this article.

ADDIE Development Procedures

During the Analysis phase, classroom observations and interviews with teachers and students were used to identify learning problems and media needs. The analysis indicated that Solar System learning was dominated by PowerPoint and conventional video and that students had difficulty visualizing astronomical objects. During Design, the researchers prepared a flowchart and storyboard, planned three user interaction pathways (view in 3D, Scan Marker, and Place in Your Room) and organized content from broad astronomical contexts such as the Milky Way and the Solar System to specific celestial objects. A deep-learning lesson plan (Rencana Pembelajaran Mendalam, RPM) was prepared in parallel so that the technological environment and the learning sequence were designed together rather than separately.

During Development, the media was constructed in Assemblr EDU under the title "Astronomy & Space Exploration." The source manuscript reports an access address of https://uns.id/AR-SitemTataSurya. The product included eight planets and other celestial objects such as asteroids, comets, meteoroids, and stars. Expert feedback led to several revisions, including an instruction panel on the main menu, a navigation button returning users to the main menu, and descriptive/data panels for individual Solar System objects. These revisions were educationally significant because they reduced navigation ambiguity and strengthened the relationship between visual exploration and conceptual information.

During Implementation, students first explored the AR material independently before the face-to-face meeting. Classroom learning then continued with guided exploration and discussion. Creative thinking was measured before and after implementation. Student and teacher response questionnaires were administered after use. The Evaluation phase synthesized evidence from expert judgments, user responses, and pretest-posttest results to judge whether the product was sufficiently valid and practical for further development and whether student creative-thinking scores showed a meaningful pattern of improvement.

Pedagogical Alignment of the Product

The learning design treated AR as a representational tool within a broader sequence of cognitively engaging activities. Three principles guided the alignment. First, meaningful representation was supported by presenting otherwise inaccessible astronomical objects as manipulable three-dimensional models. Second, active exploration was supported through multiple interaction pathways that allowed students to inspect objects rather than only watch a linear presentation. Third, classroom discussion and pre-class exploration were used to connect individual interaction with explanation and social knowledge construction. These principles are consistent with evidence that active and constructive engagement supports learning (Chi & Wylie, 2014; Freeman et al., 2014) and with AR research emphasizing the importance of pedagogical orchestration (Garzón et al., 2020). The alignment was also intended to create opportunities for the four creative-thinking dimensions measured in the study: generating multiple ideas (fluency), considering different perspectives or categories (flexibility), producing uncommon responses (originality), and developing ideas in detail (elaboration).

Instruments and Data Collection

Five sources of evidence were used. First, observation and interview protocols supported needs analysis. Second, expert-validation questionnaires used a four-point Likert format to judge the AR medium, content and language, the deep-learning lesson plan, and the content of the creative-thinking test. The media instrument comprised 18 items, the content and language instrument 28 items, the lesson-plan instrument 16 items, and the test-content review covered 10 items. Third, student response questionnaires were administered in the small-scale and limited-scale trials, and a teacher-response questionnaire was administered to two science teachers. Fourth, the creative-thinking pretest and posttest consisted of 10 items adapted to represent four dimensions commonly associated with the Torrance tradition: fluency, flexibility, originality, and elaboration. TTCT-related measures have a long history in creativity research, but psychometric interpretation depends on scoring procedures, task form, context, and reliability evidence (Kim, 2006; Said-Metwaly et al., 2017). Fifth, the development process itself generated qualitative revision information from expert comments and field observations.

Data Analysis

Content-validity judgments were summarized using the Gregory Matrix coefficient as specified in the source research protocol. Coefficients at or above .80 were interpreted by the study as very high. Practicality responses were converted into percentages from the four-point Likert scale, with values in the 81–100% range categorized as very feasible. For the creative-thinking instrument, item-level construct evidence was examined through Pearson correlations using SPSS 27. With $n = 32$, the source study used an r-table criterion of .349 at the 5% significance level. Pretest item correlations ranged from .388 to .678, and posttest item correlations ranged from .413 to .840; thus, all 10 items exceeded the stated criterion in both administrations.

The pretest and posttest distributions were examined using the Shapiro-Wilk test, an established test of normality for complete samples (Shapiro & Wilk, 1965). The reported significance values were .050 for the pretest and .071 for the posttest. Because the pretest value lies exactly at the conventional .05 threshold, normality should be described as borderline rather than unequivocally established. The source study proceeded with a paired-samples t-test to evaluate the pre-post difference and reported $p = .001$. It did not report the pretest and posttest means and standard deviations, the t statistic, degrees of freedom, confidence interval, or a standardized within-subject effect size; these quantities are therefore not reconstructed or invented here.

Normalized gain (N-Gain) was used to summarize improvement relative to the available score range. The source study interpreted values from .30 to below .70 as moderate gain.

Overall N-Gain and indicator-specific N-Gain values are reported exactly as provided in the original dataset summary. The quantitative analysis is interpreted descriptively and inferentially within the limits of the design. Specifically, a significant paired test indicates that scores differed between measurement occasions; it does not by itself demonstrate that the intervention caused the difference. No between-group comparison was available.

Scope of Inference

The implementation is best characterized as a one-group pretest-posttest preliminary evaluation embedded within an R&D study. This design is appropriate for identifying whether a developed product is usable and whether learning outcomes move in the expected direction during an early field trial, but it is vulnerable to alternative explanations such as repeated testing, maturation, novelty, teacher effects, concurrent learning experiences, and regression processes. The conclusions therefore focus on product validity, user-reported practicality, and observed pre-post improvement. Strong causal language such as "proven effective" is avoided. The acknowledgments in the source report indicate that the school granted permission and that teachers and students participated in the data-collection process; no additional institutional ethics-review information is reported in the source and none is assumed here.

Table 1. Design alignment between AR features, learning activity, and intended creative-thinking affordances

AR / learning-design feature	Pedagogical function	Potential creative-thinking affordance
View in 3D / markerless placement	Makes inaccessible celestial objects manipulable and spatially inspectable.	Provides multiple visual cues for fluency and originality.
Content from broad to specific scales	Supports connection making from galaxy/Solar System context to individual objects.	Encourages perspective shifts relevant to flexibility.
Pre-class independent exploration	Provides prior exposure and student control before guided classroom activity.	Creates time to notice, question, and generate initial ideas.
Classroom exploration and discussion	Moves activity from viewing toward constructive and interactive engagement.	Supports comparison, idea revision, and verbal elaboration.
Descriptions, instructions, and navigation	Reduces interface ambiguity and links visualization to scientific information.	Preserves cognitive resources for concept-focused creative tasks.

Kognisi: Jurnal Ilmu Keguruan
ISSN (online): 2987-0240
Volume 4 Issue 1 Year 2026
https://afeksi.id/journal3/index.php/kognisi/index

RESULTS

Developed AR Learning Medium

The ADDIE process produced an AR learning medium titled "Astronomy & Space Exploration" on the Assemblr EDU platform. The final product presented the Solar System and related astronomical content through interactive three-dimensional objects. Students could choose among View in 3D, Scan Marker, and Place in Your Room. The third pathway enabled markerless placement of digital objects in the user's surrounding environment. Content was structured from broader to more specific scales so that learners could move from the Milky Way and Solar System overview to individual planets and smaller celestial objects. Expert feedback resulted in improved instructions, return navigation, and more informative description panels. These changes made the product more self-explanatory and reduced the likelihood that interaction would become disconnected from scientific content.

Expert Validity

All four validation components reached the study's very-high category. The AR media component obtained a coefficient of .89, content and language .82, the deep-learning lesson plan .87, and the content validity of the 10-item pretest-posttest instrument .80. The pattern indicates that validators judged the product, learning plan, and assessment content as sufficiently aligned for limited field use. The result should be understood as expert-judgment evidence rather than a comprehensive psychometric validation of every instrument.

Practicality Responses

Student and teacher responses were consistently favorable. In the small-scale trial, 10 students produced an overall response percentage of 83%. During the limited-scale implementation, 32 students produced an 88% response, while two science teachers produced a 91% response. Under the study's predefined criteria, all three percentages were categorized as very feasible. The increase from the small-scale to the limited-scale student response is also consistent with the iterative revisions made after early evaluation, although the study did not test whether the difference between these percentages was statistically meaningful.

Preliminary Creative-Thinking Outcomes

Item correlations exceeded the stated r-table criterion in both administrations: .388–.678 for the pretest and .413–.840 for the posttest. Shapiro-Wilk significance values were .050 and .071, respectively. The paired-samples t-test reported $p = .001$, indicating a statistically detectable difference between the pretest and posttest measurements. Because the original report did not provide descriptive score statistics or the t statistic, the magnitude of the raw-score change cannot be independently reconstructed from the available manuscript. The normalized-gain analysis nevertheless provides a standardized descriptive indication of improvement: across 32 students, N-Gain ranged from .10 to .80 and averaged .4460, which the study classified as moderate.

All four creative-thinking indicators were also in the moderate-gain category. Fluency had the largest mean N-Gain (.4683), followed by originality (.4567) and flexibility (.4452). Elaboration had the smallest mean gain (.3495). The result suggests that the learning sequence was most strongly associated with students' capacity to produce multiple responses and somewhat less strongly with the ability to extend and detail ideas. This pattern is pedagogically plausible given that three-dimensional exploration readily supplies multiple visual prompts, while elaboration typically requires sustained explanation, justification, revision, and time-on-task.

Table 2. Expert-validation results

Component	Focus / items	Coefficient	Category
AR media	Graphics and AR media, 18 items	.89	Very high
Content and language	Content, presentation, language, contextuality, 28 items	.82	Very high
Deep-learning lesson plan (RPM)	Lesson-plan feasibility, 16 items	.87	Very high
Test content	Pretest-posttest content, 10 items	.80	Very high

Table 3. Student and teacher practicality responses

Trial stage / respondent	N	Response	Category
Small-scale / students	10	83%	Very feasible
Limited-scale / students	32	88%	Very feasible
Limited-scale / science teachers	2	91%	Very feasible

Table 4. Summary of creative-thinking measurement and pre-post evidence

Analysis	Reported result	Interpretation in this study
Pretest item correlations	.388–.678; criterion r = .349	All reported items exceeded the study criterion.
Posttest item correlations	.413–.840; criterion r = .349	All reported items exceeded the study criterion.
Shapiro-Wilk	Pretest p = .050; posttest p = .071	Pretest normality is borderline at α = .05; posttest is above .05.
Paired-samples t-test	p = .001	Pretest and posttest scores differed statistically; causal attribution is not established.
Overall N-Gain	Mean = .4460; range = .10–.80	Moderate gain under the study criterion.

Table 5. N-Gain by creative-thinking indicator

Indicator	Minimum	Maximum	Mean N-Gain	Category
Fluency	.40	.53	.4683	Moderate
Flexibility	.34	.55	.4452	Moderate
Originality	.36	.60	.4567	Moderate
Elaboration	.33	.37	.3495	Moderate

DISCUSSION

Validity, Usability, and Pedagogical Coherence

The first contribution of the study is the development of a technically interactive AR product that was explicitly linked to a pedagogical sequence. The high expert-validation coefficients and strong user-response percentages indicate that the product was acceptable for limited classroom use. This matters because AR research has repeatedly shown that usability and instructional coherence are conditions for educational value. Akçayır and Akçayır (2017) identified interaction, visualization, and motivation as frequent advantages of AR, but also reported technical and usability barriers. Wu et al. (2013) similarly emphasized that AR must be integrated with appropriate learning tasks, while Ibáñez and Delgado-Kloos (2018) argued that STEM applications benefit most when visualization is connected to inquiry and conceptual learning rather than treated as a stand-alone novelty. The revisions made in this development—clearer instructions, return navigation, and richer descriptive panels—directly address these concerns.

The practicality findings are also consistent with studies in which AR has been successfully used as an accessible support for science learning. Akçayır et al. (2016), for example, found that AR in science laboratories could support students' laboratory skills and attitudes, and Chiang et al. (2014) reported positive achievement and motivation outcomes in mobile AR-supported science inquiry. Fidan and Tuncel (2019) demonstrated that embedding AR within problem-based physics learning can strengthen both cognitive and affective outcomes. These studies differ from the present research in age, content, setting, and experimental design, but collectively they reinforce a central design lesson: positive experiences tend to emerge when the technology is tied to a coherent learning activity. In the current product, the sequence from independent exploration to classroom discussion served that coordinating function.

At the same time, favorable response percentages should not be confused with evidence of learning. Novel visual technologies often generate enthusiasm, and perceived ease or enjoyment may be important for adoption without guaranteeing conceptual change. Research comparing immersive technologies shows that stronger presence can sometimes reduce learning if extraneous processing becomes dominant (Makransky et al., 2019; Parong & Mayer, 2018). The present development partially mitigated this risk by incorporating descriptive information and a guided learning plan, yet future studies should directly measure cognitive load, navigation difficulty, and task engagement. A publication-quality evaluation of AR should therefore treat usability, affect, conceptual understanding, and higher-order thinking as related but distinct outcomes.

Interpreting the Pre-Post Improvement

The significant paired comparison and moderate mean N-Gain provide encouraging preliminary evidence that students' creative-thinking performance improved during the instructional period. The result fits the wider literature suggesting that AR can support learning by making complex relationships inspectable and by encouraging active manipulation. Meta-analytic evidence indicates positive average effects of AR on learning, although effect sizes vary substantially according to subject area, pedagogical approach, duration, and comparison condition (Chang et al., 2022; Garzón & Acevedo, 2019; Garzón et al., 2019). The current study's moderate gain should therefore be regarded as plausible and educationally meaningful without being overstated. It demonstrates a direction of change that justifies more rigorous testing, not definitive causal proof.

The interpretation is strengthened conceptually by the pedagogical design. The media did not simply display planets; it asked students to explore representations before class and then participate in exploration and discussion during class. This sequence can shift learners from passive viewing toward constructive and interactive engagement. According to the ICAP framework, generating inferences, explanations, or new representations is cognitively more productive than passively receiving information, and interactive co-construction can be more productive still (Chi & Wylie, 2014). Similarly, active-learning evidence across STEM disciplines supports the value of requiring students to act on knowledge rather than merely hear it (Freeman et al., 2014). The current deep-learning-oriented pedagogy is therefore best understood not as a separate technology but as the instructional mechanism that gives the AR representation a role in meaning making.

This interpretation also aligns with research showing that the pedagogical configuration of AR matters. Garzón et al. (2020) found that instructional approaches moderated the educational impact of AR, and Cheng and Tsai (2013) argued that science learning with AR requires careful attention to representation, inquiry, and learner guidance. Bujak et al. (2013) further proposed that AR can reduce the gap between abstract representations and concrete experience by spatially coupling information, while Lindgren et al. (2016) demonstrated that embodied interaction can be productive when bodily activity is meaningfully connected to concepts. For Solar System learning, where direct observation of scale and spatial relations is inherently constrained, these affordances provide a strong rationale for three-dimensional exploration.

Why Fluency Improved More Than Elaboration

The indicator pattern offers a more nuanced view than the overall gain alone. Fluency showed the largest improvement, followed closely by originality and flexibility, whereas elaboration was lower. One explanation is that interactive AR supplies a dense stream of perceptual cues. Students can change viewpoint, inspect multiple objects, and encounter visually differentiated planetary features; such cues may trigger more associations and make it easier to generate several responses. This is compatible with divergent-thinking theory, in which fluency reflects the capacity to produce multiple ideas and originality reflects uncommon responses (Runco & Acar, 2012). Digital-creativity research likewise suggests that technology can expand the representational resources available for ideation when learners have agency to manipulate and transform information (Tang et al., 2022).

Elaboration, by contrast, requires students to develop an idea with additional detail, explanation, or refinement. A visually impressive object does not automatically create the discourse time needed for such refinement. The lower elaboration gain may therefore reveal a design boundary rather than a weakness of AR itself. Reviews of creativity assessment caution that different creativity dimensions respond differently to tasks and scoring procedures, and that short interventions may favor idea generation over sustained development (Said-Metwaly et al., 2017). The current finding also resonates with Hidajat's (2024) review of AR for mathematical creativity, which emphasizes the importance of scaffolds that explicitly support reasoning and multiple-stage idea development. Future versions could include prompts such as "extend your explanation," "compare two plausible models," "add evidence," "revise the idea after peer feedback," or "design an alternative representation."

Evidence from related AR creativity studies supports this strategy. Yilmaz and Goktas (2017) showed that AR can contribute to narrative and creative outcomes when embedded in structured storytelling, while Yousef (2021) reported creativity-related benefits in primary learning when AR was integrated into guided activities. More recent work with gifted eighth-grade students found that AR-supported science instruction can enhance creative-thinking skills under an explicit instructional design (AlAli et al., 2025). Although these populations and interventions are not directly comparable with the present study, they suggest that creativity gains depend on what students are asked to produce with the technology. In other words, AR may provide the stimulus, but elaboration requires purposeful prompts, criteria, feedback, and sufficient time.

Deep Learning as a Pedagogical, Not Technological, Construct

A second conceptual contribution is clarification of the term deep learning. In current educational and technological discourse, "deep learning" can refer either to neural-network methods in artificial intelligence or to learning characterized by meaningful understanding and transfer. The present study concerns the latter. The learning design sought to move students beyond recognition of planetary facts toward active exploration and connection making. Mayer (2002) describes meaningful learning as learning that supports transfer, while Hattie and Donoghue (2016) distinguish strategies that are more useful during surface acquisition from those supporting deeper consolidation and transfer. The current sequence—independent preparation, AR exploration, discussion, and reflection—can be interpreted as an attempt to create conditions for that progression.

This clarification matters because the visible technology should not be mistaken for the learning mechanism. Enjoyment and curiosity may support engagement, but productive learning also requires perceived control, task value, cognitive demand, and opportunities for agency (Pekrun, 2006). Learning-science research similarly emphasizes meaningful challenge and opportunities to apply knowledge (Darling-Hammond et al., 2020). Accordingly, the deep-learning orientation is most defensible when operationalized through observable student activities rather than retained as a broad label.

Implications for Science Teaching

For science teachers, the findings suggest three practical implications. First, AR should be selected for topics where spatial, dynamic, or inaccessible phenomena create a genuine representational problem. The Solar System meets this criterion because students cannot physically observe the target objects at classroom scale. Second, AR tasks should be designed around questions and products rather than around viewing alone. Students might compare planetary conditions, construct evidence-based explanations, identify multiple possible interpretations of visual data, or redesign a representation for a younger audience. Third, creative-thinking indicators can be used as design targets. A fluency prompt asks for many possibilities; a flexibility prompt requires multiple categories or perspectives; an originality prompt asks for uncommon but defensible ideas; and an elaboration prompt requires added detail and justification.

These recommendations are consistent with broader evidence that technology has stronger educational impact when it supports active, constructive, and collaborative processes (Chi & Wylie, 2014; Freeman et al., 2014). They also respond to an important caution in AR research: students can become occupied with interface actions while failing to process the intended concept. Teacher facilitation should therefore periodically redirect attention from "what the object looks like" to "what the representation means," "what evidence it provides," and "what new question it raises." Such facilitation would help the product function as a tool for thinking rather than as a digital exhibit.

Limitations and Directions for Further Research

Several limitations constrain interpretation. Most importantly, the study did not include a control or comparison group. The significant pre-post difference therefore cannot distinguish intervention effects from maturation, repeated-testing effects, novelty, teacher influence, or other concurrent experiences. A future quasi-experimental or randomized study should compare the AR plus deep-learning-oriented pedagogy with a well-specified alternative, such as the same pedagogy using non-AR three-dimensional media or conventional multimedia. Such a design would help separate the contribution of the pedagogical sequence from the incremental contribution of AR itself. Meta-analytic literature has shown that comparison conditions strongly influence estimated technology effects, so this distinction is essential for a stronger causal claim (Chang et al., 2022; Garzón et al., 2020).

Second, the available report does not include pretest and posttest means and standard deviations, the paired t statistic and degrees of freedom, confidence intervals, or a standardized effect size. These should be added from the raw data because p values alone do not convey magnitude or precision. Third, the 32-student sample limits generalization; replication across schools, teachers, and devices is needed. Fourth, self-reported practicality responses may reflect novelty or social desirability and should be complemented by observation or usage evidence.

Fifth, the creative-thinking instrument was described as an adaptation of TTCT dimensions, but the report provides limited psychometric information beyond item correlations and expert content judgments. Creativity measurement is methodologically complex, and validity should be understood as an accumulation of evidence about score interpretation rather than a single item-correlation threshold (Kim, 2006; Said-Metwaly et al., 2017). Future work should report internal consistency where appropriate, scoring rubrics, assessor training, and inter-rater reliability if responses require human judgment. Factor-analytic evidence may also be considered with a sufficiently large sample. Finally, the pretest Shapiro-Wilk p value of .050 is borderline; with raw paired differences available, future analyses should inspect the distribution of difference scores directly and could report a robust or nonparametric sensitivity analysis when assumptions are questionable.

These limitations do not negate the development results. Rather, they define the next research phase. The product has passed initial expert and user evaluation, and students' creative-thinking scores moved in the expected direction with a moderate normalized gain. The appropriate next step is controlled, multi-class evaluation with fuller statistical reporting and a stronger measurement framework. Such research could also examine conceptual understanding, cognitive load, motivation, retention, and transfer alongside creative thinking, providing a more complete account of how AR and pedagogy interact.

CONCLUSION

This R&D study developed an Assemblr EDU-based AR medium for Solar System learning and integrated it with a deep-learning-oriented pedagogical sequence emphasizing meaningful, active, and reflective engagement. Expert judgments indicated very-high validity for the AR media (.89), content and language (.82), the lesson plan (.87), and test content (.80). Student and teacher responses also indicated strong perceived practicality, with percentages of 83% in the small-scale student trial, 88% among students in the limited-scale implementation, and 91% among teachers.

During implementation with 32 Grade VII students, creative-thinking scores differed significantly between pretest and posttest ($p = .001$), and the mean normalized gain was .4460, categorized as moderate. Fluency showed the largest indicator gain (.4683), while elaboration showed the smallest (.3495). These patterns suggest that interactive three-dimensional exploration may readily stimulate idea generation, whereas detailed development of ideas requires more explicit scaffolding and sustained discourse.

The study supports the feasibility and educational promise of integrating AR with pedagogically demanding science activities, but the evidence should be interpreted as preliminary. Because no control group was included, the design cannot attribute the observed improvement uniquely to the intervention. Future research should use comparison groups, larger and more diverse samples, fuller reporting of effect magnitude and uncertainty, and stronger psychometric evidence for creative-thinking measures. For classroom practice, the central implication is that AR should not be treated as an end in itself: its value lies in how three-dimensional representations are connected to questions, explanations, collaboration, reflection, and creative production.

Acknowledgments

The authors thank the principal and science teachers of SMP Negeri 3 Karanganyar for permission and collaboration during data collection and all students of class VII-H who participated in the study.

REFERENCES

Akçayır, M., & Akçayır, G. (2017). Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review, 20*, 1–11. https://doi.org/10.1016/j.edurev.2016.11.002

Akçayır, M., Akçayır, G., Pektaş, H. M., & Ocak, M. A. (2016). Augmented reality in science laboratories: The effects of augmented reality on university students' laboratory skills and attitudes toward science laboratories. *Computers in Human Behavior, 57*, 334–342. https://doi.org/10.1016/j.chb.2015.12.054

AlAli, R., Wardat, Y., Zaki Aboud, Y., & Alhayek, K. A. (2025). The effectiveness of using augmented reality technology in science education to enhance creative thinking skills among gifted eighth-grade students. *Eurasia Journal of Mathematics, Science and Technology Education, 21*(6), em2644. https://doi.org/10.29333/ejmste/16416

Arici, F., Yildirim, P., Caliklar, Ş., & Yilmaz, R. M. (2019). Research trends in the use of augmented reality in science education: Content and bibliometric mapping analysis. *Computers & Education, 142*, 103647. https://doi.org/10.1016/j.compedu.2019.103647

Azuma, R. T. (1997). A survey of augmented reality. *Presence: Teleoperators and Virtual Environments, 6*(4), 355–385. https://doi.org/10.1162/pres.1997.6.4.355

Bacca, J., Baldiris, S., Fabregat, R., Graf, S., & Kinshuk. (2014). Augmented reality trends in education: A systematic review of research and applications. *Educational Technology & Society, 17*(4), 133–149. https://www.jstor.org/stable/jeductechsoci.17.4.133

Billinghurst, M., & Duenser, A. (2012). Augmented reality in the classroom. *Computer, 45*(7), 56–63. https://doi.org/10.1109/MC.2012.111

Bower, M., Howe, C., McCredie, N., Robinson, A., & Grover, D. (2014). Augmented reality in education—Cases, places and potentials. *Educational Media International, 51*(1), 1–15. https://doi.org/10.1080/09523987.2014.889400

Bujak, K. R., Radu, I., Catrambone, R., MacIntyre, B., Zheng, R., & Golubski, G. (2013). A psychological perspective on augmented reality in the mathematics classroom. *Computers & Education, 68*, 536–544. https://doi.org/10.1016/j.compedu.2013.02.017

Chang, H.-Y., Binali, T., Liang, J.-C., Chiou, G.-L., Cheng, K.-H., Lee, S. W.-Y., & Tsai, C.-C. (2022). Ten years of augmented reality in education: A meta-analysis of (quasi-)experimental studies to investigate the impact. *Computers & Education, 191*, 104641. https://doi.org/10.1016/j.compedu.2022.104641

Cheng, K.-H., & Tsai, C.-C. (2013). Affordances of augmented reality in science learning: Suggestions for future research. *Journal of Science Education and Technology, 22*(4), 449–462. https://doi.org/10.1007/s10956-012-9405-9

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Chiang, T. H. C., Yang, S. J. H., & Hwang, G.-J. (2014). An augmented reality-based mobile learning system to improve students' learning achievements and motivations in natural science inquiry activities. *Educational Technology & Society, 17*(4), 352–365. https://www.jstor.org/stable/jeductechsoci.17.4.352

Darling-Hammond, L., Flook, L., Cook-Harvey, C., Barron, B., & Osher, D. (2020). Implications for educational practice of the science of learning and development. *Applied Developmental Science, 24*(2), 97–140. https://doi.org/10.1080/10888691.2018.1537791

Demircioglu, T., Karakus, M., & Ucar, S. (2023). Developing students' critical thinking skills and argumentation abilities through augmented reality–based argumentation activities in science classes. *Science & Education, 32*, 1165–1195. https://doi.org/10.1007/s11191-022-00369-5

Dunleavy, M., Dede, C., & Mitchell, R. (2009). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of Science Education and Technology, 18*(1), 7–22. https://doi.org/10.1007/s10956-008-9119-1

Erbas, C., & Demirer, V. (2019). The effects of augmented reality on students' academic achievement and motivation in a biology course. *Journal of Computer Assisted Learning, 35*(3), 450–458. https://doi.org/10.1111/jcal.12350

Fidan, M., & Tuncel, M. (2019). Integrating augmented reality into problem based learning: The effects on learning achievement and attitude in physics education. *Computers & Education, 142*, 103635. https://doi.org/10.1016/j.compedu.2019.103635

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences, 111*(23), 8410–8415. https://doi.org/10.1073/pnas.1319030111

Garzón, J., & Acevedo, J. (2019). Meta-analysis of the impact of augmented reality on students' learning gains. *Educational Research Review, 27*, 244–260. https://doi.org/10.1016/j.edurev.2019.04.001

Garzón, J., Kinshuk, Baldiris, S., Gutiérrez, J., & Pavón, J. (2020). How do pedagogical approaches affect the impact of augmented reality on education? A meta-analysis and research synthesis. *Educational Research Review, 31*, 100334. https://doi.org/10.1016/j.edurev.2020.100334

Garzón, J., Pavón, J., & Baldiris, S. (2019). Systematic review and meta-analysis of augmented reality in educational settings. *Virtual Reality, 23*(4), 447–459. https://doi.org/10.1007/s10055-019-00379-9

Hattie, J. A. C., & Donoghue, G. M. (2016). Learning strategies: A synthesis and conceptual model. *npj Science of Learning, 1*, 16013. https://doi.org/10.1038/npjscilearn.2016.13

Hidajat, F. A. (2024). Augmented reality applications for mathematical creativity: A systematic review. *Journal of Computers in Education, 11*, 991–1040. https://doi.org/10.1007/s40692-023-00287-7

Ibáñez, M.-B., & Delgado-Kloos, C. (2018). Augmented reality for STEM learning: A systematic review. *Computers & Education, 123*, 109–123. https://doi.org/10.1016/j.compedu.2018.05.002

Kim, K. H. (2006). Can we trust creativity tests? A review of the Torrance Tests of Creative Thinking (TTCT). *Creativity Research Journal, 18*(1), 3–14. https://doi.org/10.1207/s15326934crj1801_2

Lindgren, R., Tscholl, M., Wang, S., & Johnson, E. (2016). Enhancing learning and engagement through embodied interaction within a mixed reality simulation. *Computers & Education, 95*, 174–187. https://doi.org/10.1016/j.compedu.2016.01.001

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and Instruction, 60*, 225–236. https://doi.org/10.1016/j.learninstruc.2017.12.007

Mayer, R. E. (2002). Rote versus meaningful learning. *Theory Into Practice, 41*(4), 226–232. https://doi.org/10.1207/s15430421tip4104_4

OECD. (2024). *PISA 2022 Results (Volume III): Creative minds, creative schools.* OECD Publishing. https://doi.org/10.1787/765ee8c2-en

Parong, J., & Mayer, R. E. (2018). Learning science in immersive virtual reality. *Journal of Educational Psychology, 110*(6), 785–797. https://doi.org/10.1037/edu0000241

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Radu, I. (2014). Augmented reality in education: A meta-review and cross-media analysis. *Personal and Ubiquitous Computing, 18*(6), 1533–1543. https://doi.org/10.1007/s00779-013-0747-y

Runco, M. A., & Acar, S. (2012). Divergent thinking as an indicator of creative potential. *Creativity Research Journal, 24*(1), 66–75. https://doi.org/10.1080/10400419.2012.652929

Said-Metwaly, S., Van den Noortgate, W., & Kyndt, E. (2017). Methodological issues in measuring creativity: A systematic literature review. *Creativity. Theories – Research – Applications, 4*(2), 276–301. https://doi.org/10.1515/ctra-2017-0014

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika, 52*(3–4), 591–611. https://doi.org/10.1093/biomet/52.3-4.591

Stolaki, A., & Economides, A. A. (2018). The Creativity Challenge Game: An educational intervention for creativity enhancement with the integration of Information and Communication Technologies (ICTs). *Computers & Education, 123*, 195–211. https://doi.org/10.1016/j.compedu.2018.05.009

Tang, C., Mao, S., Naumann, S. E., & Xing, Z. (2022). Improving student creativity through digital technology products: A literature review. *Thinking Skills and Creativity, 44*, 101032. https://doi.org/10.1016/j.tsc.2022.101032

Wang, Y., Liu, W., Yu, X., Li, B., & Wang, Q. (2024). The impact of virtual technology on students' creativity: A meta-analysis. *Computers & Education, 215*, 105044. https://doi.org/10.1016/j.compedu.2024.105044

Wu, H.-K., Lee, S. W.-Y., Chang, H.-Y., & Liang, J.-C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education, 62*, 41–49. https://doi.org/10.1016/j.compedu.2012.10.024

Yilmaz, R. M., & Goktas, Y. (2017). Using augmented reality technology in storytelling activities: Examining elementary students' narrative skill and creativity. *Virtual Reality, 21*(2), 75–89. https://doi.org/10.1007/s10055-016-0300-1

Yousef, A. M. F. (2021). Augmented reality assisted learning achievement, motivation, and creativity for children of low-grade in primary school. *Journal of Computer Assisted Learning, 37*(4), 966–977. https://doi.org/10.1111/jcal.12536