

Google Earth-Supported Landform Learning: Product Development and Initial Evidence of Students' Spatial-Thinking Gains

Wa Ode Samlia¹, Sunarty Suly Eraku¹, Nurfaika Nurfaika¹

¹Master's Program in Geography Education, Universitas Negeri Gorontalo, Gorontalo, Indonesia

ARTICLE INFO	ABSTRACT
Article History: Received: 2026-08-28 Accepted: 2026-09-28 Published: 2026-09-30 Keywords: Geography education; Google Earth; Landforms; Spatial thinking; Virtual globe Corresponding author: Sunarty Suly Eraku Email: sunarty.eraku@ung.ac.id DOI: 10.37905/jgej.v7i2.41158 Copyright © 2026 The Authors  This open access article is distributed under a Creative Commons Attribution-NonCommercial (CC-BY-NC) 4.0 International License	Spatial thinking is a core competence in geography learning, yet landform topics are often taught through verbal explanation, textbooks, two-dimensional images, and static maps that provide limited opportunities for students to explore relief, scale, distance, and spatial relationships. This study developed and initially evaluated a Google Earth-supported learning environment for landform learning using the Analysis, Design, Development, Implementation, and Evaluation (ADDIE) model. The product integrated Google Earth projects, QR-based access, 2D/3D visualization, zoom and measurement functions, location marking, and guided spatial-thinking activities. A single expert appraised the material, media, and evaluation/instrument aspects. The product obtained an overall feasibility score of 85.38%, with the media aspect receiving the highest score (91.67%), followed by evaluation (86.67%) and material (78.75%). Implementation involved 30 student respondents at SMA Negeri 2 Konawe Selatan. Mean scores increased from 74.53 on the pretest to 86.50 on the posttest. All respondents obtained higher posttest scores, and the mean normalized gain was 0.488, indicating a moderate increase; 10.0% of students were in the high-gain category, 86.7% in the moderate category, and 3.3% in the low category. These results indicate that the developed learning environment is feasible and can support spatial exploration and learning gains. Because the implementation used a one-group pretest–posttest design, the score increase should be interpreted as initial evidence rather than a causal treatment effect.

How to cite: Samlia, W. O., Eraku, S. S., & Nurfaika. (2026). Google Earth-Supported Landform Learning: Product Development and Initial Evidence of Students' Spatial-Thinking Gains. *Jambura Geo Education Journal*, 7(2), 359–371. <https://doi.org/10.37905/jgej.v7i2.41158>

1. Introduction

Spatial thinking is a disciplinary foundation of geography because geographic understanding requires learners to locate phenomena, work with distance and scale, interpret representations, recognize spatial patterns, and reason about relationships among places and processes. It is therefore not a single visualization skill, but a coordination of spatial concepts, representational tools, and reasoning processes. Ishikawa (2016) empirically distinguishes geospatial conceptions, map-based reasoning, and spatial visualization as related but non-identical dimensions. At the school level, Thayaseelan et al. (2024) further show that spatial-thinking measures require construct validation and attention to item invariance. Meta-analytic evidence also indicates that spatial skills are malleable, can improve through training, and can transfer beyond the specific task practiced (Uttal et al., 2013). Developing these dimensions is especially important in secondary geography, where students must move beyond naming geographic features toward explaining how and why they are spatially organized.

Landform learning makes this representational demand particularly visible. To understand mountains, valleys, highlands, lowlands, slopes, and drainage settings, students must coordinate plan-view maps, satellite images, elevation information, profiles, and three-dimensional terrain while also relating visible form to geomorphological processes. Static pictures and verbal descriptions can introduce terminology, but they provide limited opportunities to rotate terrain, change scale, compare viewpoints, or measure relationships directly. Importantly, 3D visualization alone does not guarantee improvement across all spatial abilities. In a Google Earth study, students improved in achievement and mental rotation but not in spatial visualization, while cognitive load remained high (Koç & Topu, 2022). This finding underlines the need for carefully sequenced tasks rather than technology-led demonstrations.

Interactive geospatial environments nevertheless provide strong affordances for addressing this problem. Google Earth enables learners to search for real locations, move between local and regional scales, tilt and rotate terrain, compare 2D and 3D views, measure distance, and examine associations

among physical and human features. A quasi-experimental study in secondary geography found that Google Earth supported students' analysis of spatial and temporal change more effectively than conventional instruction (Xiang & Liu, 2017). A later review similarly identified data visualization and flexible reasoning about spatial patterns and relations as plausible mechanisms through which virtual globes can support geospatial learning (Jaeger, 2024). These studies establish the educational potential of Google Earth, but they also imply that learning depends on what students are asked to do with its functions.

Adjacent research on augmented and immersive learning clarifies both the promise and the risk of interactive visualization. Reviews identify benefits for representation, engagement, and learning, but also recurrent problems involving usability, distraction, technical instability, and weak pedagogical integration (Akçayır & Akçayır, 2017; Wu et al., 2013). Meta-analytic evidence suggests a positive average association between AR use and learning gains, while also showing that effects vary by learner, subject, and implementation context (Garzón & Acevedo, 2019). More recent syntheses demonstrate that cognitive-load outcomes remain inconsistent and depend on the relationship among content complexity, interface demands, learner expertise, and instructional support (Buchner et al., 2022; Poupard et al., 2025). Experimental evidence that stronger immersion can increase presence while reducing learning further cautions against treating visual intensity as instructional quality (Makransky et al., 2019). Although the present product is a mobile 3D virtual globe configuration rather than a camera-overlay AR, this evidence is relevant to its design because both require deliberate alignment between visual affordances and learning actions. A recent geography-specific synthesis similarly emphasizes that spatial-learning benefits depend on balancing interactivity and spatial anchoring with instructional scaffolding (Pratama et al., 2025).

Classroom integration is therefore a design problem rather than a device-selection problem. Work in Computers & Education emphasizes that educational AR must accommodate classroom constraints, teacher orchestration, collaboration, and the visibility of learner activity (Cuendet et al., 2013). A recent design-based study in the British Journal of Educational Technology likewise found that coherent tasks, progressive challenge, adequate exploration time, clear representations, and useful feedback distinguish productive AR-supported learning environments from technically impressive but instructionally weak ones (Li et al., 2025). Evidence comparing marker-based and markerless AR further shows that different technical modes elicit different engagement and navigation patterns (Na et al., 2025). These findings support a design chain in which the selected technology, spatial action, disciplinary concept, and assessment evidence are specified together.

This need was sharpened by Schmidt and Stumpe's (2025) systematic review of 18 mobile AR studies in primary and secondary geography. Fourteen studies did not refer to a learning theory, 12 did not specify a pedagogical approach, and only three explicitly employed 3D manipulation. The evidence was also dominated by small, selected groups and short interventions, while higher-order thinking, validated competence measures, and learning-process variables remained insufficiently examined. Thus, the central literature gap is not simply a shortage of digital or immersive geography applications. It is the limited instructional alignment among technology-specific affordances, geographic content, theoretically purposeful learning activities, and competence-specific assessment.

Taken together, the Google Earth and adjacent immersive-learning literature leaves a narrower development gap for school-based landform learning. The literature has not adequately documented a traceable design chain in which each virtual-globe operation is assigned to a specific spatial-thinking indicator, enacted through a guided landform task, and evaluated with geospatial rather than predominantly verbal stimuli. This gap was also evident in the study context. The needs analysis at SMA Negeri 2 Konawe Selatan showed that instruction relied mainly on teacher explanation, textbooks, two-dimensional pictures, and static maps, providing limited opportunities for students to observe relief from multiple perspectives, measure spatial relationships, or connect landforms with real geographic settings.

Accordingly, this study aimed to develop and conduct an initial evaluation of a Google Earth-supported landform learning environment using the Analysis, Design, Development, Implementation, and Evaluation (ADDIE) model. The product combined QR-based access, Google Earth's 2D/3D visualization and measurement functions, guided spatial tasks, a user guide, and geospatial assessment stimuli. Two questions guided the study: (1) To what extent is the developed product feasible based on a single-expert appraisal of its material, media, and evaluation/instrument aspects? and (2) What pattern of change occurs in students' spatial-thinking scores after implementation, as indicated by

pretest–posttest scores and normalized gain? The study addresses the instructional-alignment component of the current research agenda by making the relationship among geospatial functions, learning actions, spatial-thinking indicators, and assessment explicit. Because the implementation used a one-group design, it offers product-development and initial score-change evidence rather than a causal estimate of effectiveness.

2. Method

2.1 Research Design and Study Context

This study employed an educational research and development (R&D) design structured by the five ADDIE phases: Analyze, Design, Develop, Implement, and Evaluate. ADDIE was used as a development framework because it connects needs analysis, specification of instructional tasks, production of learning resources, classroom enactment, and iterative evaluation (Branch, 2009). Consistent with educational design research, the inquiry treated the learning product and evidence from its initial use as mutually informative for revision (McKenney & Reeves, 2018).

This study was conducted at SMA Negeri 2 Konawe Selatan. The analytic sample comprised 30 students with complete pretest and posttest scores. The classroom implementation followed a one-group pretest–posttest arrangement in a regular geography-learning context, without a comparison group or random assignment. Accordingly, the study was designed to report product feasibility and initial within-group score change, not to estimate a causal treatment effect (Shadish et al., 2002).

2.2 ADDIE Development Procedure

In the Analysis phase, the researchers examined the learning needs, landform content, learner and device characteristics, existing instructional practices, and spatial thinking demands of the topic. The researchers consulted the geography teacher to clarify classroom conditions and the difficulties students encountered when interpreting relief, scale, direction, and relationships among landform features. The analysis specified five target domains: (1) location, direction, distance, and scale; (2) transformation between two- and three-dimensional representations; (3) interpretation of maps, imagery, and three-dimensional terrain; (4) recognition of spatial patterns, distributions, associations, and relations; and (5) reasoning about landforms, geomorphic processes, and environmental context.

In the Design phase, the five domains were translated into observable learning actions and media features. The design included Google Earth project navigation; selected representations of Mount Mekongga, Lohe Valley, highlands, and lowlands; transitions between 2D and 3D views; zoom, measurement, and place-marking functions; QR-based access; worksheets; guided questions; group discussion; product-review instruments; and a spatial-thinking test. The intended sequence moved from locating and orienting a feature to changing viewpoint, measuring and comparing terrain, interpreting spatial relations, and explaining the associated geomorphic process. Table 1 presents the alignment between the design components and the intended spatial-thinking evidence.

Table 1. Alignment of Google Earth design components with spatial-thinking indicators

Design component	Instructional function	Spatial-thinking indicator
Location search and orientation	Identify the position of a landform and its regional context.	Location, relative position, direction/orientation
Zoom, scale, and measurement tools	Compare dimensions and measure distance or spatial extent.	Distance and scale
Switching between 2D and 3D views	Compare representations and understand relief.	2D-3D transformation, relief
Observation of rivers, roads, vegetation, and settlements	Identify relationships among features and patterns of land use.	Pattern, association, spatial relationships
Geomorphological process questions	Explain relationships among landform, process, and environment.	Spatial reasoning

Source: Research data, 2026

During the Development phase, the design specifications were converted into a working Google Earth learning project and a user guide. The prototype integrated QR access, 2D/3D visualization, zoom and measurement tools, location marking, landform-learning activities, and operating instructions. The researchers checked access and navigation, submitted the product and instruments for expert appraisal, and revised the materials in response to the review, including the depth of geomorphological explanations and the clarity of spatial-thinking prompts.

The Implementation phase placed the revised product in classroom use and generated paired pretest and posttest observations. The Evaluation phase operated throughout the cycle: formative evidence from the initial analysis, functional checking, and expert comments informed revision, whereas the final product-review percentages and student score summaries provided an initial summative description. This staged but iterative use of ADDIE follows the framework's emphasis on evaluation and revision across the development process (Branch, 2009).

2.3 Product Review and Research Instruments

The product-review package comprised three four-point rating forms completed by one reviewer: 20 items for material quality, 18 items for media quality, and 15 items for evaluation/instrument quality. The forms addressed the relevance and organization of the landform content, the usability and technical presentation of the Google Earth medium, and the suitability of the assessment components. Written comments were used as formative evidence for revision. A related ADDIE study of AR media for lithospheric-dynamics learning used separate material, media, and learning-expert validation before testing the product with students (Bokingo et al., 2023).

Because the appraisal was conducted by one reviewer, the resulting percentages were treated as a single-expert formative appraisal rather than as an inter-rater estimate of content validity. This distinction is important because the interpretation of content validity evidence depends on transparent reporting of the number of experts, rating procedure, and aggregation method (Polit & Beck, 2006).

Student outcomes were assessed using a spatial-thinking test administered before and after use of the product. The test blueprint represented the five domains defined during Analysis and used landform-related maps, imagery, 2D/3D representations, measurement, relief profiles, and alternative viewpoints. The blueprint drew on Ishikawa's (2016) account of geospatial conceptions and map-based reasoning. Its multidomain structure also responds to Thayaseelan et al.'s (2024) evidence that school-level spatial measurement requires explicit construct validation and attention to item behavior. Scores were reported on a 0–100 scale.

2.4 Implementation and Data Collection

Students first completed the spatial-thinking pretest. They then used the developed medium in guided geography-learning activities that required them to access the Google Earth project, locate selected landforms, change between 2D and 3D views, zoom and measure terrain, mark locations, compare representations, discuss spatial patterns, and connect visible features with geomorphological processes. A posttest was administered after the learning activities. The analysis used the 30 cases for which both pretest and posttest scores were complete.

2.5 Data Analysis

For each product-review component, the obtained item scores were summed and divided by the maximum possible score. The overall product percentage was calculated from the sum of all obtained scores divided by the sum of all maximum scores:

$$\text{Feasibility (\%)} = (\text{obtained score} / \text{maximum possible score}) \times 100$$

Pretest and posttest distributions were summarized using the mean, standard deviation, minimum, and maximum. Raw score change was retained alongside normalized gain so that the observed change could be interpreted in relation to the baseline distribution. Individual normalized gain was calculated before averaging across students and groups (Hake, 1998):

$$g_i = (\text{posttest}_i - \text{pretest}_i) / (100 - \text{pretest}_i)$$

Individual values were classified as low ($g < 0.30$), moderate ($0.30 \leq g < 0.70$), or high ($g \geq 0.70$), following Hake (1998). The mean individual N-gain was reported for the full sample. For the eight classroom groups, N-gain was calculated from each group's mean pretest and posttest scores. Because normalized gain can vary with the pretest distribution, it was interpreted together with the raw pretest and posttest statistics rather than as a standalone effect-size estimate (Nissen et al., 2018).

All analyses were descriptive. No inferential significance test, covariate adjustment, or between-group treatment comparison was performed. Therefore, the score changes provide preliminary implementation evidence for this sample and should not be interpreted as a causal effect of the learning medium (Shadish et al., 2002).

3. Results and Discussion

3.1 Results

3.1.1 Analysis and Design Outputs

The Analysis stage identified a mismatch between the representational demands of landform learning and the classroom resources used to support those demands. Students were expected to interpret relief, scale, direction, distribution, and relations among landforms, whereas static representations provided limited opportunities to rotate terrain, change viewpoints, compare two- and three-dimensional representations, or measure spatial relations directly. The analysis therefore specified five target domains: positional reasoning; transformation between 2D and 3D representations; interpretation of maps, imagery, and terrain; recognition of spatial patterns and relations; and explanation of landforms in relation to geomorphic processes and environmental context.

The Design stage translated those domains into a task sequence and product specification. Search and orientation functions supported location and direction tasks; zoom and measurement supported distance, scale, and terrain comparison; changes between 2D and 3D views supported representational transformation; place marks and selected landform examples supported recognition of patterns and associations; and worksheets, prompts, and group discussions linked the visible terrain to geomorphic explanations. The resulting prototype combined a Google Earth project, smartphone access through a QR code, operating guidance, learning activities, and a spatial-thinking assessment. Figure 1 shows the developed 3D view and access configuration.



Figure 1. Example of the developed learning product: a 3D Google Earth view on a smartphone and QR-based access to the project.

3.1.2 Product Appraisal and Revision

One expert appraised the prototype using three four-point rating forms. The material component obtained 63 of 80 possible points (78.75%), the media component obtained 66 of 72 points (91.67%),

and the evaluation/instrument component obtained 52 of 60 points (86.67%). Across the 53 items, the total was 181 of 212 points (85.38%). Media received the highest percentage, followed by evaluation/instrument and material. Because all ratings came from one reviewer, the percentages in Table 2 represent a single-expert product appraisal rather than agreement across experts.

Table 2. Single-expert product appraisal results

Aspect	Items	Score	Maximum	Percentage	Decision
Material	20	63	80	78.75%	Feasible with revision
Media	18	66	72	91.67%	Highly feasible; minor revision
Evaluation/instrument	15	52	60	86.67%	Feasible with revision
Overall	53	181	212	85.38%	Feasible after revision

Source: Research data, 2026

The appraisal generated two principal revision priorities. First, the landform material was expanded so that the visible terrain was connected more explicitly to geomorphological processes and environmental context. Second, the spatial-thinking prompts and evaluation items were clarified so that students were required to locate, measure, compare viewpoints, identify spatial relations, and explain the observed forms rather than only identify features. These changes addressed the comparatively lower material rating and strengthened the alignment between content, learning actions, and assessment.

The revised package retained the positively appraised technical features: QR-based access, smartphone compatibility, 2D/3D navigation, zoom and measurement, place marking, and a guided sequence for classroom use. Functional checking was completed after revision, and the revised version was used in the classroom implementation.

3.1.3 Pretest–Posttest Score Patterns

Complete paired scores were available for 30 students. The mean increased from 74.53 (SD = 5.82; range = 60–83) on the pretest to 86.50 (SD = 5.75; range = 76–98) on the posttest, a raw mean increase of 11.97 points. Every student obtained a higher posttest score. The mean of the 30 individual N-gain values was 0.488 (SD = 0.155; range = 0.273–0.889). Based on the stated thresholds, one student was in the low category, 26 were in the moderate category, and three were in the high category. Table 3 presents the full-sample descriptive statistics.

Table 3. Descriptive statistics for pretest, posttest, and N-gain (n = 30)

Indicator	Mean	SD	Minimum	Maximum
Pretest	74.53	5.82	60.00	83.00
Posttest	86.50	5.75	76.00	98.00
N-gain	0.488	0.155	0.273	0.889

Source: Research data, 2026

Table 4. Mean pretest, posttest, score increase, and N-gain by group

Group	n	Pretest	Posttest	Increase	N-gain
1	4	77.50	87.25	9.75	0.434 (Moderate)
2	4	80.25	94.00	13.75	0.705 (High)
3	4	79.00	90.25	11.25	0.554 (Moderate)
4	4	69.75	85.50	15.75	0.522 (Moderate)
5	4	74.00	86.25	12.25	0.486 (Moderate)
6	3	72.67	84.67	12.00	0.443 (Moderate)
7	4	70.00	81.25	11.25	0.381 (Moderate)

Group	n	Pretest	Posttest	Increase	N-gain
8	3	72.00	81.00	9.00	0.324 (Moderate)

Source: Research data, 2026

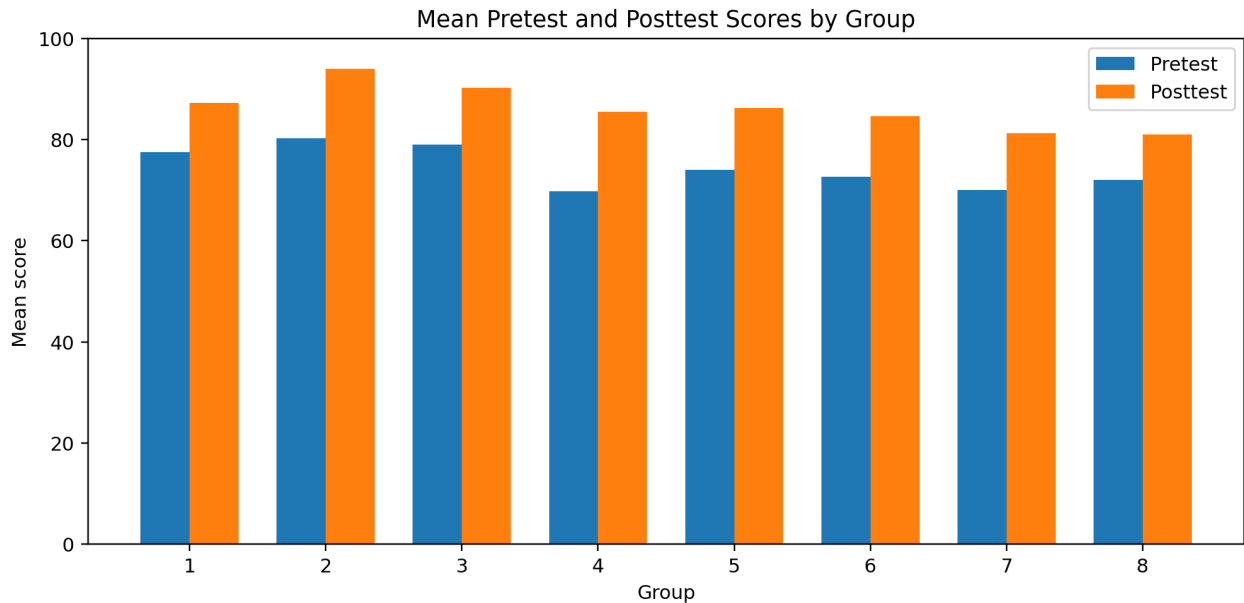


Figure 2. Mean pretest and posttest scores by group.

At the group level, N-gain calculated from each group's mean pretest and posttest scores ranged from 0.324 to 0.705 (Table 4 and Figures 2–3). Group 2 had the highest mean posttest score (94.00) and the highest group-level N-gain (0.705), whereas Group 8 had the lowest group-level N-gain (0.324). Seven group-level values fell in the moderate range, and one reached the high threshold. The raw mean increases ranged from 9.00 points in Group 8 to 15.75 points in Group 4. The Evaluation stage consolidated the revised product, expert-appraisal record, classroom materials, and paired-score summaries. These outputs provide a traceable account of product readiness and of score patterns observed during the initial implementation, illustrated in Figure 3. The Discussion interprets these results.

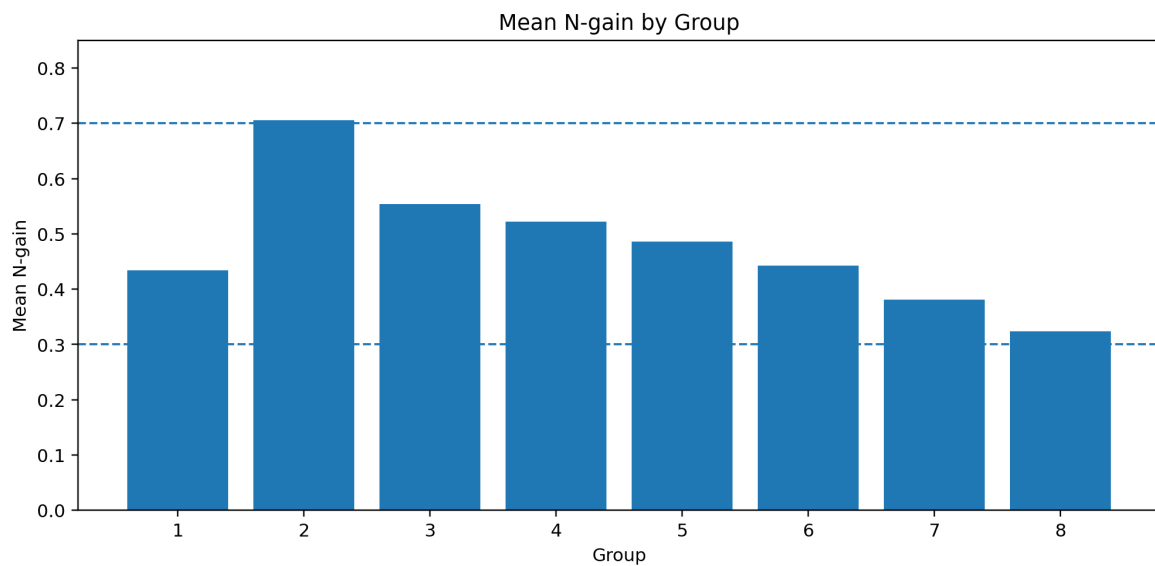


Figure 3. Mean N-gain by group; dashed lines indicate the moderate (0.30) and high (0.70) thresholds.

3.2 Discussion

3.2.1 Product Feasibility, Instructional Alignment, and Technical Readiness

The appraisal profile indicates that technical presentation and usability were the most positively judged aspects of the product. The media score (91.67%) exceeded the evaluation/instrument score (86.67%) and the material score (78.75%), while the combined score was 85.38%. This pattern supports readiness for initial classroom use, but it should be read as formative evidence from one expert, not as a stable estimate of inter-rater content validity. Transparent separation of reviewer count, rating dimensions, and aggregation is necessary when content-validity claims are made (Polit & Beck, 2006).

The strong media rating is educationally meaningful only because the technical functions were embedded in an instructional sequence. Research on classroom technology design argues that visibility of learner activity, teacher orchestration, collaboration, and fit with ordinary classroom constraints should be treated as design requirements rather than implementation details (Cuendet et al., 2013; Wu et al., 2013). Recent design-based evidence similarly identifies coherent tasks, progressive challenge, sufficient exploration time, clear representations, and useful feedback as conditions for productive AR-supported learning (Li et al., 2025). In the present product, the QR code shortened the access path, but the intended learning value lay in using search, viewpoint change, zoom, measurement, and place marking to generate spatial evidence. The product's design logic is consistent with recent geospatial and immersive-learning work that treats disciplinary task design, classroom framing, and orchestration as conditions for productive use rather than automatic consequences of the technology (Bolkas et al., 2024; Edler et al., 2025; Jaeger, 2024; Schmidt & Stumpe, 2025; Spangenberg et al., 2025).

The comparatively lower material rating identifies an important boundary: visually accessible terrain does not by itself provide a sufficiently developed explanation of landform processes. Research on spatial thinking shows that position, scale, representation, pattern, and causal interpretation must be coordinated across maps and digital geographic information (Duarte et al., 2022; Ishikawa, 2016). Recent intervention and measurement studies further show that conclusions depend on the spatial subskill assessed and on how learners interact with the representation (Na et al., 2025; Thayaseelan et al., 2024; Wang et al., 2025). Expanding the geomorphological explanations and rewriting prompts as observable spatial actions therefore strengthened the substantive layer of the product, not merely its wording.

A second design boundary concerns cognitive demand. Interactive 3D and augmented-reality environments may help learners externalize rotation, perspective change, and spatial comparison; however, controls, navigation, split attention, and unfamiliar interfaces can also add extraneous load (Elford et al., 2022; Jian & Abu Bakar, 2024; Koumpouros, 2024; Mokmin et al., 2024; Şimşek et al., 2025). Systematic reviews report that the direction of cognitive load effects varies across learners and designs (Buchner et al., 2022; Poupard et al., 2025), while experimental work shows that stronger presence does not necessarily produce stronger learning (Makransky et al., 2019). Even pre-training may not remove these demands uniformly, and visuospatial working-memory differences can remain consequential (Arztmann et al., 2025). The guided sequence and operating instructions in the present product are compatible with recommendations to manage these demands. However, cognitive load was not measured; therefore, reduced load cannot be claimed from the appraisal or score data.

3.2.2 Score Patterns as Initial Spatial-Learning Evidence

The implementation produced a clear descriptive pattern: the mean score rose by 11.97 points, all 30 students improved, and the mean individual N-gain was 0.488. The concentration of 26 students in the moderate N-gain category, with 3 high and 1 low, indicates that improvement was widespread within this sample. Nevertheless, the one-group pretest–posttest arrangement cannot separate the contribution of the product from testing, maturation, concurrent instruction, teacher support, or other time-varying influences. The findings should therefore be described as initial within-group implementation evidence rather than a causal effect (Shadish et al., 2002).

The N-gain result also requires calibration. Normalized gain expresses improvement relative to the remaining distance from a maximum score and is useful for describing change across different baselines (Hake, 1998). Yet it can be associated with the pretest distribution and should not replace raw change, uncertainty estimates, or a design-based effect estimate (Nissen et al., 2018). Reporting the 11.97-point raw increase alongside the individual N-gain distribution is therefore more informative than treating

0.488 as an isolated measure of effectiveness. The group-level values require the same caution because they were calculated from group means and involved only three or four students per group.

Several mechanisms nevertheless make the observed pattern educationally plausible. Google Earth allows repeated changes of scale and viewpoint, movement between regional context and local form, and observation of places that may be inaccessible during ordinary lessons. Google Earth research links these affordances to geospatial reasoning and attitudes when they are embedded in guided tasks (Jaeger, 2024; Ortiz-Guerrero & Loizzo, 2024). Complementary virtual-field literature documents disciplinary engagement and improved access, while emphasizing the need for instructional framing and support (Cook et al., 2025; Semken et al., 2025). The present sequence used these affordances to connect location, measurement, terrain comparison, and geomorphic explanation, offering students external representations on which to base discussion.

The 2D/3D transition is particularly relevant because landform interpretation requires learners to coordinate map symbols, imagery, profiles, and perspective views. The broader evidence establishes that spatial skills can respond to training (Uttal et al., 2013) and that AR interventions have produced positive average learning gains (Garzón & Acevedo, 2019), but neither finding warrants a uniform-effect assumption. Research with three-dimensional geospatial technology documents embodiment and representational advantages alongside uneven or construct-specific cognitive outcomes (Bagher et al., 2023; Koç & Topu, 2022). For example, an AR-integrated mathematics curriculum did not improve every broad spatial outcome, although a movement-related dimension improved (Wang et al., 2025), and a recent scoping review found that effects vary with learner characteristics, task design, spatial subskill, and assessment choice (Yang et al., 2026). The overall score increase in this study is therefore compatible with, but does not demonstrate, change in each spatial domain.

Variation among the eight groups is descriptive rather than evidence of differential treatment response. Group 2 reached a group-level N-gain of 0.705, whereas Group 8 obtained 0.324, but the groups were small, were not randomized conditions, and had different starting scores. The variation may reflect baseline composition, interaction quality, navigation fluency, teacher attention, or other unmeasured processes. Process evidence such as screen recordings, interaction logs, observations, and group discourse would be required to connect learning actions to score changes.

This cautious interpretation is reinforced by recent methodological reviews. Lawson et al. (2024) found that many media-comparison studies involving immersive virtual reality for STEM education did not fully control instructional and technological differences, making it difficult to attribute outcomes to immersion itself. Reviews in geography education likewise call for larger samples, longer interventions, validated measures, and attention to process variables (Schmidt & Stumpe, 2025). The present evidence is therefore most useful for refining the product and specifying a stronger subsequent evaluation, rather than for ranking the learning environment against conventional instruction.

3.2.3 Practical Implications, Terminological Boundary, and Limitations

For classroom use, the results support a structured sequence: establish location and orientation; compare 2D and 3D representations; manipulate scale and viewpoint; measure and mark features; identify spatial patterns; and conclude with an explanation of the relevant geomorphic process. The teacher remains central in modelling controls, focusing attention, probing spatial language, and linking observations to concepts. Classroom-design research supports making learner activity visible, preparing students for the interaction mechanics, and coordinating technical access with guided tasks and debriefing (Arztmann et al., 2025; Cuendet et al., 2013; Li et al., 2025). This implication also aligns with recent geography and immersive-learning studies that present disciplinary framing, teacher guidance, and orchestration as necessary design considerations (Bolkas et al., 2024; Edler et al., 2025; Schmidt & Stumpe, 2025; Spangenberg et al., 2025). Local development studies provide contextual support for organizing interactive geography media around explicit curriculum objectives (Dumako et al., 2025; Meeradji et al., 2024).

Terminological precision is required. The documented product combines Google Earth 3D visualization, mobile access, a QR link, and guided learning tasks; the manuscript does not report camera-based superimposition of virtual objects on the learner's immediate physical surroundings. The revised title therefore foregrounds Google Earth rather than AR. Reviews distinguish AR from immersive virtual reality, desktop or mobile 3D environments, and other extended-reality configurations because each entails different affordances, interaction requirements, and potential

confound (Poupard et al., 2025; Schmidt & Stumpe, 2025; Wu et al., 2013). Future reporting should specify the hardware, software mode, tracking or registration features, and learner interaction pathway before assigning a technology label.

Four limitations constrain interpretation. First, one reviewer completed the appraisal; therefore, the breadth and agreement of expert judgments are unknown. Second, the implementation involved 30 students in one school without random assignment or a comparison condition. Third, the analysis was descriptive, and the N-gain categories do not provide an inferential test or causal estimate. Fourth, only an overall spatial-thinking score was reported; the data do not show whether positional reasoning, scale, representation, pattern recognition, or geomorphic explanation changed differently.

The source record also does not provide several details needed for replication and risk-of-bias assessment: grade or age distribution, sampling procedure, intervention duration and lesson frequency, spatial-test item count and reliability, reviewer qualifications, implementation-fidelity evidence, cognitive-load or usability measures, and ethics or consent procedures. These items should be documented in any submission version if verified records are available; they should not be inferred retrospectively.

A stronger next study should use multiple content, media, and measurement experts; report item-level validity and reliability; examine gains by spatial-thinking domain; and capture process measures such as task completion, navigation logs, classroom observation, group discourse, cognitive load, and usability. A comparison condition, baseline-equivalent groups, longer implementation, and recruitment across schools would improve causal and external validity. These additions would test whether the intended sequence produces durable spatial learning, for whom it works, and which geospatial actions account for the change.

4. Conclusion

This study developed a Google Earth-supported learning environment for landform learning using the ADDIE model and explicitly aligned geospatial functions with spatial-thinking indicators. The final product integrated QR-based project access, 2D/3D visualization, zoom and measurement tools, location marking, guided spatial tasks, and a supporting user guide. A single-expert appraisal produced an overall feasibility score of 85.38%, with the media aspect rated highest at 91.67%, followed by evaluation/instrument at 86.67% and material at 78.75%. Reviewer feedback led to substantive revisions in geomorphology content, spatial-thinking activities, geospatial assessment stimuli, and technical-access procedures.

Implementation with 30 respondents showed that the mean score increased from 74.53 to 86.50 and that all students obtained higher posttest scores. The mean N-gain was 0.488, indicating a moderate increase, with 10.0% of respondents in the high category, 86.7% in the moderate category, and 3.3% in the low category. These findings support the feasibility of using the product as an interactive environment for spatial exploration and provide initial evidence of learning-score improvement. However, the absence of a comparison group means that the results should not be interpreted as a causal estimate of effectiveness.

Future research should strengthen the validation process with multiple experts, test the medium in more schools, analyze gains by spatial-thinking indicator, and use quasi-experimental or experimental designs to compare the developed medium with alternative instructional approaches. Further development should also examine device compatibility and connectivity requirements and explore the application of the same design logic to other geography topics that require multiscale and three-dimensional spatial reasoning.

5. Acknowledgments

The authors thank the geography teacher and students of SMA Negeri 2 Konawe Selatan for their cooperation during the needs analysis and classroom implementation. The authors also acknowledge the reviewer who provided formative feedback on the learning materials, media, and assessment instruments.

Funding. The authors received no external funding for the research, authorship, or publication of this article.

Conflict of interest. The authors declare no conflict of interest.

Data availability. De-identified data may be made available by the corresponding author on reasonable request, subject to institutional requirements.

AI Use Declaration. The authors used ChatGPT-5 to assist with translation into English and language editing. The authors reviewed and revised the AI-assisted text and take full responsibility for the final manuscript.

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>
- Arztmann, M., Domínguez Alfaro, J. L., Hornstra, L., Wong, J., Jeuring, J., & Kester, L. (2025). To move or not to move? The effect of active versus passive pre-training on cognitive load and in-game performance in an AR game. *British Journal of Educational Technology*, 56(6), 2251–2270. <https://doi.org/10.1111/bjet.13565>
- Bagher, M. M., Sajjadi, P., Wallgrün, J. O., LaFemina, P., & Klippel, A. (2023). Virtual reality for geospatial education: Immersive technologies enhance sense of embodiment. *Cartography and Geographic Information Science*, 50(3), 233–248. <https://doi.org/10.1080/15230406.2022.2122569>
- Bokingo, M. F., Rusiyah, R., Nurfaika, N., Lahay, R. J., & Hendra, H. (2023). Pengembangan media pembelajaran berbasis augmented reality pada pembelajaran geografi materi dinamika litosfer. *Social Landscape Journal*, 4(3), 212–223. <https://doi.org/10.56680/slj.v4i3.53977>
- Bolkas, D., O'Banion, M., Parisi, E. I., Bauer, P., & Detchev, I. (2024). Demonstration of immersive technologies for geospatial learning. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-5-2024, 1–7. <https://doi.org/10.5194/isprs-annals-X-5-2024-1-2024>
- Branch, R. M. (2009). *Instructional design: The ADDIE approach*. Springer. <https://doi.org/10.1007/978-0-387-09506-6>
- Buchner, J., Buntins, K., & Kerres, M. (2022). The impact of augmented reality on cognitive load and performance: A systematic review. *Journal of Computer Assisted Learning*, 38(1), 285–303. <https://doi.org/10.1111/jcal.12617>
- Cook, M. L., Sheffield, S. L., Ryan, J. G., Bebeau, C. M., & Rodgers, M. (2025). Enhancing geoscience education through immersive virtual field trips: Student perspectives and implications for broadening participation. *Journal of Geoscience Education*, 73(4), 352–369. <https://doi.org/10.1080/10899995.2025.2537995>
- Cuendet, S., Bonnard, Q., Do-Lenh, S., & Dillenbourg, P. (2013). Designing augmented reality for the classroom. *Computers & Education*, 68, 557–569. <https://doi.org/10.1016/j.compedu.2013.02.015>
- Duarte, L., Teodoro, A. C., & Gonçalves, H. (2022). Evaluation of spatial thinking ability based on exposure to geographical information systems (GIS) concepts in the context of higher education. *ISPRS International Journal of Geo-Information*, 11(8), Article 417. <https://doi.org/10.3390/ijgi11080417>
- Dumako, N. O., Eraku, S. S., & Pambudi, M. R. (2025). Pengembangan media pembelajaran geografi pada materi sistem informasi geografis (SIG) untuk meningkatkan hasil belajar siswa kelas X SMA Negeri 6 Kota Gorontalo menggunakan Edpuzzle. *Jurnal Riset dan Pengabdian Interdisipliner*, 2(3), 468–476. <https://ejurnal.ung.ac.id/index.php/jrpi/article/view/31970>
- Edler, D., Wiktorin, D., Becker, G., Schweins, U., Reichert, K., Keil, J., & Dickmann, F. (2025). The potential of immersive virtual reality for teaching geography in schools. *KN - Journal of Cartography and Geographic Information*, 75(4), 205–217. <https://doi.org/10.1007/s42489-025-00204-1>
- Elford, D., Lancaster, S. J., & Jones, G. A. (2022). Exploring the effect of augmented reality on cognitive load, attitude, spatial ability, and stereochemical perception. *Journal of Science Education and Technology*, 31(3), 322–339. <https://doi.org/10.1007/s10956-022-09957-0>

- 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>
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64–74. <https://doi.org/10.1119/1.18809>
- Ishikawa, T. (2016). Spatial thinking in geographic information science: Students' geospatial conceptions, map-based reasoning, and spatial visualization ability. *Annals of the American Association of Geographers*, 106(1), 76–95. <https://doi.org/10.1080/00045608.2015.1064342>
- Jaeger, A. J. (2024). Google Earth as a tool for supporting geospatial thinking. *Land*, 13(12), Article 2218. <https://doi.org/10.3390/land13122218>
- Jian, Y., & Abu Bakar, J. A. (2024). Comparing cognitive load in learning spatial ability: Immersive learning environment vs. digital learning media. *Discover Sustainability*, 5, Article 111. <https://doi.org/10.1007/s43621-024-00310-6>
- Koç, T., & Topu, F. B. (2022). Using three-dimensional geospatial technology in primary school: Students' achievements, spatial thinking skills, cognitive load levels, experiences and teachers' opinions. *Education and Information Technologies*, 27(4), 4925–4954. <https://doi.org/10.1007/s10639-021-10810-x>
- Koumpourous, Y. (2024). Revealing the true potential and prospects of augmented reality in education. *Smart Learning Environments*, 11, Article 2. <https://doi.org/10.1186/s40561-023-00288-0>
- Lawson, A. P., Martella, A. M., LaBonte, K., Delgado, C. Y., Zhao, F., Gluck, J. A., Munns, M. E., Wells LeRoy, A., & Mayer, R. E. (2024). Confounded or controlled? A systematic review of media comparison studies involving immersive virtual reality for STEM education. *Educational Psychology Review*, 36(3), Article 69. <https://doi.org/10.1007/s10648-024-09908-8>
- Li, S., Jiao, X., Cai, S., & Shen, Y. (2025). Enhancing AR-based learning environments for STEM education: A design-based study on design features, kinematics learning and mathematics self-efficacy. *British Journal of Educational Technology*, 56(4), 1438–1462. <https://doi.org/10.1111/bjet.13528>
- 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>
- McKenney, S., & Reeves, T. C. (2018). *Conducting educational design research* (2nd ed.). Routledge. <https://doi.org/10.4324/9781315105642>
- Meeradji, H., Eraku, S. S., & Masruroh. (2024). Pengembangan media pembelajaran berbantuan Nearpod pada mata pelajaran geografi materi dinamika litosfer di kelas X SMA Negeri 5 Gorontalo. *Jurnal Samudra Geografi*, 7(2), 99–108. <https://doi.org/10.33059/jsg.v7i2.10890>
- Mokmin, N. A. M., Hanjun, S., Jing, C., & Qi, S. (2024). Impact of an AR-based learning approach on the learning achievement, motivation, and cognitive load of students on a design course. *Journal of Computers in Education*, 11(2), 557–574. <https://doi.org/10.1007/s40692-023-00270-2>
- Na, H., Staudt Willet, K. B., & Kim, C. (2025). Investigating the impact of AR technologies on geometric learning in primary school: A comparison between marker-based and markerless AR. *British Journal of Educational Technology*, 56(6), 2502–2521. <https://doi.org/10.1111/bjet.13584>
- Nissen, J. M., Talbot, R. M., Thompson, A. N., & Van Dusen, B. (2018). Comparison of normalized gain and Cohen's d for analyzing gains on concept inventories. *Physical Review Physics Education Research*, 14(1), Article 010115. <https://doi.org/10.1103/PhysRevPhysEducRes.14.010115>
- Ortiz-Guerrero, C., & Loizzo, J. (2024). Rocks really rock: Electronic field trips via Web Google Earth can generate positive impacts in attitudes toward Earth sciences in middle- and high-school students. *Geoscience Communication*, 7(2), 101–119. <https://doi.org/10.5194/gc-7-101-2024>
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489–497. <https://doi.org/10.1002/nur.20147>
- Pratama, M. I. L., Yusuf, D., Maryati, S., Rusiyah, R., Kobi, W., Melo, R. H., Pambudi, M. R., Masruroh, M., Hendra, H., & Asrul, A. (2025). Expanding spatial understanding through mobile augmented reality: A contemporary synthesis of pedagogical, technological, and cognitive dimensions in geography education. *Geosfera: Jurnal Penelitian Geografi*, 4(2), 242–252. <https://doi.org/10.37905/geojpg.v4i2.35155>

- Poupard, M., Larrue, F., Sauzéon, H., & Tricot, A. (2025). A systematic review of immersive technologies for education: Learning performance, cognitive load and intrinsic motivation. *British Journal of Educational Technology*, 56(1), 5–41. <https://doi.org/10.1111/bjet.13503>
- Schmidt, R., & Stumpe, B. (2025). Systematic review of mobile augmented reality applications in geography education. *Review of Education*, 13(1), e70042. <https://doi.org/10.1002/rev3.70042>
- Semken, S., Mead, C., Foley, K., Ruberto, T., Bruce, G., & Anbar, A. D. (2025). Research on teaching geoscience with virtual field experiences. *Annual Review of Earth and Planetary Sciences*, 53, 425–442. <https://doi.org/10.1146/annurev-earth-060923-115406>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Şimşek, B., Direkci, B., Koparan, B., Canbulat, M., Gülmez, M., & Nalçacıgil, E. (2025). Examining the effect of augmented reality experience duration on reading comprehension and cognitive load. *Education and Information Technologies*, 30(2), 1445–1464. <https://doi.org/10.1007/s10639-024-12864-z>
- Spangenberg, P., Matthes, N., Kapp, F., Kruse, L., & Plass, J. L. (2025). Orchestrating iVR technology in an authentic classroom setting and its effects on factual knowledge, comprehension and transfer. *Educational Technology Research and Development*, 73(1), 387–413. <https://doi.org/10.1007/s11423-024-10409-2>
- Thayaseelan, K., Zhai, Y., Li, S., & Liu, X. (2024). Revalidating a measurement instrument of spatial thinking ability for junior and high school students. *Disciplinary and Interdisciplinary Science Education Research*, 6(1), Article 3. <https://doi.org/10.1186/s43031-024-00095-8>
- Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, 139(2), 352–402. <https://doi.org/10.1037/a0028446>
- Wang, L., Zhang, Q., & Sun, D. (2025). Exploring the impact of an augmented reality-integrated mathematics curriculum on students' spatial skills in elementary school. *International Journal of Science and Mathematics Education*, 23(1), 387–414. <https://doi.org/10.1007/s10763-024-10473-3>
- 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>
- Xiang, X., & Liu, Y. (2017). Understanding 'change' through spatial thinking using Google Earth in secondary geography. *Journal of Computer Assisted Learning*, 33(1), 65–78. <https://doi.org/10.1111/jcal.12166>
- Yang, Y., Du, W., Mavrikis, M., & Geraniou, E. (2026). Spatial skill development through augmented reality in mathematics education: A scoping review. *Digital Experiences in Mathematics Education*, 12(2), 277–310. <https://doi.org/10.1007/s40751-025-00187-8>