12th International Conference of the Immersive Learning Research Network
Selected Academic papers for the iLRN Proceedings
DOI: https://doi.org/10.56198/S7USIZPOE

# An Augmented Reality-Based Proposal for Embodied Spatial Comprehension

Diego Mauricio Rivera Pinzón[1] [0000-0003-3952-295X] and Edward Geovanny García Guerrero[1]

[1] Universidad Pedagógica Nacional, Bogotá, Colombia
dmrivera@pedagogica.edu.co

**Abstract.** This full paper proposes an augmented reality (AR) learning solution for orthographic projections that goes beyond static 3D visualization by deliberately incorporating full-body movement as part of the learning activity. Recent advances in educational technology highlight the need for more interactive and experiential approaches to teaching spatial concepts in technical drawing, architecture, and engineering, as students often struggle to bridge the gap between two-dimensional representations and the three-dimensional objects they depict. The proposal integrates anchored virtual models into real-world environments and guides learners through progressively complex activities. Early tasks help students interpret views from simple segments or planes, while later tasks require reconstructing complex solids from multiple 2D projections and identifying hidden features and sectional views. Immediate, incremental feedback supports real-time error correction, and light gamification (e.g., badges and progress tracking) reinforces motivation and self-paced exploration. A key design contribution is the requirement to physically move around anchored objects by changing angles and perspectives to achieve alignment and access specific views, linking sensorimotor exploration with graphic conventions. While this paper focuses on the design and implementation, it also outlines how logged interaction data (e.g., time on task and movement traces) can support future empirical validation of usability and learning outcomes in higher education contexts.

**Keywords:** Augmented Reality, Spatial Perception, Orthographic Projections, Kinesthetic Learning, Educational Technology.

## 1 Introduction

The visualization and mental manipulation of three-dimensional objects from their two-dimensional representations is a fundamental skill in the training of engineers, architects, and other STEM professionals. However, numerous studies have documented the cognitive difficulties students face when transitioning from 2D views to conceiving 3D models, such as in the topic of orthographic projections [1]. Pre-university and early university-level students often struggle to interpret plans and mentally generate the corresponding three-dimensional shape [2]. Instructors report that many students "are unable to properly understand or process visual images and mental models" of engineering objects using traditional methods. These deficiencies in the 2D–3D transition are attributed to insufficient spatial training, which overloads working memory and hinders the learning of graphical and design subjects. As a result, it is common for novice students to lack the spatial perception needed to successfully engage in technical drawing and modeling courses, highlighting the need for educational strategies that address this issue from the early stages of higher education.

The relevance of spatial skills in university contexts is widely supported by literature. Various authors agree that spatial visualization ability is a significant predictor of academic success in STEM disciplines [3]. For example, it has been found that spatial ability strongly correlates with performance in engineering and science, as well as with career choice in these fields [4]. Moreover, mastering these skills facilitates the understanding of abstract concepts and improves the solving of complex problems [5]. Nevertheless, spatial competence varies among students depending on factors such as prior experience, motivation, and even gender [6]. Studies with engineering students show that those with better spatial imagery (the ability to form detailed mental images) achieve higher performance in 2D–3D projection tasks than their peers with less efficient strategies [1]. Likewise, researchers such as Sorby have demonstrated that instruction focused on spatial visualization leads to significant improvements in student performance and retention in engineering "the answer is yes" to whether training these skills improves STEM outcomes [6]. In contrast, the lack of adequate spatial training can result in academic disadvantages: for example, Gold et al. report that weaknesses in spatial skills (influenced by gender differences, motivation, and childhood play opportunities) are reflected in lower performance in early courses. This underscores the importance of integrating visuo-spatial activities into traditional teaching, which has historically placed more emphasis on the memorization of procedures than on the development of spatial intuition. In sum, developing spatial perception and visualization in university students not only addresses a specific difficulty in interpreting projections, but also establishes crucial cognitive foundations for their success in scientific and technical fields.

Considering this situation, educational innovation has explored technological tools such as Augmented Reality (AR) to enhance spatial skills and ease cognitive load during the 2D–3D transition. AR merges 3D virtual objects with the real environment, allowing students to visualize three-dimensional models overlaid on traditional instructional materials [6]. In recent years, numerous studies have been published on the use of AR in educational settings, showing positive trends in learning outcomes. On one hand, systematic reviews and meta-analyses confirm that AR provides measurable pedagogical benefits: Garzón et al. found in 61 studies that educational systems using AR tend to produce moderate improvements in learning effectiveness ($d \approx 0.64$) compared to conventional methods [7]. Among the most frequently reported advantages of incorporating AR in the classroom are increased learning achievement and enhanced student motivation. These gains are evident across different educational levels and subjects, suggesting that AR has reached a certain level of maturity and cross-disciplinary adoption in education.

On the other hand, recent experimental studies reinforce the effectiveness of AR specifically for developing spatial skills. For instance, [8] implemented a mobile AR application in an orthographic projection course and found that the group of students using AR significantly improved their mental rotation skills (measured with standardized tests) more than the control group receiving traditional instruction. Additionally, students gave positive feedback on the augmented experience, perceiving it as a useful and engaging tool for visualizing 3D objects, low-cost and classroom viable. Similarly, [9] developed EDINAR, an AR application for teaching engineering drawing, and documented in a sample of 392 students that those who learned with AR obtained better grades and conceptual understanding than those taught with conventional methods. These empirical results support the inclusion of AR as an instructional resource in graphics courses, by improving both academic performance and student satisfaction.

Equally important, recent literature emphasizes pedagogical design considerations when creating educational AR applications. It is recommended to include features that support incremental understanding (e.g., manipulable 3D models, color-coded faces and edges, and step-by-step animations) to facilitate the connection between orthographic and isometric views [6]. Gamification elements—such as rewards or interactive challenges—have also been incorporated into AR to increase active student engagement. Evidence suggests that when AR is carefully designed to leverage its interactive 3D visualization capabilities, it can significantly enhance the development of spatial skills and reduce traditional barriers to accessing three-dimensional materials [10]. Nonetheless, challenges have also been identified: some studies point to usability issues and technical difficulties (e.g., poorly detected markers or inadequate mobile devices) that can undermine AR's educational effectiveness [11]. Consequently, researchers urge a rigorous evaluation of AR design strategies—considering intervention duration, type of interaction, and level of immersion—to maximize its positive impact on learning. In summary, AR emerges as a promising tool to "bridge" the gap between physical models and spatial conceptualization, offering more intuitive learning environments where students can virtually see and manipulate what they previously could only imagine.

Another contemporary line of educational research highlights the role of kinesthetic and embodied learning in the development of spatial cognition. The theory of embodied cognition posits that the human mind does not operate in isolation from the body, but that cognitive processes, particularly spatial processes, are grounded in sensorimotor experience and physical interaction with the environment [4]. Applied to STEM education, this implies that actively involving students' bodies in visualization tasks could improve their spatial understanding. DeSutter and Stieff (2017) review evidence showing how bodily actions aligned with spatial goals (e.g., gestures, movements, or concrete manipulations) can reinforce spatial thinking in science and engineering. In particular, it has been observed that encouraging students to gesture or move while reasoning about geometric structures helps them build stronger mental representations of those structures. This is consistent with findings in cognitive psychology: when people link bodily movements to a spatial task, their performance in that task often improves [12]. For instance, studies report that performing rotational movements with the hands or body can facilitate mental rotation operations of objects in the mind, reducing cognitive load for beginners. In [3], these principles were applied in the design of tangible learning interfaces, demonstrating that physical and immersive interfaces that require students to move and interact with virtual objects can leverage the body–spatial cognition connection to enhance skills such as 3D visualization. Likewise, in ref. [13] implemented an educational program with embodied activities (such as theater and manipulation of concrete objects) to promote spatial reasoning in elementary school girls, finding that such sensorimotor experiences significantly enrich participants' visual–spatial thinking.

In higher education, integrating kinesthetic elements can be as simple as allowing students to use physical models, make gestures while explaining concepts, or use AR/VR environments where they must move their bodies to explore 3D models. These approaches capitalize on the fact that spatial perception is intrinsically linked to the body: learning by doing is more effective than by observing alone. Therefore, combining AR with embodied learning methods could yield synergistic benefits. In fact, AR naturally involves some degree of physical interaction (e.g., moving a device to view an object from different angles, aligning markers with the camera), making it a promising technology for active and kinesthetic learning. In sum, the contributions of embodied learning suggest that spatial skills are not developed solely through abstract exercises on paper, but also by moving the body and objects to explore space tangibly. This approach complements digital tools: while AR provides enriched visualization, physical activities provide sensorimotor intuition, together contributing to a deeper understanding of three-dimensional space.

Considering all of the above, the search for comprehensive educational solutions that address the problem from multiple angles is fully justified. The present work focuses on the technological presentation of an AR application called AR-Views, designed for teaching orthographic projections in higher education. AR-Views aims to combine the advantages of augmented visualization with kinesthetic learning: the application allows students to interact with 3D models overlaid on the real world, encouraging them to view, manipulate, and explore projections using their own mobile devices. This work contributes a methodological design proposal in which full-body movement is not incidental but required to obtain and align viewpoints around spatially anchored virtual models. Building on this principle, the paper details a scaffolded activity sequence for orthographic projection practice, specifies the multimodal feedback and light gamification mechanisms used to support learning, and documents the prototype implementation to enable reuse and future empirical validation.

The following sections outline the theoretical and technological foundations of AR-Views and describe its implementation as a design proposal. The paper focuses on the interaction principles, with particular emphasis on full-body exploration around spatially anchored models, and outlines how future studies can evaluate its usability and potential contribution to learning orthographic projections and spatial skills in STEM education.

## 2 Methodological Proposal

This methodological proposal envisions an augmented reality learning environment in which virtual models are anchored rigidly to the physical world. Once placed, each object remains fixed in orientation, and only its initial scale may be adjusted, so that every change of viewpoint must be achieved by the learner's own movement, such as walking, bending, or raising the device. This constraint turns mental rotation into embodied action and reveals how the observer's position governs the dimensional information preserved in each orthographic view. Building on that interaction principle, the sequence of activities alternates between generating two-dimensional projections from a stationary solid and, inversely, reconstructing the solid from given projections. An algorithm continuously checks whether the device lies within the required sightline cone; when alignment is reached, visual cues and subtle haptic pulses confirm success and log the event for later analysis. The design thus offers a time-flexible space for exploratory practice, blending bodily engagement, immediate feedback, and iterative refinement to nurture robust spatial reasoning.

### 2.1 Spatial-Anchored Interaction Principle

**World-fixed placement:** The virtual object is anchored in the room and cannot be rotated with on-screen gestures; only uniform scale may be adjusted during placement.

**Body-mediated exploration:** Learners physically walk, bend, or raise the device to obtain new perspectives. Mental rotation is externalized as real movement, making the transformation from 2D view to 3D form perceptible.

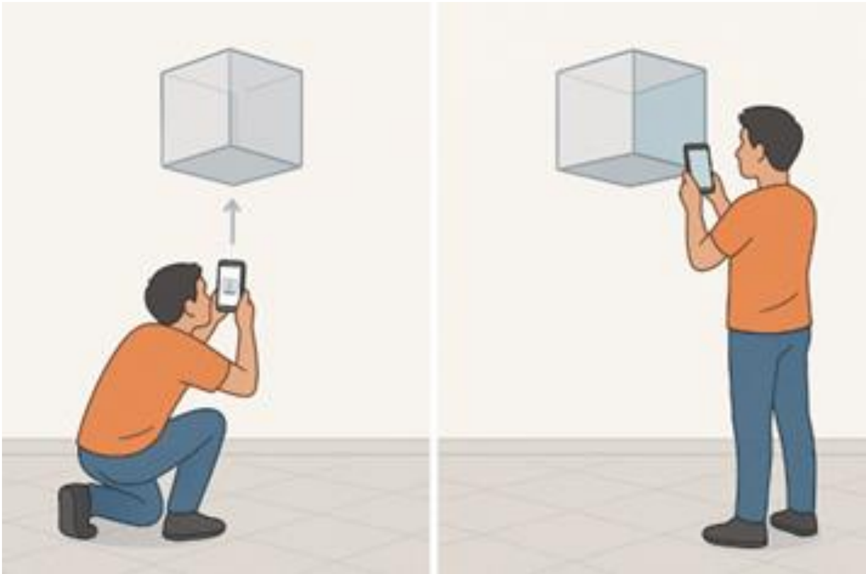

**Fig. 1.** Two viewpoints of a floating cube.

### 2.2 Key Viewing Situations

**Table 1.** Key Viewing Situations.

| Situation | What the learner perceives | Why it matters |
|---|---|---|
| Critical viewpoint of a line | A straight line collapses to a single point when the learner looks exactly along its direction. | Highlights how projection removes one dimension. |
| Critical viewpoint of a plane | A plane collapses to a straight line when viewed along its normal. | Demonstrates the loss of one degree of freedom in projection. |
| Principal views of a solid | Front, top, and right views appear only from three mutually perpendicular sightlines. | Forms the basis of orthographic drawing conventions. |

Older literature sometimes calls these “degenerate cases”; here we prefer critical viewpoints because they represent the precise positions where dimensional information vanishes.

### 2.3 Learning Task Families

**View-finding (3D → 2D).** Objective: Locate the exact positions from which the anchored solid produces its canonical orthographic views.
Experience: The system highlights a target view name (“Top”), and the learner moves until a visual ring turns green, indicating correct alignment.

**Reconstruction (2D → 3D).** Objective: Start with three fixed 2D projections and edit a provisional 3D model until its silhouettes match all views.
Experience: Learners combine primitives or adjust Boolean cuts while walking around the object to observe the effect from each projection direction. Alignment indicators confirm progress.

**Critical-view discovery.** Objective: Find the unique position where a displayed line or plane collapses to a lower-dimensional shape (point or line).
Experience: Subtle colour cues fade in as the learner approaches the critical cone; a final chime signals success.

**Projection-selection challenge.** Objective: With the object fixed, choose—on a side panel—the set of 2D projections that correspond to the live 3D pose.
Experience: The learner circles the model, comparing possible answer thumbnails with the real-time view on screen.

**Object-modification challenge.** Objective: With three projections locked in place, adjust the 3D object (scale, extrude, subtract) so that the live view matches each projection when the device is positioned at the indicated spots.
Experience: This variant deepens modelling skills: each physical relocation is paired with an immediate geometric edit.

These task families can be sequenced from guided demonstration to open exploration; there is no time pressure—learners are encouraged to linger, test hypotheses, and repeat viewpoints freely.

### 2.4 Automatic View Validation

For every activity the application continuously compares the device's camera axis with the pre-defined target axis for the desired view. When the angular difference falls within a small tolerance (configurable in the authoring tool), the view is considered "achieved" and the corresponding step is marked complete. This programmable check:

- eliminates guesswork for the learner,
- provides unambiguous triggers for feedback, and
- records achievement data for later analytics.

### 2.5 Feedback Channels

Visual feedback

- A thin outline around the model shifts from red to green as alignment improves.
- Subtle glow effects appear on correctly matched silhouettes during reconstruction tasks.

Haptic feedback

- Short vibration bursts (e.g., two pulses per second) indicate that the learner is outside the acceptable cone of vision.
- A single long pulse confirms successful view attainment.
  Haptic cues ensure that attention is not monopolized by the screen—learners can focus on bodily alignment while receiving tactile guidance.

### 2.6 Bidirectional Application of Orthographic Skills

All task families deliberately support the two complementary competencies that underpin engineering graphics:

- Interpretation of orthographic views (inferring 3D form from given 2D drawings).
- Generation of orthographic views (producing accurate 2D drawings from a 3D form).

By alternating object-fixed and projection-fixed scenarios, AR-Views ensures that learners practice both directions of the spatial-reasoning cycle, strengthening the mental link between tangible forms and their symbolic projections.

## 3 Results

The methodological principles were instantiated in a working prototype developed during the thesis project with Unity 2022 LTS and ARFoundation for ARCore/ARKit smartphones. After an initial calibration that lets learners place and scale the model, the object's orientation is permanently anchored to the physical room; every subsequent viewpoint must be achieved by bodily displacement.

### 3.1 Sequenced Activities Implemented

The prototype delivers a five-step progression (RA 1–RA 5) that mirrors the cognitive ladder proposed in the methodological section while embedding a single motivational mechanic: a badge awarded at the end of each activity. A persistent dashboard shows both badge collection and percentage completion, but no timers or failure states are imposed, learners may revisit any task until they feel confident.

**RA 1 – Seeing a Line as a Point.**
A floating segment AB appears after plane detection. The learner must align the device so that the whole segment collapses to a single point—first along the X axis, then Z, then Y. Each successful alignment is confirmed by a green ring and a long vibration; completing the three orientations unlocks the first badge. To avoid rote repetition, the segment AB is re-oriented at random around its centroid each time the task reloads; learners must therefore discover a new alignment for every attempt. (see Fig. 2).

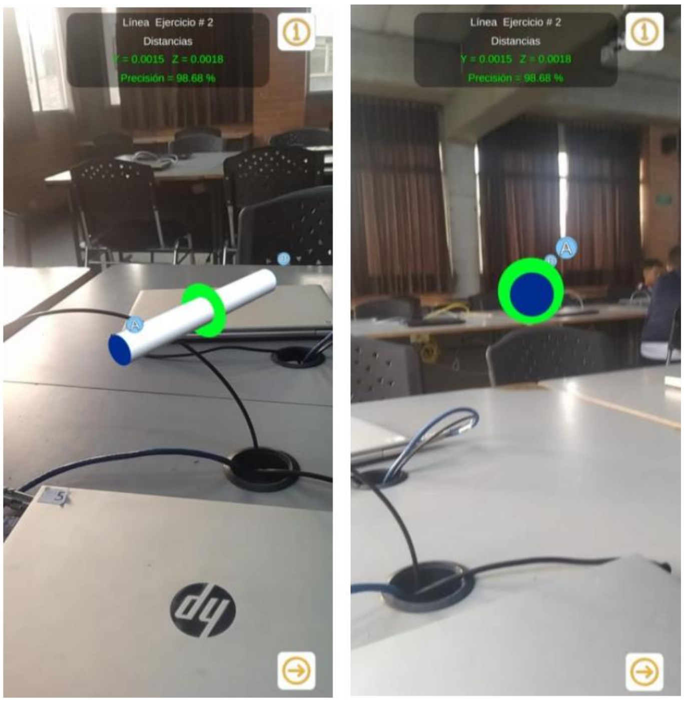


**Fig. 2.** Two screenshots of activity 1.

**RA 2 – Seeing a Plane as a Line.**
The next activity anchors a square plane ABCD. Again, the learner aligns the camera three times, but now the goal is to view the plane edge-on so that it reduces to a straight line. Progress is tracked exactly as in RA 1; a second badge is issued after the vertical-axis attempt is validated. The procedure mirrors RA 1: the square plane is randomly rotated before each trial, requiring the learner to locate a fresh edge-on viewpoint in every session (see Fig. 3).

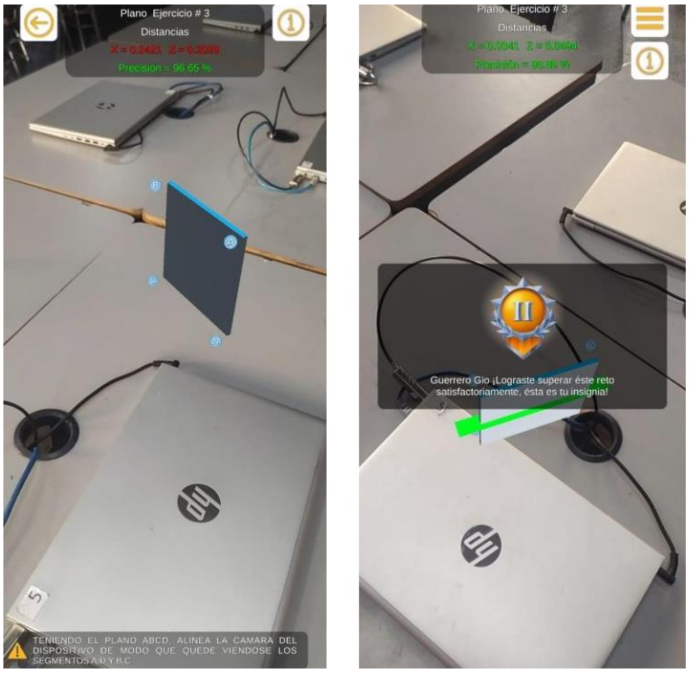


**Fig. 3.** Two screenshots of activity 2.

**RA 3 – Locating the Three Principal Views (3D → 2D).**
A polyhedral solid with cavities and hidden features is presented. The learner must physically reach the right, front and top sight-lines. When alignment is correct, the corresponding orthographic projection is overlaid on screen, including hidden-line notation, reinforcing the mapping from spatial pose to 2D drawing. Beyond identifying the canonical front, top, and right views of a fixed polyhedral solid, this stage emphasises the dynamic mapping between bodily position and 2D representation. As the learner circles the object, a semi-transparent overlay of the emerging projection fades in and out, allowing real-time comparison between the live silhouette and its orthographic counterpart. Hidden features are revealed in dashed lines, reinforcing conventional drawing rules while maintaining the embodied exploration paradigm (see Fig. 4).

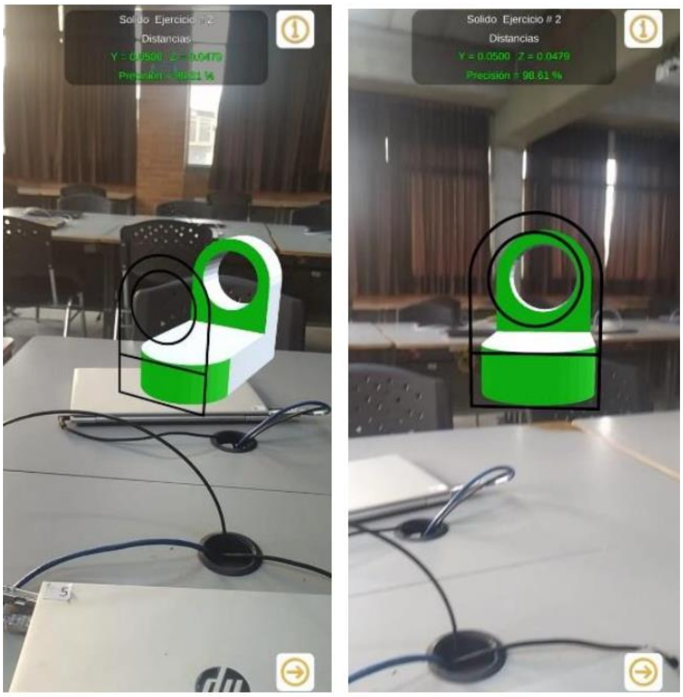


**Fig. 4.** Two screenshots of activity 3.

**RA 4 – Ten Practice Drills (3D → 2D and 2D → 3D).**
This module alternates between two modes. In view-selection mode the solid is fixed; the learner walks around it and chooses, from four thumbnails, the triplet of projections that matches the live object.

In modification mode ten voxel-based solids must be sculpted—adding or removing cubes—so that the live silhouettes coincide with static drawings pinned on the interface. Incorrect attempts trigger on-screen hints and motivational messages, but the learner may restart freely until all ten figures are cleared and the fourth badge is awarded. Each drill alternates between projection-selection and object-modification modes, forcing the learner to switch constantly between interpreting and generating projections; the variety of voxel-based models and the absence of time limits foster deliberate practice (see Fig. 5).

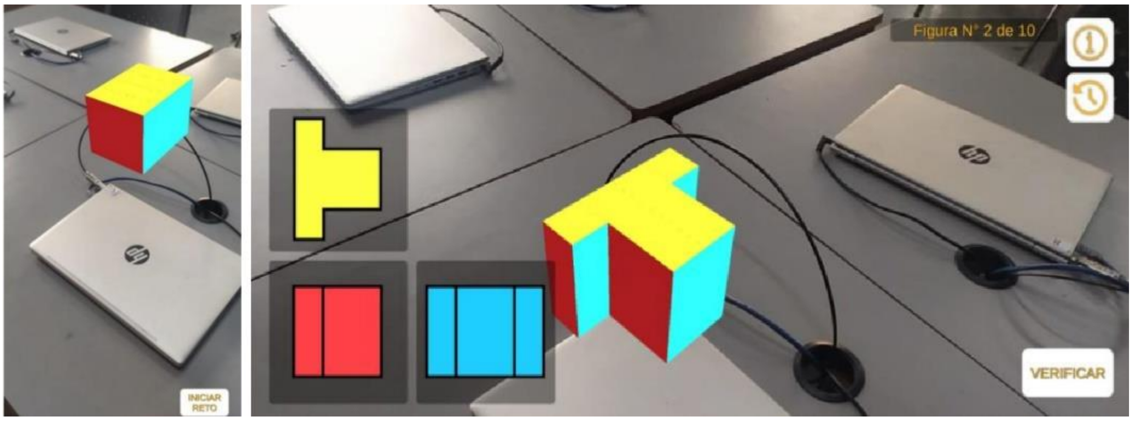


**Fig. 5.** Two screenshots of activity 4. On the left is the cube before being edited by the user, on the right is a modified version.

**RA 5 – Interactive Examination.**
The capstone presents five unseen solids, one at a time, together with four possible orthographic solutions. The learner orbits the model, zooming or crouching as needed, and selects a single answer per item. After each choice, immediate feedback appears; at the end, a results screen displays the score, recommends a short explanatory video and pins the final badge to the collection. For every item, an unfamiliar solid appears with a floating arrow that marks the reference sightline. Four candidate triplets of orthographic views are displayed below; the learner orbits the object, aligns the camera with the arrow, and taps the projection set that correctly matches the live silhouette. Immediate feedback confirms the choice and advances to the next item (see Fig. 6).

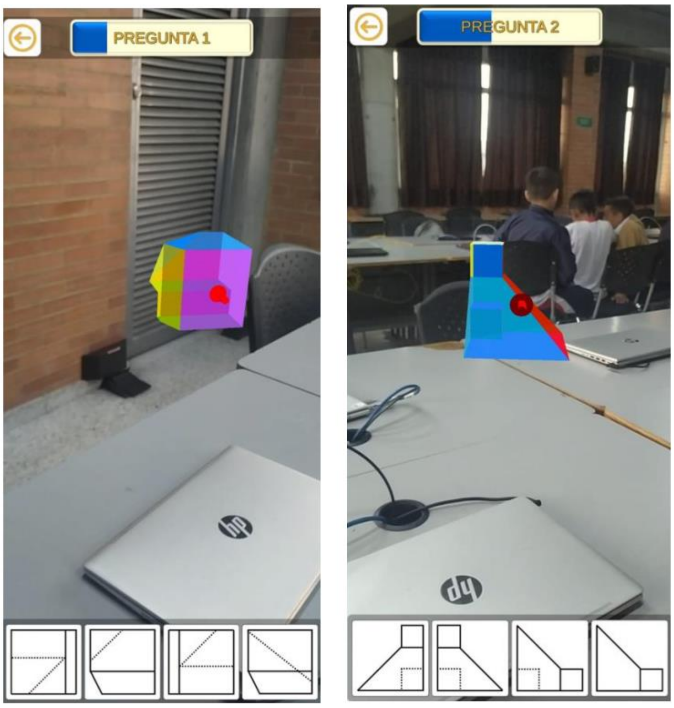


**Fig. 6.** Two screenshots of activity 5.

Across all activities, automatic view-detection (angular tolerance) drives completion events, while brief vibration bursts keep attention on body alignment rather than interface widgets. The absence of countdowns or punitive scoring transforms the sequence into an open exploration space that privileges mastery by repetition over speed.

### 3.2 Gamification and Feedback

A lightweight reward layer accompanies the sequence. Each validated view or silhouette match yields experience points that fill a persistent progress bar; completing every task in a module awards a badge that unlocks optional "assist-off" variants. Consecutive days of practice increment a streak multiplier, encouraging habitual engagement without penalizing absences. Feedback is multimodal: the object outline transitions from red to green in real time, while short vibration pulses signal misalignment and a single long pulse confirms success. These cues keep attention on bodily positioning rather than on interface widgets.

### 3.3 Data Logging

To enable later analysis of navigation strategies, the application records anonymized tuples containing timestamp, camera pose, angular error and task identifier. Logs remain on the device and can be exported in JSON format for research without capturing personal imagery, satisfying institutional privacy requirements.

## 4 Conclusions

The prototype suggests that fixing virtual objects in physical space and inviting learners to move—rather than only manipulating models on-screen—provides a promising setting for investigating the role of embodied action in spatial cognition. Informal trial sessions, although outside the scope of this article's results, indicate that crouching, orbiting, and other full-body motions may become part of the learner's problem-solving strategies, motivating future exploratory studies on kinesthetic learning within AR environments.

Beyond its immediate application to orthographic projection, the work contributes a transferable design pattern: combining automatic view detection, multimodal feedback, and an exploration-friendly reward system can help connect abstract graphical conventions with sensorimotor exploration during learning activities. Future research can extend these ideas to curved-surface geometry, intersecting solids or even non-engineering domains where spatial reasoning is critical, refining our understanding of how embodied interaction and augmented reality may jointly support conceptual understanding.

Although a systematic evaluation is outside the scope of this paper, the system is instrumented to support future analysis through interaction logs. These logs can be used to compute aggregate indicators such as time-on-task per activity, number of viewpoint adjustments before alignment, and movement traces around the anchored model, enabling both usability analysis and exploratory links between sensorimotor exploration and task performance.

## Acknowledgments

Project was developed with the support and financing of the Research Center of the Pedagogical University - CIUP and the Department of Technology of the National Pedagogical University of Colombia (National Pedagogical University, UPN).

## Disclosure of Interests.

The authors have no competing interests to declare that are relevant to the content of this article.

## References

1. Kok, P.J.: Pre-Service Teachers' Visuospatial Cognition: 2D to 3D Transition. *African Journal of Research in Mathematics, Science and Technology Education* 24(3), 293–306 (2020).

2. Alqahtani, A.S., Daghestani, L.F., Ibrahim, L.F.: Techniques used to Improve Spatial Visualization Skills of Students in Engineering Graphics Course: A Survey. *International Journal of Advanced Computer Science and Applications* 8(3), 91–100 (2017).
3. Clifton, P.G., et al.: Design of embodied interfaces for engaging spatial cognition. *Cognitive Research: Principles and Implications* 1, Art. 24 (2016).
4. DeSutter, D., Stieff, M.: Teaching students to think spatially through embodied actions: Design principles for learning environments in STEM. *Cognitive Research: Principles and Implications* 2, Art. 22 (2017).
5. Re, A., Perna, S., Città, G., Tosto, C., Gentile, M.: Embodied Cognition in learning STEM disciplines. An overview on the role of spatial skills. *QTimes – Journal of Education, Technology and Social Studies* (2024).
6. Sun, L., Chandramouli, M., Chew, K.J.: Augmented Reality: Bridging the Inclusivity Gap in Engineering Graphics Education through Spatial Skills Enhancement. In: *2024 South East Section Meeting*, Marietta, Georgia (2024).
7. Garzón, J., Pavón, J., Baldiris, S.: Systematic review and meta-analysis of augmented reality in educational settings. *Virtual Reality* 23(4), 447–459 (2019).
8. Omar, M., Farzeeha, D., Mokhtar, M., Zaid, N.M.: Effects of Mobile Augmented Reality (MAR) towards students' visualization skills when learning orthographic projection. *International Journal of Emerging Technologies in Learning* 14(20), 106–121 (2019).
9. Tiwari, A.S., Bhagat, K.K., Lampropoulos, G.: Designing and evaluating an augmented reality system for an engineering drawing course. *Smart Learning Environments* 11, Art. no. 1 (2024).
10. Akçayır, M., Akçayır, G.: Advantages and challenges associated with augmented reality for education: A systematic review of the literature. *Educational Research Review* 20, 1–11 (2017).
11. Chang, H.-Y., et al.: Ten years of augmented reality in education: A meta-analysis of (quasi-)experimental studies to investigate the impact. *Computers & Education* 191, Art. 104641 (2022).
12. Amico, G., Schaefer, S.: Negative Effects of Embodiment in a Visuo-Spatial Working Memory Task in Children, Young Adults, and Older Adults. *Frontiers in Psychology* 12, 688174 (2021).
13. Frazier, A.D., Bryant, C.: Developing visual-spatial thinking in youth using sensorimotor experiences: Approaches from a Piagetian cognitive framework. *Journal of Pedagogical Research* 3(3), 99–112 (2019).