European Journal of Science and Mathematics Education, 2026, **14**(3), 471-488
ISSN: 2301-251X (Online)

OPEN ACCESS

Research Article

AI-supported hybrid learning in exploratory geometry: For higher education

Sara Cruz [1,2,3*]
0000-0002-9918-9290

Floriano Viseu [3]
0000-0002-8221-6870

[1] Applied Artificial Intelligence Laboratory, Polytechnic Institute of Cávado and Ave, Barcelos, PORTUGAL
[2] Center for Research and Innovation in Education, Polytechnic Institute of Porto, Porto, PORTUGAL
[3] Research Center on Education, University of Minho, Braga, PORTUGAL
[*] Corresponding author: scruz@ipca.pt

Citation: Cruz, S., & Viseu, F. (2026). AI-supported hybrid learning in exploratory geometry: For higher education. *European Journal of Science and Mathematics Education, 14*(3), 471-488. https://doi.org/10.30935/scimath/18798

ARTICLE INFO

Received: 7 Oct 2025

Accepted: 17 May 2026

ABSTRACT

The abstract nature of mathematical concepts often impedes learners' comprehension, particularly within geometry, a difficulty accentuated at the transition from secondary to higher education, where students shift from procedural fluency to formal, axiomatic reasoning. The goal of this study was to understand how a hybrid flipped learning design, mediated by Moodle and supported by artificial intelligence (AI), articulating tangram, GeoGebra, and Polypad, improves performance in geometry and qualifies students' geometric reasoning, by mapping the appropriation of AI-generated explanations in asynchronous interactions. Methodologically, we adopted a mixed-methods, single-group pre-/post-design (*n* = 22) within a hybrid flipped learning cycle streamlined to asynchronous preparation via Moodle and AI tools and studio-style, in-class sessions; qualitative data comprised forum posts analyzed through directed content analysis. Pre-/post-comparisons showed statistically significant gains on all four domains, robust to parametric and non-parametric tests; effect sizes ranged from large to very large, with distributional shifts across outcomes. Individually, improvement was most widespread for spatial reasoning and mathematical problem-solving; geometric properties improved for 15 students; geometric deduction was heterogeneous. Qualitatively, students increasingly named and justified properties, described transformations with greater precision, and used AI-generated explanations as scaffolds to verify reasoning, explore alternative representations, and correct misconceptions while maintaining authorship of arguments. These findings indicate the promise of multimodal, AI-supported hybrid designs for early undergraduate geometry learning, while acknowledging limits of causal inference, small sample size, and absent follow-up.

Keywords: higher education, generative AI, hybrid learning, flipped learning, dynamic geometry, spatial reasoning

INTRODUCTION

In contemporary mathematics education, there is a broad consensus that learners progress most powerfully by acting, representing, and reflecting within rich tasks rather than passively receiving information. Socio-constructivist perspectives frame learning as socially mediated and tool-supported: Vygotsky's zone of proximal development (ZPD) foregrounds guided assistance and cultural tools such as language, while Bruner's discovery learning and scaffolding emphasize temporary supports that fade as independence grows (Margolis, 2020; Tohari & Rahman, 2024). In parallel, constructionism positions learning-through-making as a route to conceptual growth, long reflected in technology-enhanced mathematics, including dynamic geometry and the design of tangible-digital tasks. Empirically, concrete manipulatives and structured digital environments catalyze conceptual understanding by linking perception, action, and abstraction (Brum et al., 2023; Costa, 2023); reviews of work with GeoGebra and tangrams report gains in interest and geometric reasoning, including movement along van Hiele levels (Rocha & da Silva, 2021). These strands converge on hybrid, multimodal learning in which students alternate between physical manipulation, digital construction, and explicit reflection. Hoffmann et al. (2009) characterize this orchestration as a pedagogical architecture that seeks didactic adequacy through the coordination of tools and spaces. Research on embodied interaction shows that bodily engagement and tangible manipulation play a distinctive cognitive role, complementing rather than competing with dynamic digital representations (Abrahamson et al., 2023). Experimental evidence cautions against purely virtual approaches: balanced, multimodal designs can reduce cognitive load and improve learning relative to digital-only conditions (Al Hakim et al., 2024).

Concurrently, higher-education students encounter ubiquitous information and AI-mediated explanations; cultivating the capacity to interrogate, verify, and articulate mathematical reasoning is therefore central. When responsibly integrated, AI can function as a formative scaffold, providing criterion-referenced feedback, prompting metacognition, and generating task variants while leaving pedagogical judgement with the instructor. Evidence suggests benefits for self-regulated learning in mathematics (Rebolledo-Mendez et al., 2022) and feasibility in early and primary contexts under teacher guidance (Holmes et al., 2019). Effective implementation, however, demands teacher knowledge at the intersection of content, pedagogy, and technology (technological pedagogical content knowledge [TPACK]) and sustained professional development; ethical considerations (privacy, bias, and transparency) require privacy-by-design and careful governance (Jain et al., 2024; Meylani, 2024; Wang et al., 2023).

Against this backdrop, we report a pilot mixed-methods study with students enrolled in a higher professional technical program in educational technology services at a Portuguese higher-education institution. The study focuses on engagement with exploratory and investigative geometry within a Moodle-mediated hybrid design interleaving tangram manipulation, dynamic-geometry work in GeoGebra, and AI-supported reflective inquiry. The objective of this study was to understand how a hybrid design inspired by SOFLA, mediated by Moodle and supported by AI, articulating tangram, GeoGebra, and Polypad, improves performance in geometry (geometric decomposition [GD], geometric properties [GP], mathematical problem-solving [MPS], and spatial reasoning [SR]) and qualifies students' geometric reasoning, mapping the appropriation of AI-generated explanations in asynchronous interactions. Our research questions (RQs) were:

(1) To what extent does this hybrid, AI-supported design improve performance across GD, GP, MPS, and SR?

(2) In what ways do higher-education students articulate and develop their geometrical reasoning in asynchronous forum exchanges, including their appropriation of AI-generated explanations for sense-making?

We triangulate pre- and post-change scores with a directed content analysis of forum posts, positioning this work as a first step toward principled, AI-enabled hybrid designs in undergraduate geometry learning.

THEORETICAL FRAMEWORK

Mathematical learning experiences in the primary years are strengthened when designed as hybrid, multimodal sequences integrating physical manipulatives, dynamic digital representations, and pupils' verbal and visual production. Such integration supports transduction across representations, develops GD and SR, and can reduce extraneous cognitive load by distributing attention across complementary modes (Abrahamson et al., 2023; Rocha & da Silva, 2021). Here artificial intelligence (AI) serves as a formative scaffold generating criterion-referenced feedback and contrastive explanations, proposing mathematically constrained task variants, and returning orchestration signals to teachers, without displacing professional judgement. This role accords with socio-constructivist accounts of guided assistance and with fading scaffolds as learners gain independence and aligns with evidence that metacognitive prompts and guidance strengthen self-regulated learning in mathematics (Rebolledo-Mendez et al., 2022). Emerging evidence shows that, when designed for teacher-led use, AI is both feasible and pedagogically meaningful in early and primary schooling (Holmes et al., 2019). A coherent theory of change follows across educational phases: a progression that blends hands-on geometric work, dynamic digital exploration and structured reflection creates sustained opportunities for sense-making and movement along van Hiele levels. Crucially, laying this foundation in the first years of schooling, with teachers mediating tasks and AI providing targeted feedback and metacognitive scaffolds, builds GD, SR, and property reasoning that later underpin the formal, axiomatic practices demanded in higher education. In this view, pupils do not merely complete activities; they accumulate durable conceptual resources and greater autonomy that transfer to undergraduate geometry, where success depends on exactly these early-cultivated competencies.

For educational-technology research, this linkage is productive on three fronts. Theoretically, it connects embodiment and multi-representation with responsible algorithmic scaffolds while centering teacher expertise through frameworks such as TPACK (Clark-Wilson et al., 2020). Methodologically, it motivates mixed-methods studies pairing validated learning measures, such as, SR, quality of justifications, with AI-use telemetry and joint displays to relate effects to mechanisms. Practically and in policy terms, it foregrounds portable design principles, graduated hints, transparency, privacy-by-default and a meaningful opt-out, plus guardrails to avoid over-help and dependency, audit bias, and protect pupil data; these concerns are well documented in contemporary debates on ethics, privacy and fairness in AI-enhanced education (Jain et al., 2024; Meylani, 2024; Wang et al., 2023).

Mathematical Learning Experiences in Education

Theoretical frameworks in constructivist and socio-constructivist traditions continue to shape contemporary understandings of mathematical learning. Vygotsky, Bruner, and Papert emphasize learners' active construction of knowledge through interaction, exploration, and creation, informing practices that prioritize experiential, investigative environments. Vygotsky's ZPD motivates tasks just beyond current capability tackled through teacher-mediated, collaborative problem-solving (Janisch & Jelinek, 2023; Tohari & Rahman, 2024). Bruner's discovery learning advocates active engagement, intuitive exploration, and scaffolding that fades with growing independence (Margolis, 2020; Tohari & Rahman, 2024). Papert's constructionism extends this stance via learning-by-making with digital tools and manipulatives.

Despite pedagogical value, ZPD-aligned and discovery approaches demand time, resources, and specialized teacher training; learner variability necessitates differentiation (Tohari & Rahman, 2024). Empirical work reinforces the importance of concrete, digital experience: Costa (2023) evidences concrete materials for integers; Brum et al. (2023) highlight benefits of manipulatives; Janisch and Jelinek (2023) stress language and thought organization; Kampff et al. (2004) show computerized environments supporting significant mathematical knowledge, together underscoring hybrid spaces blending physical and digital modalities.

Specific tools show strong promise. Rocha and Silva (2021) find GeoGebra stimulates interest in plane isometries and, with tangrams, fosters geometric understanding and spatial perception. Advanced tools also develop spatial visualization: Herrera et al. (2024) report a 25% improvement with virtual environments and 3D printing versus minimal control gains; "25%" denotes the mean (M) normalized gain on the revised PSVT:R (30-item rotations): 0.254 (experimental) vs. 0.050 (control) over a 10-week calculus course; the experimental group showed a medium pre-/post-effect (Hedges' $g = 0.325$; $p < .001$), the control change non-significant ($g = 0.035$; $p = .163$). Complementing this, Portuguese pre-service teachers integrating Engineering Design, Tinkercad, and 3D printing deepened content knowledge, consolidated CAD fluency, and articulated classroom uses, while surfacing needs for precision, alignment, and targeted scaffolding (Barbosa & Vale, 2025).

Problem-solving research reflects a broader shift toward technology-enhanced environments. Santos-Trigo (2024) identifies problem formulation and solution strategies as central and argues that digital tools shape posing and exploration by influencing representation, reasoning, and engagement. Dynamic-geometry systems support modelling, connection-making, and problem-posing; a "problematizing principle" places inquiry at the center through cycles of modelling, testing, and redesign (Santos-Trigo, 2024), mobilizing prior knowledge and forging new connections (Leikin & Guberman, 2023). This aligns with Kahneman's (2011) distinction between fast and slow thinking to promote deliberate, reflective engagement. Practical supports, digital walls or notebooks, help students document and monitor work, questions, and ideas, fostering peer interaction and shared evaluation (Santos-Trigo, 2024). Converging theory and evidence thus endorse exploratory, hands-on, multimodal mathematical learning that integrates manipulatives, digital tools, and constructivist principles to cultivate conceptual understanding, autonomy, and engagement.

Hybrid and Multimodal Learning Environments

The integration of digital technologies into mathematics education offers substantial promise but also exposes context-dependent constraints. Contemporary evidence supports hybrid, multimodal learning ecologies, deliberately combining physical, digital, and virtual resources, to deepen engagement, strengthen conceptual understanding, and cultivate autonomy. At a macro-policy level, the Indian experience during and after COVID-19 illustrates this opportunity-challenge duality: rapid scaling revealed deficits in infrastructure, staff development, and curricular alignment, underscoring sustained investment for flexible, equitable access. Simultaneously, hybrid delivery afforded flexibility and self-paced progression but raised risks of social isolation, device/network dependence, and variable course design. Student preference data (Rafee et al., 2024) indicate a turn towards blended arrangements, suggesting technology is most productive when orchestrated with institutional capacity, teacher support, and pedagogical integration, not as a wholesale substitute for embodied, collaborative learning.

Lieban (2023) highlights interactive, creative pedagogies within STEAM/maker culture, advocating dynamic combinations of physical and digital resources for mathematical thinking and problem-solving. This aligns with multimodal learning and the TPACK framework (Clark-Wilson et al., 2020), which locates effective design at the intersection of technological fluency, pedagogy, and content. Effective blended approaches, including flipped learning, depend on well-prepared educators and thoughtful scenarios, mirroring system-level recommendations on accessibility, inclusivity, and retention via investment in spaces, infrastructure, and staff support (Rafee et al., 2024).

At the heart of hybrid designs is the integration of learning spaces, enabling alternation between concrete manipulation and digital construction. This alternation is cognitive as well as functional: tasks mobilize distinct forms of representation, interaction, and reflection. Hoffmann et al. (2009) describe a pedagogical architecture in which orchestrated physical-digital resources increase didactic adequacy across contexts. Such integration supports differentiation and inclusion, critical where visualization, manipulation, and symbolic reasoning intersect, while requiring coherent orchestration to mitigate hybrid frictions (timetabling, connectivity, remote engagement) and to leverage flexibility, self-regulation, and soft-skill development (Rafee et al., 2024).

Research on embodied interaction reinforces these benefits. Abrahamson et al. (2023) argue that tangible manipulation is cognitively consequential, especially in early learning; digital technologies should complement rather than replace it through simulations, dynamic representations, and real-time feedback. Skill and Young (2002) similarly advocate hybrid models that, when well implemented, improve outcomes.

Experimental evidence (Al Hakim et al., 2024) underscores the value of multimodal balance: learners using physical utensils within simulated environments reported lower cognitive load and higher gains than peers using purely virtual tools. This principle extends beyond early schooling to higher education, where tactile experience continues to scaffold abstraction, linking spatial transformation in dynamic geometry to proof, modelling, and problem-solving. Over-reliance on screen-only workflows risks attenuating kinesthetic richness. Accordingly, effective design seeks equilibrium, integrating manipulatives and fabrication (e.g., tangrams and 3D printing) and studio tasks with virtual tools (GeoGebra, AR/VR, and AI-supported inquiry), sequencing hands-on → digital → formal, aligning representations, and making embodied reasoning explicit.

In sum, the literature affirms that hybrid, multimodal environments, intentionally planned, teacher-mediated, and balanced between tactile and digital experiences, enhance engagement, deepen conceptual understanding, and accommodate diverse learners across spaces, provided policy and institutional strategies address equity, infrastructure, and staff development for sustainable scale (Rafee et al., 2024).

The Role of Artificial Intelligence in Mathematics Education

National Council of Teachers of Mathematics (2024) frames AI as a consequential development for mathematics teaching because AI-driven tools can respond to students' thinking and interests in ways that earlier technologies could not, while still requiring the mathematical, pedagogical, and relational expertise of teachers to ensure meaningful use. This position gives institutional weight to a human-centered view of AI integration: AI should augment mathematical sense-making, feedback, access to multiple representations, and instructional responsiveness, rather than replace teacher judgement or students' own reasoning. The integration of AI into mathematics education can personalize instruction, increase student engagement, and enhance academic performance. Recent advances in adaptive systems, intelligent tutoring and machine-learning models, tailor experiences to individual learners' needs, addressing long-standing limitations of uniform instruction that fails to accommodate diverse capabilities and trajectories.

AI-based adaptive learning facilitates personalized pathways, enabling self-paced learning with real-time, tailored feedback. Such individualization improves comprehension of complex concepts and promotes deeper cognitive engagement (Mustafa, 2024). Evidence indicates significant post-assessment gains, enhancing mathematical competence and retention (Dabingaya, 2022; Meylani, 2024). Scalability and cost-effectiveness are additional advantages, supporting large cohorts without compromising instructional quality (Meylani, 2024).

An emergent strand involves AI chatbots, which increasingly support learning in real time. These agents adapt to cognitive styles, promoting inclusivity and responsiveness (Boltayevich et al., 2024), and enhance engagement by providing instant feedback, extending availability beyond classroom hours, and fostering active participation in problem-solving (A'ini & Khoiriyah, 2024; Roca et al., 2024). Chatbots also contribute to administrative efficiency by automating reminders and feedback, allowing educators to focus on pedagogy (Sain et al., 2024).

Complementing these developments, domain-specific large language models (e.g., DeepSeek R1 and MathCoder) couple structured logic, stepwise mathematical reasoning, and contextualized programming guidance with adaptive feedback, addressing limitations of general-purpose LLMs and informing policies for responsible classroom integration (Cohn et al., 2025; Neha & Bhati, 2025). Beyond logistical and instructional benefits, AI chatbots can scaffold metacognition, prompting justification, suggesting alternative strategies, and helping learners monitor and regulate their thinking, particularly relevant in mathematics, where awareness of problem-solving strategies is critical (Rebolledo-Mendez et al., 2022).

Recent research has begun to introduce AI tools in early years and primary education, recognizing their potential to cultivate foundational competencies, such as representational fluency, exploratory problem-solving, metacognitive regulation, and collaborative inquiry, through age-appropriate, teacher-guided applications. When well designed, AI-supported environments benefit younger learners, particularly within exploration and collaborative tasks aligned with constructivist pedagogies (Cohn et al., 2025; Holmes et al., 2019). Crucially, these competencies are not confined to primary schooling: the same dispositions and skills underpin deep learning in higher education, where students must integrate multiple representations, sustain self-regulated inquiry, and transition from exploratory reasoning to formal, axiomatic justification. Despite the promise of AI in mathematics education, several challenges must be addressed to ensure its responsible and effective implementation. A key concern lies in ethical considerations, particularly around data privacy, algorithmic bias, and transparency. Given the sensitivity of educational data, robust safeguards are essential to ensure ethical protection and use, while poor data quality can compromise AI efficacy through flawed outputs and ineffective personalization, and institutional resistance, including limited infrastructure and low willingness to adopt, remains a persistent barrier (Meylani, 2024). Further concerns arise regarding social interaction limitations. AI-mediated learning environments may reduce meaningful student-teacher and peer-to-peer interactions, potentially impeding the development of social skills and emotional intelligence, competencies essential for holistic education (Gökçearslan et al., 2024; Subiyantoro et al., 2023). A preference for human interaction, particularly in complex mathematical reasoning, may also reduce learner motivation when AI tools are used in isolation. Academic integrity is another pressing concern, as students may misuse chatbots to complete assessments, thereby undermining the authenticity of their learning outcomes (Tang & Chaw, 2024). Finally, privacy and equity issues are central to ongoing debates about AI in education. Questions of consent, data ownership, and bias in algorithmic design raise significant concerns, especially regarding how AI systems might unintentionally disadvantage marginalized student populations (Jain et al., 2024; Wang et al., 2023). In this respect, National Council of Teachers of Mathematics (2024) position is especially pertinent because it explicitly cautions against treating AI as a substitute for the teacher, reinforcing that responsible adoption in mathematics education must preserve teacher agency, disciplinary judgement, and the centrality of students' mathematical reasoning. AI-powered adaptive learning systems and chatbots hold significant potential to transform mathematics education by fostering personalization, metacognitive engagement, and instructional efficiency. However, their integration must be carefully managed to address the ethical, social, and pedagogical complexities they present.

Two complementary approaches illustrate how geometry teaching can be connected to notions of AI through digital tools. Abar and de Almeida (2024) report that using GeoGebra to cultivate computational thinking establishes a conceptual linkage with AI, as described by primary-school teachers. In a different line of work, Dalmon et al. (2010) describe an intelligent tutoring system integrated with iGeom that employs pattern-recognition algorithms and adaptive learning to guide learners at secondary and university levels. Both studies underscore the need for teacher professional development and robust technological infrastructure to implement such strategies effectively. Notably, however, neither line of research addresses the use of tangrams, focusing instead exclusively on digital environments (Abar & de Almeida, 2024).

METHOD

Methodological Options

This pilot study, situated in higher education mathematics, employed a SOFLA-informed hybrid learning design (Marshall & Kostka, 2020) mediated by Moodle and purposefully integrated with tangram, GeoGebra, and Polypad, alongside AI-supported reflective inquiry. Consistent with the nature of the study aims, we adopted a mixed-methods approach. The quantitative strand was designed to estimate the magnitude and precision of learning gains across four geometric domains, GD, GP, MPS, and SR via parallel pre-/post-assessments and effect-size inference. The qualitative strand was designed to characterize the forms and trajectories of students' geometric reasoning, with particular attention to how learners appropriate AI-generated explanations during asynchronous interactions (Moodle forums), thereby explaining how and why change occurs beyond numerical outcomes.

To ensure interpretive clarity, we use the following operational definitions for the analytic dimensions examined in this study:

1. **GD:** part-whole analyses of figures, including composition, tiling, dissection, and reconfiguration, used to reveal structure, establish equivalences, and scaffold generalization (Abrahamson et al., 2023),
2. **GP:** knowledge and principled use of definitions, theorems, and invariants (congruence, symmetry, angle/area relations) to support explanation and proof within Euclidean settings (Santos-Trigo, 2024),
3. **MPS:** the strategic formulation, monitoring, and justification of solution paths for non-routine tasks, encompassing the coordination of representations, heuristics, and warrants for claims (Santos-Trigo, 2024), and
4. **SR:** the capacity to encode, represent, mentally manipulate, and transform spatial information (such as rotation, reflection, translation), integrating visual-motor imagery with rule-governed operations; following cognitive science usage, we understand SR as "the ability to generate, retain, retrieve, and transform well-structured visual images" (Uttal et al., 2013, p. 352).

Pedagogically, although SOFLA comprises eight phases, our implementation operated as a repeated cycle across the semester using a streamlined two-moment sequence:

(1) asynchronous preparation (Moodle activities and AI tools) and

(2) studio-style, in-class sessions for mathematical exploration with physical-digital resources.

The same instructional cycle was applied iteratively to several geometry topics; the present article reports outcomes from one representative cycle.

The objective of this work was to understand how a SOFLA-inspired hybrid design, mediated by Moodle and supported by AI, articulating tangram, GeoGebra, and Polypad, improves performance in geometry (GD, GP, MPS, and SR) and qualifies students' geometric reasoning, by mapping the appropriation of AI-generated explanations in asynchronous interactions.

The RQs are as follows:

1. **RQ1.** To what extent does the hybrid, AI-supported design improve performance across GD, GP, MPS, and SR in higher education students?

2. **RQ2.** In what ways do students articulate and develop their geometrical reasoning in asynchronous forum exchanges, including their appropriation of AI-generated explanations for sense-making?

This design foregrounds complementarity: the quantitative evidence establishes how much change occurred at the domain level, while the qualitative evidence elucidates how reasoning evolved, and which AI-mediated mechanisms supported that evolution, together providing a rigorous, policy-relevant account of impact and process.

Mathematical learning experiences in the primary years are strengthened when designed as hybrid, multimodal sequences integrating physical manipulatives.

Twenty-two students enrolled in a higher professional technical course in educational technology services that provides access to an initial teacher education program at a Portuguese higher education institution participated in this study. The cohort comprised two male and twenty female students, aged 18-22 years. In terms of prior pathways, only two students had progressed through the regular (general) secondary route, whereas the remaining twenty had completed vocational/professional programs. Overall, all participants reported or demonstrated difficulties in mathematics and in geometry, especially SR, geometric construction, and problem-solving. Only one student had prior experience of using GeoGebra at secondary school, and none had previously encountered Polypad or comparable dynamic-geometry environments. This lack of prior exposure underscores the need to integrate technological tools into foundational mathematics preparation for future education professionals. All students had personal access to a digital device (laptop) and were familiar with Moodle-based online learning environments, which facilitated the implementation of the study's asynchronous components.

Research Design and Layout

A distinctive feature of the design was the use of generative AI to scaffold exploration and promote learner autonomy. AI chatbots, such as ChatGPT, Copilot, and DeepSeek, were not treated as simple retrieval engines but as interactive agents to stimulate inquiry, provide access to conceptual knowledge, and encourage reflection, an approach consistent with evidence that AI can support self-regulated learning and student engagement in hybrid environments (Zawacki-Richter et al., 2019). In the reported intervention unfolded across four interconnected stages, each combining digital and physical resources with clear pedagogical aims:

1. **Step 1. Introduction to tangram (asynchronous and pre-class learning):** Students accessed a Moodle forum titled "introduction to tangram". They used AI assistants to respond to four guided prompts about the tangram's origin, geometric composition, mathematical significance, and potential for creative construction. Students created discussion threads to share findings and examples, initiating reflective learning and developing critical digital literacies. This phase aligns with SOFLA's "pre-work" and "sign-in activity" steps, encouraging asynchronous engagement and topic anticipation.
2. **Step 2. Building the tangram in GeoGebra (synchronous, in-class, and individual):** In a classroom session, students used GeoGebra to reconstruct the seven tangram pieces through guided digital constructions. This stage emphasized the application of geometric vocabulary and construction techniques, such as types of triangles, square and parallelogram, enhancing SR and accuracy. The activity corresponds to SOFLA's "whole-class application" and "breakout" elements, combining interactive modeling with individual exploration.
3. **Step 3. Exploring the digital tangram in Polypad (synchronous, in-class, and collaborative):** Working in small groups, students used Polypad to recreate both basic and advanced tangram figures. They explored geometric transformations such as translation, rotation, and reflection. Their creations were documented via screenshots and discussed in class. This task reflects the "share-out" stage of SOFLA, promoting peer explanation, critique, and collaborative knowledge building.
4. **Step 4. Properties of the tangram pieces (asynchronous and reflective closure):** Students used AI chatbots to investigate and document properties of quadrilaterals present in the tangram. They completed a worksheet and shared their insights in a Moodle forum post. This final stage mirrors SOFLA's "reflection" step, fostering consolidation and metacognitive engagement. Taken together, this methodological framework demonstrates pedagogical innovation by intentionally blending physical manipulatives, dynamic digital tools, and AI-driven inquiry; adapting the SOFLA structure to a higher-education context; and fostering autonomous,

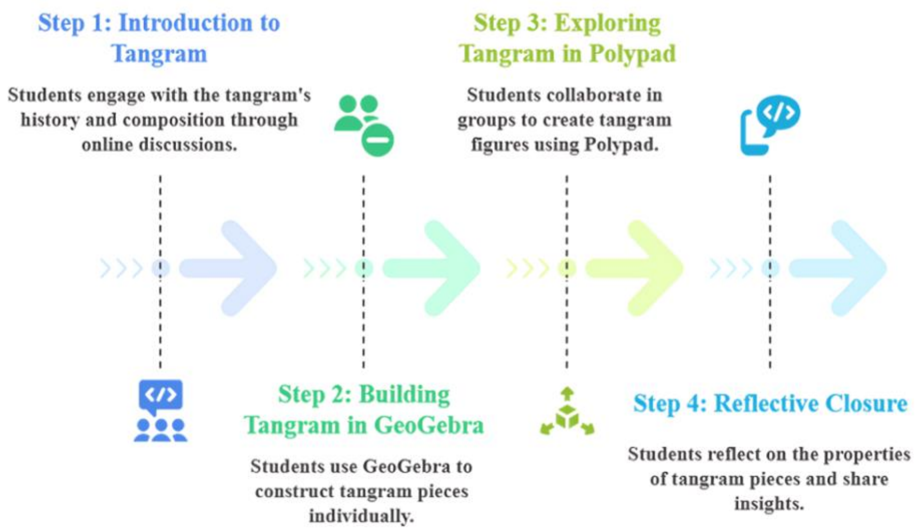

Figure 1. Tangram hybrid learning journey (the authors' own creation)

reflective, and collaborative learning aligned with contemporary digital-competence frameworks (Redecker, 2017).

Figure 1 illustrates the structured pedagogical sequence developed in this study, highlighting the integration of asynchronous and synchronous phases inspired by the SOFLA model.

Data Collection and Analysis Methods

A multi-method approach was employed to collect both quantitative and qualitative data throughout the intervention, ensuring a comprehensive understanding of students' learning processes and experiences. Two parallel assessments were administered immediately before (pre-test) and after (post-test) the intervention. Each comprised eight items, two per analytical dimension aligned with the study objectives:

(1) GD,

(2) GP,

(3) MPS, and

(4) SR.

Items were scored with a criterion-referenced rubric comprising three anchored categories: 0 = incorrect/insufficient evidence; 2 = incomplete or partially correct solution; 2.5 = fully correct solution (no intermediate scores). For each dimension, a domain score was computed as M of its two items at pre- and post-test (range 0.0-2.5). Using two parallel indicators scored on the same rubric reduces item-level noise and yields a more stable estimate of domain-specific competence while preserving the original measurement scale. Parallel forms were designed to follow the same content blueprint and represent comparable difficulty and representation across pre- and post-test. Student responses were pseudonymized as S_i (i = 1, 2, ..., 22) to individual-level tracking of conceptual development over time.

To address the study aims, namely, to characterize changes in geometric competencies and to understand how students articulate and develop their reasoning, the qualitative component captured the mechanisms of learning (e.g., how students used language, coordinated representations, and appropriated AI-generated explanations) through directed content analysis with inductive refinement, thereby illuminating pathways and barriers that numerical scores alone cannot reveal. In parallel, the quantitative component provided estimable evidence of change: domain scores enabled the estimation of magnitude and precision of pre-/post-gains, supported hypothesis testing (paired comparisons with effect sizes and confidence intervals [CIs]), and facilitated the examination of cross-domain relations (e.g., correlations between GD and MPS). Taken together, the two strands offered complementary warrant, qualitative data explaining why/how change occurred, and quantitative data establishing how much change occurred, thus strengthening internal coherence and enhancing the practical interpretability of findings for course design and policy.

During the asynchronous phase, students posted brief written reflections in Moodle. These posts were collated to gauge conceptual engagement and the articulation of mathematical ideas. Qualitative analysis followed Bardin's content analysis framework: pre-analysis, exploration/coding, and treatment of results and inference. Qualitative analysis was operationalized as a directed approach with a priori categories:

(1) clarity and correctness of mathematical language,

(2) depth of geometric reasoning,

(3) integration of insights from AI-generated explanations, and

(4) self-reflection on learning, while permitting inductive subthemes to emerge where warranted.

A codebook was iteratively refined during a pilot calibration; two analysts independently coded all excerpts, disagreements were resolved by discussion, and inter-rater agreement was estimated using Cohen's κ. To enhance trustworthiness, we maintained an audit trail (decisions, memos, code revisions), conducted periodic peer debriefs, and used a structured instructor logbook (session-level entries of pedagogical decisions, observed behaviors, obstacles, and notable learning episodes). These materials supported triangulation and context-sensitive interpretation of learning trajectories and classroom dynamics.

Quantitative data comprised domain scores for GD, GP, MPS, and SR at pre- and post-test. After computing change scores (Δ = post-pre), we examined distributional assumptions (Shapiro-Wilk and Kolmogorov-Smirnov with Lilliefors correction). Primary inferences used paired-samples t-tests with 95% CIs and effect sizes (Cohen's d and Hedges' g where appropriate); Wilcoxon signed-rank tests provided robustness to non-normality. Associations among improvements were explored via Pearson correlations on Δ scores.

Finally, evidence from multiple sources, tests, forum posts, and logbook entries was jointly analyzed and categorized into higher-order dimensions (geometric competencies and student perceptions), which structured the presentation of results and discussion. This convergent design preserved methodological rigor while enabling coherent, evidence-based narrative synthesis.

RESULTS

Geometric Competences

The diagnostic test revealed heterogeneous starting points across all four domains. Difficulties were most pronounced in MPS and SR. Typical errors involved failing to identify component shapes within composite figures and struggling to anticipate transformations, in line with evidence on dynamic-geometry supported reasoning and technology-enhanced problem exploration (Rocha & da Silva, 2021; Santos-Trigo, 2024).

After the intervention, improvements were widespread: SR increased for 18/22 students, MPS for 16/22, and GP for 16/22. GD showed a more mixed pattern, because 14/22 improved, 3/22 were unchanged, and 5/22 declined, indicating that decomposition/abstraction remained challenging for a subset. Overall, the data point to the strongest pre-/post-gains in SR and MPS, with GD persisting as a focal area for targeted support (Meylani, 2024).

Between the diagnostic and final tests (**Table 1**), every student improved in at least one dimension. Gains were most widespread in SR and MPS. GP improved for 15/22 students, whereas GD showed a more varied pattern, 14/22 improved, 3/22 were unchanged, and 5/22 declined, suggesting that decomposition/abstraction remained challenging for a subset. These patterns were confirmed through student-level comparisons (students coded S_i, $i = 1, 2, \ldots, 22$), combining the numerical change scores with qualitative evidence from the study.

The study evaluated the effectiveness of the intervention across four cognitive domains: GD, GP, MPS, and SR. We combined parametric and non-parametric analyses to ensure robustness. Regarding the descriptive statistics, pre-test performance varied across domains, with MPS showing the lowest baseline (M = 0.875, standard deviation [SD] = 0.731) and GD the highest (M = 1.443, SD = 0.756). Post-test scores improved across all domains; the largest gain was in MPS (M = 1.614, SD = 0.560), representing an increase of approximately 84% over baseline (Δ = 0.739).

Assumption checks in **Table 2**, normality of the pre-/post-difference scores was assessed using Shapiro-Wilk and Kolmogorov-Smirnov tests with the Lilliefors correction (two-tailed p values): SR, W = 0.880, p = 0.012;

Table 1. Individual student scores in pre- and post-test by dimension of geometric competence

Student	GD			GP			MPS			SR		
	Pre	Post	Difference	Pre	Post	Difference	Pre	Post	Difference	Pre	Post	Difference
S1	0	2.50	2.50	1	1	0	0	1	1	0	1	1
S2	1	2	1	1	2	1	1	2.25	1.25	1	2.25	1.25
S3	1.25	1	-0.25	1.25	2	0.75	1	1	0	1	1	0
S4	2.25	2.50	0.25	2.50	2	-0.50	1	2	1	1	2	1
S5	0	1	1	2.25	2.50	0.25	0	1	1	0	1	1
S6	2.25	2	-0.25	1	2	1	1	1	0	1	1	0
S7	2	2.50	0.50	2	1.25	-0.75	0	1	1	0	1	1
S8	0	0	0	2.25	2.25	0	2	2.25	0.25	2	2.25	0.25
S9	2	2.25	0.25	1.25	2.50	1.25	0	2.25	2.25	0	2.25	2.25
S10	2.25	1.25	-1	1	1	0	2	2.25	0.25	2	2.25	0.25
S11	2.50	2	-0.50	1	1.25	0.25	2	2	0	2	2	0
S12	1	1.25	0.25	2	2.50	0.50	1	1	0	1	1	0
S13	1	2	1	1.25	2.25	1	1	2	1	1	2	1
S14	2	2.50	0.50	1	1.25	0.25	1	2	1	1	2	1
S15	2	2.50	0.50	2.25	2.25	0	0	1.25	1.25	0	1.25	1.25
S16	1.25	2	0.75	1	1	0	0	1	1	0	1	1
S17	1.25	1.25	0	1	2.50	1.50	2.25	2.25	0	2.25	2.25	0
S18	1.25	2	0.75	1	2.25	1.25	0	1	1	0	1	1
S19	2	2	0	1	2	1	1	2	1	1	2	1
S20	1.25	1	-0.25	1.25	2	0.75	1	1	0	1	1	0
S21	2	2.25	0.25	1.25	2	0.75	1	2	1	1	2	1
S22	1.25	2	0.75	0	1.25	1.25	1	2	1	1	2	1

Table 2. Test of normality

	Kolmogorov-Smirnov^{a}			Shapiro-Wilk		
	Statistics	df	Significance	Statistics	df	Significance
GDDif	.138	22	.200*	.923	22	.086
GPDif	.146	22	.200*	.945	.22	.252
MPSDif	.307	22	.000	.810	22	.001
SRPDif	.188	22	.042	.880	22	.012

Note. *This is a lower bound of the true significance & ^{a}Lilliefors significance correction

GP, W = 0.945, p = 0.252; GD, W = 0.923, p = 0.086; MPS, W = 0.810, p < 0.001. The corresponding Kolmogorov-Smirnov results were SR, D = 0.188, p = 0.044; GP, D = 0.146, p = 0.255; GD, D = 0.138, p = 0.342; MPS, D = 0.307, p = 0.001. Given these modest departures and the sample size (n = 22), we corroborated all inferences with Wilcoxon signed-rank tests, which led to the same substantive conclusions.

Paired samples t-tests revealed statistically significant improvements in all four cognitive domains. MPS demonstrated the most substantial improvement (t[21] = 5.872, p < 0.001, Cohen's d = 1.252, 95% CI [0.66, 1.84]), classified as a large effect according to Cohen's conventions (**Figure 2**). SR (t[21] = 4.169, p < 0.001, d = 0.889, 95% CI [0.37, 1.41]) and GP (t[21] = 4.007, p < 0.001, d = 0.854, 95% CI [0.34, 1.37]) also exhibited very large effect sizes. GD, while showing significant improvement (t[21] = 2.416, p = 0.025, d = 0.515, 95% CI [0.04, 0.99]), demonstrated a large effect size. Non-parametric Wilcoxon signed-rank tests corroborated these findings, with all domains showing significant improvements (p ≤ 0.021), confirming the robustness of results regardless of distributional assumptions.

Individual student analysis revealed heterogeneous response patterns to the intervention. Of the 22 participants, 8 students (36.4%) exhibited negative progress in at least one domain, indicating differential responsiveness to the intervention. The top-performing student (S21) achieved a total improvement score of 4.5 points across all domains, while the most substantial individual domain improvement was observed in MPS (2.25 points for student S9). In **Table 3** we present Pearson correlations among pre-/post-change scores (Δ = post-pre) for SRPDif, GPDif, GDDif, and MPSDif (n = 22).

Correlation analysis of improvement scores revealed a significant positive correlation between GD and MPS improvements (r = 0.518, p = 0.014), suggesting potential cognitive transfer between these domains. This finding aligns with theoretical frameworks proposing shared cognitive mechanisms underlying spatial-geometric reasoning and MPS abilities. Pre-test correlations showed minimal inter-domain relationships, with

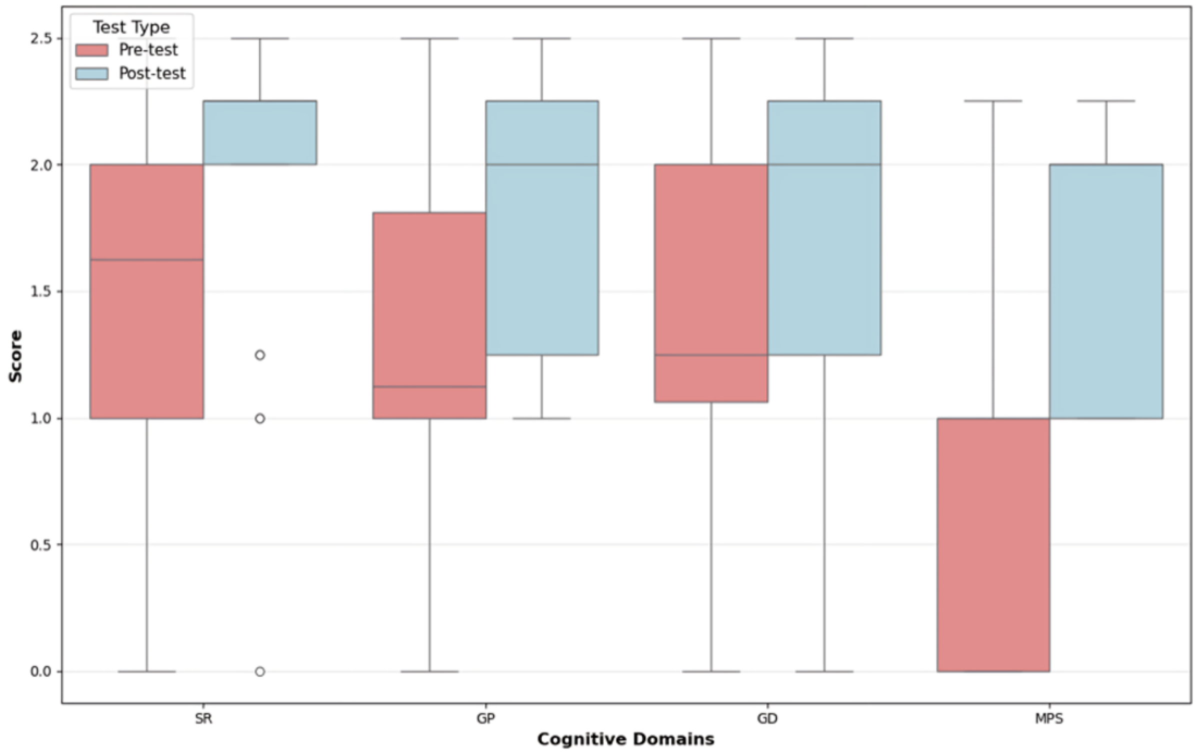

Figure 2. Tests of normality for pre-post difference scores (Δ post-pre): Shapiro-Wilk and Kolmogorov-Smirnov (Lilliefors-corrected) (df = 22) (SPSS)

Table 3. Pearson intercorrelations among change scores (GDDif, GPDif, MPSDif, and SRPDif)

		GDDif	GPDif	MPSDif	SRPDif
GDDif	Pearson correlation	1	-.061	.518*	-.150
	Significance (2-tailed)		.786	.014	.506
	N	22	22	22	22
GPDif	Pearson correlation	-.061	1	.009	-.213
	Significance (2-tailed)	.786		.968	.341
	N	22	22	22	22
MPSDif	Pearson correlation	.518*	.009	1	-.014
	Significance (2-tailed)	.014	.968		.950
	N	22	22	22	22
SRPDif	Pearson correlation	-.150	-.213	-.014	1
	Significance (2-tailed)	.506	.341	.950	
	N	22	22	22	22

Note. *Correlation is significant at the 0.05 level (2-tailed)

the strongest negative correlation observed between GP and SR (r = -0.349). Post-test correlations remained generally weak, except for a moderate negative correlation between GD and GP (r = -0.344), suggesting domain-specific developmental trajectories.

The observed effect sizes exceed conventional thresholds for educational interventions, with three of four domains achieving very large effects (d > 0.8) (Hill et al., 2008). The significant correlation between GD and MPS improvements suggests that interventions targeting spatial-geometric skills may yield broader mathematical benefits, supporting integrated pedagogical approaches (Uttal et al., 2013). The heterogeneity in individual responses indicates the need for personalized intervention strategies and highlights the importance of identifying student characteristics that predict intervention responsiveness. Analysis of forum contributions curiosity, and the ability to navigate between physical, digital, and AI-supported environments.

Students used geometric vocabulary confidently and demonstrated growing SR through shape decomposition and reconstruction. Using AI chatbots was reported as “fun” and “helpful” in student comments.

Table 4. Engagement and motivation–Student quotations from the final focus group

Dimension	Student
Engagement and motivation	“*It was fun trying to guess the shapes by myself*” (S2). “*Using GeoGebra was like … building with puzzle pieces, a bit like a game*” (S5). “*Tangram made geometry lessons feel like a proper puzzle, way more interesting and engaging than usual*” (S3). “*I felt much more motivated to learn geometry when we used the tangram; it made me want to keep trying different arrangements*” (S11). “*Polypad made lessons properly interactive, dragging and rotating pieces was fun and kept me focused*” (S15). “*We had more activities like this … not just numbers but thinking with shapes*” (S16). “*When we used GeoGebra to build shapes, it felt like exploring rather than just copying from the board*” (S17). “*Are we doing one of those different tasks today? They go quicker, miss … it made me want to do next one*” (S18). “*I’d never used GeoGebra or Polypad before, but it was simple and easier to learn … it was cool and fun. Using these tools together made the class feel more modern, and I wasn’t afraid to make mistakes because I could try things quickly*” (S20). “*I liked moving the pieces around more than just … and then it’s easier than writing the answers*” (S21).

Student Perceptions of the Experience

The integration of forum-based assignments within Moodle promoted structured reflection and task ownership. Throughout the lessons, we recorded field notes and documented classroom observations, including spontaneous comments and remarks made by students during activities, as well as feedback accompanying the work they submitted on the Moodle platform supporting the tasks. These qualitative records were then analyzed and categorized thematically. On the basis of this analysis, we organized students’ comments into three categories:

(1) engagement and motivation,

(2) perceived learning gains, and

(3) role of AI, together with illustrative excerpts presented in **Table 4**.

Students’ perceptions seem to indicate a strong preference for geometry instruction that incorporates digital manipulatives and puzzle-based activities, with particular emphasis on engagement and motivation. For instance, one student expressed enjoyment in solving geometric problems independently: “*it was fun trying to guess the shapes by myself*” (S2). This aligns with Papert’s constructionist principles, where learners construct knowledge through active experimentation and play (Margolis, 2020). Similarly, another participant described GeoGebra as “*building with puzzle pieces, a bit like a game*” (S5), reflecting Bruner’s discovery learning theory, which advocates for intuitive, exploratory approaches to mathematical understanding (Tohari & Rahman, 2024). The integration of tangrams and digital tools emerged as a key factor in enhancing student interest in geometry. As one participant noted: “*the tangram made geometry lessons feel like a proper puzzle, way more interesting and engaging than usual*” (S3). This sentiment resonates with Vygotsky’s ZPD. The student further elaborated that “*I felt much more motivated to learn geometry when we used the tangram; it made me want to keep trying different arrangements*” (S11), underscoring the importance of intrinsic motivation in mathematical learning. The interactive nature of Polypad was similarly valued: “*polypad made the lessons properly interactive, dragging and rotating pieces was fun and kept me focused*” (S15). This aligns with recent findings by Rocha and Silva (2021), who found that GeoGebra stimulates student interest in geometric transformations through its dynamic, hands-on interface. The same student added: “*I wish we had more activities like this … not just numbers but thinking with shapes*” (S16), suggesting a preference for visual and SR over abstract algebraic manipulation. Another participant reflected on the cognitive and emotional benefits of GeoGebra: “*when we used GeoGebra to build shapes, it felt like exploring rather than just copying from the board*” (S17). This perspective supports Costa’s (2023) application of the theory of conceptual fields, which highlights the role of concrete, manipulative experiences in facilitating conceptual understanding. The student continued: “*are we doing one of those different tasks today? They go quicker, miss … it made me want to do the next one*” (S18), illustrating the self-perpetuating cycle of engagement and motivation that arises from successful problem-solving.

Table 5 presents student quotations attesting to perceived learning gains within the hybrid, SOFLA-informed geometry design.

Table 5. Perceived learning gains–Student quotations from the final focus group

Dimension	Student
Perceived learning gains	*"I came away with a better understanding of geometric ideas because of the tangram, seeing the pieces fit helped it click"* (S2). *"The interactive tangram helped me solve problems faster; once I could move the parts around, the steps made sense"* (S5). *"It's easier like this to turn shapes and flip them, and to see different positions and possibilities … I didn't know that before. It helped me understand shapes and their properties, like how triangles combine to make other figures."* (S7). *"I'd say the tangram is a genuinely useful tool for learning geometry, not just something to play with"* (S9). *"Now I can make patterns … like, three or four different ways. My understanding of geometric figures improved with these activities; I could finally explain why"* (S10). *"I used to think tangrams were just toys, but they're maths too"* (S11). *"Now with Polypad, I saw how symmetry works … it's much easier when you can see it moving"* (S14). *"GeoGebra let me visualize and construct things dynamically, so I could test ideas instead of guessing"* (S18).

Table 6. Role of AI–Student quotations from the final focus group

Dimension	Student
Role of AI	*"With ChatGPT it's much easier to ask what a trapezium is … and complete the properties. It explains with drawings, and you can understand it"* (S5). *"Just put it in the chat and it explains it to you in a different way"* (S8). *"Sure, it's been much better, AI tools helped me clear up doubts straight away, almost like having a tutor there to check if I was on the right track"* (S10). *"I asked ChatGPT if I was right … and it gave me other ways to think, like rotating the pieces and that. It's easy to apply the concepts when we have an AI chat to talk and explain. If I get stuck, I ask for a hint and try again. It's not just copying, is we learn better"* (S11). *"This way it worked well, GeoGebra and Polypad with AI worked well together, I can build a shape and then ask the AI to explain the rules behind it"* (S12). *"I didn't really understand what it said … so I copied what I wrote in Gemini, pasted it into Copilot, and it used simpler words"* (S15). *"It's easy to apply the concepts when we have an AI chat to talk and explain. If I get stuck, I ask for a hint and try again. It's not just copying, is we learn better"* (S17). *"It's a lot easier to measure the sides using GeoGebra"* (S19). *"If you rotate or move them, the shapes turn into others"* (S20). *"You can solve things in lots of ways … there isn't just one right answer"* (S22).

Novelty and accessibility were also key themes. One student noted: "*I'd never used GeoGebra or Polypad before, but it was simple and easier to learn … it was cool and fun*" (S20). This reflects the potential of digital tools to lower barriers to experimentation, as learners can iteratively test hypotheses without fear of failure (Kampff et al., 2004). Finally, the tactile and kinesthetic aspects of the activities were particularly appealing: "*I liked moving the pieces around more than just … and then it's easier than writing the answers*" (S21). This sentiment aligns with Brum et al.'s (2023) findings on the cognitive benefits of manipulatives, which promote mental representation through physical interaction.

The students' testimonies cohere with evidence that hybrid, multimodal sequences, weaving physical manipulatives with dynamic digital representations, enable transduction across representations, strengthen GD and SR, and can lower extraneous cognitive load (Abrahamson et al., 2023; Rocha & da Silva, 2021). Thus, "*I came away with a better understanding … seeing the pieces fit helped it click*" (S2) and faster problem-solving with interactive tangram (S5) reflect embodied action-perception loops; ease of rotating/flipping shapes and explaining patterns (S7, S9, S10, S11) aligns with guided assistance/ZPD and fading scaffolding (Rebolledo-Mendez et al., 2022). Dynamic environments make invariants visible, "*with Polypad, I saw how symmetry works*" (S14), and support hypothesis testing in GeoGebra (S18), consistent with reviews of dynamic geometry's efficacy (Rocha & da Silva, 2021). Within teacher-led orchestration (TPACK), AI can supply criterion-referenced feedback and metacognitive prompts without displacing professional judgement, consolidating these perceived gains (Da Costa & Prado, 2015; Holmes et al., 2019).

Table 6 presents student quotations illustrating the role of AI as an explanatory and scaffolding partner within the multimodal workflow.

The testimonies depict AI as a formative scaffold that supplies on-demand, criterion-referenced explanations and metacognitive prompting: asking “what a trapezium is … and complete the properties” with drawings (S5), receiving an alternative explanation “in a different way” (S8), and tutor-like checking “if I was on the right track” (S10). Synergy with dynamic geometry is salient: “GeoGebra and Polypad with AI worked well together … build a shape and then ask the AI to explain the rules” (S12); measuring and transforming figures (S19, S20) and recognizing multiple solution paths (S22) exemplify transduction across representations and hypothesis testing (Abrahamson et al., 2023; Rocha & da Silva, 2021). Multi-agent paraphrasing, “copied … into Copilot, and it used simpler words” (S15), signals accessibility and language-level adaptation. Situated within teacher-led orchestration (Clark-Wilson et al., 2020) and primary-phase feasibility (Holmes et al., 2019), these uses advance conceptual clarity without displacing professional judgement, while underscoring the need for transparency, privacy-by-default, and bias auditing in AI-enhanced classrooms (Jain et al., 2024; Meylani, 2024; Wang et al., 2023).

DISCUSSION

This pilot study suggests, rather than demonstrates, illustrates the educational potential of combining physical manipulatives, dynamic geometry tools, and AI-assisted inquiry within a hybrid learning environment mediated by Moodle. The integration of these multimodal resources aligns with constructivist and constructionist principles, which advocate experiential, exploratory, and technology-enhanced learning (Brum et al., 2023; Kampff et al., 2004; Santos-Trigo, 2024). Such hybrid configurations reflect the growing relevance of blended pedagogies in mathematics education, where digital and physical modalities converge to support conceptual development and learner autonomy (Lieban, 2023).

Concerning the extent of improvement across GD, GP, MPS, and SR, pre-/post-analyses revealed statistically significant gains in all four domains. These results were robust across parametric and non-parametric tests. Effect sizes ranged from large to very large, with upward distributional shifts observed in every outcome. At the individual level, all students improved in at least one domain, with the most widespread gains occurring in SR (18/22) and MPS (16/22). GP improved for 15 students. GD presented a more heterogeneous pattern, with 14 students improving, 3 remaining unchanged, and 5 declining. Importantly, 8 students (36.4%) exhibited negative progress in at least one domain, a proportion that is substantial given the small sample and that tempers any straightforward positive interpretation of the intervention. This pattern suggests that, although the overall tendency was favorable, the intervention did not benefit all learners consistently and may have posed difficulties for a non-trivial subgroup, particularly in tasks requiring decomposition, abstraction, and sustained geometric description. A positive correlation between gains in GD and MPS ($r \approx 0.52$) indicates potential transfer effects from spatial-geometric skills to broader problem-solving competencies. These findings resonate with prior empirical research demonstrating the cognitive benefits of concrete and digital manipulatives in mathematics learning. Costa (2023), drawing on the theory of conceptual fields, showed the efficacy of concrete materials in teaching signed numbers. Brum et al. (2023) highlighted the value of manipulatives grounded in the theories of Bruner, Ausubel, Piaget, and Vygotsky, reinforcing the notion that physical interaction promotes abstraction and mental representation. The use of tangrams and GeoGebra has also shown promise in enhancing spatial perception and geometric understanding (Rocha & da Silva, 2021). Moreover, Herrera et al. (2024) demonstrated that virtual environments and 3D printing significantly improved spatial visualization skills, with a 25% increase among students exposed to these tools.

Which examined how students articulate and develop reasoning in forums and how they appropriate AI-generated explanations, a directed content analysis of asynchronous posts revealed notable qualitative shifts. Students increasingly named and justified mathematical properties, described transformations with greater linguistic precision, and utilized AI-generated explanations as scaffolds to verify reasoning, explore alternative representations, and correct misconceptions. AI functioned not merely as a retrieval engine but as a promotable companion, supporting sense-making while preserving the integrity of students’ own argumentation. This dynamic use of AI aligns with recent literature on adaptive learning systems and AI-enhanced instruction. Dabingaya (2022) and Mustafa (2024) report that AI-based platforms facilitate personalized learning pathways, offering real-time feedback and improving comprehension of complex mathematical concepts. Meylani (2024) further underscores the scalability and cost-effectiveness of such systems, which can support large student cohorts without compromising instructional quality. The integration of AI chatbots has also been shown to enhance student engagement, inclusivity, and responsiveness (A'ini & Khoiriyah, 2024; Boltayevich et al., 2024; Roca et al., 2024).

The pedagogical implications of this pilot are supported by the TPACK framework (Clark-Wilson et al., 2025), which emphasizes the intersection of technological fluency, pedagogical strategy, and content mastery. However, as noted by Tohari and Rahman (2024), constructivist approaches such as ZPD and discovery learning require substantial investment in time, resources, and teacher training. Individual variability among learners also necessitates differentiated strategies to ensure equitable access to meaningful learning experiences. This pilot contributes to the growing body of evidence supporting the integration of physical manipulatives, dynamic geometry tools, and AI-assisted inquiry in hybrid mathematics education. However, this contribution should be interpreted cautiously: the statistically significant gains at group level coexisted with heterogeneous individual trajectories, including declines in at least one domain for over one third of participants. Accordingly, the present results are pointless to a uniformly effective model than to a potentially promising approach whose benefits, conditions of success, and risks of uneven impact require further investigation. Rather than evidencing transformative effects, the study indicates that multimodal and adaptive learning environments may support geometric reasoning for many students, while simultaneously revealing the need for targeted pedagogical support for those who do not respond positively to the design. These findings advocate for continued investment in teacher training, structured planning, and intentional technology integration together with closer monitoring of learner variability and mechanisms for early pedagogical intervention.

This study has several limitations that temper the strength and scope of its conclusions. First, the absence of a control group and the single-group pre-/post-design precludes strong causal inference and do not rule out alternative explanations such as maturation, history, or practice effects. Secondly, the sample was relatively small (n = 22) which constrains statistical power, widens CIs and limits generalizability, the lack of demographic information further prevents subgroup or equity-sensitive analyses. These limitations are especially important in light of the finding that 8 students showed negative progress in at least one domain, since in a small sample even a modest number of adverse trajectories becomes educationally consequential and cannot be dismissed as noise. Finally, without longer-term follow-up, no claims can be made about retention or transfer of learning. Future research should continue to explore how these tools can be designed and implemented to support human-centered, inclusive, and developmentally appropriate learning environments. Future studies on the involvement of prospective elementary school teachers in exploratory/investigative geometry with manipulatives, dynamic geometry, and generative AI should use stratified randomized trials with active control, controlling confounders, and, if possible, longitudinal protocols, multi-school samples sized a priori; validated measures, complemented by qualitative and process data. Particular attention should be given to identifying which learners benefit, which learners struggle, and which design features may account for negative or uneven progress across domains.

Author contributions: SC: conceptualization, data curation, formal analysis, investigation, methodology, writing–original draft; and writing–review & editing & **FV:** conceptualization, methodology, supervision, validation, and writing–review & editing. Both authors approved the final version of the article.

Funding: The authors received no financial support for the research and/or authorship of this article.

Acknowledgments: The authors would like to thank the students who participated in this study and share their reflections.

Ethics declaration: All participants were informed about the purpose of the study and their voluntary participation. The data reported in the manuscript were treated confidentially and presented in anonymized form. No personally identifiable information is included.

AI statement: During the preparation of this work, the authors used ChatGPT to improve language and readability. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the content of the published article.

Declaration of interest: The authors declared no competing interests.

Data availability: Data generated or analyzed during this study are available from the authors on request.

REFERENCES

A'ini, Q., & Khoiriyah, R. (2024). Merevolusi pendidikan dengan kecerdasan buatan chatbots: Meningkatkan pembelajaran dan penilaian [Revolutionizing education with artificial intelligence chatbots: Improving learning and assessment]. *Jurnal Multidisiplin Ibrahimy, 2*(1), 54-71. https://doi.org/10.35316/jummy.v2i1.5510

Abar, C. A. A. P., & de Almeida, M. V. (2024). Contributos do GeoGebra para exploração do pensamento computacional no contexto da geometria [Contributions of GeoGebra to the exploration of computational thinking in the context of geometry]. *REMATEC, 19*(48), Article e2024003. https://doi.org/10.37084/REMATEC.1980-3141.2024.n48.e2024003.id590

Abrahamson, D., Tancredi, S., Chen, R. S. Y., Flood, V. J., & Dutton, E. (2023). Embodied design of digital resources for mathematics education: Theory, methodology, and framework of a pedagogical research program. In B. Pepin, G. Gueudet, & J. Choppin (Eds.), *Handbook of digital resources in mathematics education* (pp. 1-34). Springer. https://doi.org/10.1007/978-3-030-95060-6_8-1

Al Hakim, V. G., Meng-Heng, L. I. N., Chang, C. Y., Jen-Hang, W. A. N. G., Chang, C. K., Zhuang, Y., & Su-Hang, Y. A. N. G. (2024). Marrying physical and virtual realms: An embodied, multi-modal approach to situational learning in digital reality. *The 32nd International Conference on Computers in Education*, *2024*, 1-11. https://doi.org/10.58459/icce.2024.4882

Barbosa, A., & Vale, I. (2025). Rebuilding manipulatives through digital making in teacher education. *STEM Education, 5*(4), 515-545. https://doi.org/10.3934/steme.2025025

Boltayevich, E. B., Ergashevna, Y. N., Ganiyevna, A. N., Masharipovich, M. M., Zaynobidinovna, K. M., & Batyrovna, U. M. (2024). LMS information system development concept and principles. *International Conference of Computer and Informatics Engineering, 2024,* 1-7. https://doi.org/10.1109/IC2IE63342.2024.10748038

Brum, E. D. C. M., Viera, M. A., & Ferreira, R. K. A. (2023). Aprendizagem significativa em matemática por meio da utilização de materiais concretos no ensino médio: Um ensaio em construção [Meaningful learning in mathematics through the use of concrete materials in high school: An essay in progress]. *Revista Ibero-Americana de Humanidades, Ciências e Educação, 9*(3), 365-380. https://doi.org/10.51891/rease.v9i3.8794

Clark-Wilson, A., Robutti, O., & Thomas, M. (2020). Teaching with digital technology. *ZDM Mathematics Education, 52*, 1223-1242. https://doi.org/10.1007/s11858-020-01196-0

Cohn, C., Snyder, C., Fonteles, J. H., TS, A., Montenegro, J., & Biswas, G. (2025). A multimodal approach to support teacher, researcher and AI collaboration in STEM+C learning environments. *British Journal of Educational Technology, 56*(2), 595-620. https://doi.org/10.1111/bjet.13518

Costa, J. F. S. (2023). Aprendendo números inteiros: Uma experiência baseada na teoria dos campos conceituais [Learning integers: An experience based on the theory of conceptual fields]. *Revista de Educação Matemática, 20*(1), Article e023104. https://doi.org/10.37001/remat25269062v20id483

Da Costa, N. M. L., & Prado, M. E. B. B. (2015). A integração das tecnologias digitais ao ensino de matemática: Desafio constante no cotidiano escolar do professor [The integration of digital technologies into mathematics teaching: A constant challenge in the teacher's daily school life]. *Perspectivas da Educação Matemática, 8*(16). https://periodicos.ufms.br/index.php/pedmat/article/view/1392

Dabingaya, M. (2022). Analyzing the effectiveness of AI-powered adaptive learning platforms in mathematics education. *Interdisciplinary Journal Papier Human Review, 3*(1), 1-7. https://doi.org/10.47667/ijphr.v3i1.226

Dalmon, D. L., Isotani, S., & de Oliveira Brandão, L. (2010). Melhorando a geometria interativa com o uso de tutores rastreadores de padrões: iGeom e CTAT [Improving interactive geometry with the use of pattern-tracing tutors: iGeom and CTAT]. Anais do XVI Workshop de Informática na Escola, *2020*, 1413-1416. https://doi.org/10.5753/wie.2010.25420

Gökçearslan, S., Tosun, C., & Erdemir, Z. G. (2024). Benefits, challenges, and methods of artificial intelligence (AI) chatbots in education: A systematic literature review. *International Journal of Technology in Education, 7*(1), 19-39. https://doi.org/10.46328/ijte.600

Herrera, L. M., Juárez Ordóñez, S., & Ruiz-Loza, S. (2024). Enhancing mathematical education with spatial visualization tools. *Frontiers in Education, 9*, Article 1229126. https://doi.org/10.3389/feduc.2024.1229126

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives, 2*(3), 172-177. https://doi.org/10.1111/j.1750-8606.2008.00061.x

Hoffmann, D. S., Martins, E. F., & Basso, M. V. D. A. (2009). *Experiências física e lógico-matemática em espaço e forma: Uma arquitetura pedagógica de uso integrado de recursos manipulativos digitais e não-digitais* [*Physical and logical-mathematical experiences in space and form: A pedagogical architecture for the integrated use of digital and non-digital manipulative resources*]. Anais Do Simposio Brasileiro De Informatica *Na Educacao*. https://bitly.cx/bkfv0

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education promises and implications for teaching and learning*. Center for Curriculum Redesign. https://curriculumredesign.org/wp-content/uploads/AIED-Book-Excerpt-CCR.pdf

Jain, V., Singh, I., Syed, M., Mondal, S., & Palai, D. R. (2024). Enhancing educational interactions: A comprehensive review of AI chatbots in learning environments. *11th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)*, 2024, 1-5. IEEE. https://doi.org/10.1109/ICRITO61523.2024.10522392

Janisch, A. B. L., & Jelinek, K. R. (2023). Atividades experimentais no ensino da matemática mediadas pela prática docente e nas perspectivas vygotskiana, piagetiana e ausubeliana [Experimental activities in mathematics teaching mediated by teaching practice and from Vygotskian, Piagetian, and Ausubelian perspectives]. *Revista de Iniciação à Docência, 8*(1), Article e12608. https://doi.org/10.22481/riduesb.v8i1.12608

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux. https://www.bibsonomy.org/bibtex/a1400a299a00de009ec8eda73e6289af

Kampff, A. J. C., Machado, J. C., & Cavedini, P. (2004). Novas tecnologias e educação matemática [New technologies and mathematics education]. *Revista Novas Tecnologias na Educação, 2*(2), 1-11. https://doi.org/10.22456/1679-1916.13703

Leikin, R., & Guberman, R. (2023). Creativity and challenge: Task complexity as a function of insight and multiplicity of solutions. R. Leikin (Ed.), *Mathematical challenges for all* (pp. 325-342). https://doi.org/10.1007/978-3-031-18868-8

Lieban, D. (2023). Entre o digital e o físico: Integrando recursos com o GeoGebra para práticas criativas em espaços de aprendizagem [Between the digital and the physical: Integrating resources with GeoGebra for creative practices in learning spaces]. *Revista do Instituto GeoGebra Internacional de São Paulo, 12*(2), 67-88. https://doi.org/10.23925/2237-9657.2023.v12i2p067-088

Margolis, J. (2020). Multiple team membership: An integrative review. *Small Group Research, 51*(1), 48-86. https://doi.org/10.1177/1046496419883702

Marshall, H. W., & Kostka, I. (2020). Fostering teaching presence through the synchronous online flipped learning approach. *TESL-EJ, 24*(2), 1-14. https://eric.ed.gov/?id=EJ1268565

Meylani, R. (2024). A critical glance at adaptive learning systems using artificial intelligence: A systematic review and qualitative synthesis of contemporary research literature. *Batı Anadolu Eğitim Bilimleri Dergisi, 15*(3), 3519-3547. https://doi.org/10.51460/baebd.1525452

Mustafa, A. N. (2024). The future of mathematics education: Adaptive learning technologies and artificial intelligence. *International Journal of Science and Research Archive, 12*(1), 2594-2599. https://doi.org/10.30574/ijsra.2024.12.1.1134

National Council of Teachers of Mathematics. (2024). Artificial intelligence and mathematics teaching: A position of the National Council of Teachers of Mathematics. *National Council of Teachers of Mathematics*. https://www.nctm.org/standards-and-positions/Position-Statements/Artificial-Intelligence-and-Mathematics-Teaching/

Neha, F., & Bhati, D. (2025). *DeepSeek models in STEM education: Capabilities, applications, and challenges*. TechRxiv. https://doi.org/10.36227/techrxiv.174198447.71647707/v1

Rafee, B. M., Ramesh, V., Zaheed, S. M. & Khatoon, S. (2024). Can hybrid learning change education? *International Journal of Social Science, 1*(2), 1-16. https://nsfjournals.com/article-files/pdf/h3UiHSGPbITDkMqnRa3A-1715971162.pdf

Rebolledo-Mendez, G., Huerta-Pacheco, N. S., Baker, R. S., & du Boulay, B. (2022). Meta-affective behaviour within an intelligent tutoring system for mathematics. *International Journal of Artificial Intelligence in Education, 32*(1), 174-195. https://doi.org/10.1007/s40593-021-00247-1

Redecker, C. (2017). *European framework for the digital competence of educators: Digcompedu*. Publications Office of the European Union. https://doi.org/10.2760/159770

Roca, M. D. L., Chan, M. M., Garcia-Cabot, A., Garcia-Lopez, E., & Amado-Salvatierra, H. (2024). The impact of a chatbot working as an assistant in a course for supporting student learning and engagement. *Computer Applications in Engineering Education, 32*(5), Article e22750. https://doi.org/10.1002/cae.22750

Rocha, R. P., & da Silva, M. D. F. (2021). Uma revisão sistemática abordando o tangram, o GeoGebra e as opções de isometria do plano [A systematic review addressing the tangram, in GeoGebra, and the options for isometric plane calculations]. *Educação Matemática Pesquisa Revista do Programa de Estudos Pós-Graduados em Educação Matemática, 23*(1), 741-768. https://doi.org/10.23925/1983-3156.2021v23i1p741-768

Sain, Z. H., Vasudevan, A., & Lama, A. V. (2024). The emerging future of AI chatbots in higher education. *Jurnal Ilmiah Didaktika, 25*(1), 93-107. https://doi.org/10.22373/jid.v25i1.25583

Santos-Trigo, M. (2024). Problem solving in mathematics education: Tracing its foundations and current research-practice trends. *ZDM Mathematics Education, 56*, 211-222. https://doi.org/10.1007/s11858-024-01578-8

Skill, T. D., & Young, B. A. (2002). Embracing the hybrid model: Working at the intersections of virtual and physical learning spaces. *New Directions for Teaching and Learning, 2002*(92), 23-32. https://doi.org/10.1002/tl.76

Subiyantoro, S., Degeng, I. N. S., Kuswandi, D., & Ulfa, S. (2023). Exploring the impact of AI-powered chatbots (ChatGPT) on education: A qualitative study on benefits and drawbacks. *Jurnal Pekommas, 8*(2), 157-168. https://www.semanticscholar.org/paper/Exploring-the-Impact-of-AI-Powered-Chatbots-(Chat-A-Subiyantoro-Degeng/b05f684a379727a47e279e12f23b1ae4308f3ae7

Tang, C. M., & Chaw, L. Y. (2024). Student learning performance evaluation: Mitigating the challenges of generative AI chatbot misuse in student assessments. *Proceedings of the European Conference on E-Learning, 23*(1), 357-364. https://doi.org/10.34190/ecel.23.1.2567

Tohari, B., & Rahman, A. (2024). Konstruktivisme Lev Semonovich Vygotsky dan Jerome Bruner: Model pembelajaran aktif dalam pengembangan kemampuan kognitif anak [Constructivism of Lev Semonovich Vygotsky and Jerome Bruner: An active learning model in the development of children's cognitive abilities]. *Nusantara: Jurnal Pendidikan Indonesia, 4*(1), 209-228. https://doi.org/10.14421/njpi.2024.v4i1-13

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, T., Lund, B. D., Marengo, A., Pagano, A., Mannuru, N. R., Teel, Z. A., & Pange, J. (2023). Exploring the potential impact of artificial intelligence (AI) on international students in higher education: Generative AI, chatbots, analytics, and international student success. *Applied Sciences, 13*(11), Article 6716. https://doi.org/10.3390/app13116716

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–Where are the educators? *International Journal of Educational Technology in Higher Education, 16*, Article 39. https://doi.org/10.1186/s41239-019-0171-0

◆❖◆