AI-assisted Socratic Questioning for Formal Proof Competence in Plane Geometry: A Literature Review

I Putu Wisna Ariawan*, I Made Candiasa, Ni Made Sri Mertasari, Dewa Gede Hendra Divayana, I Made Sugiarta

Universitas Pendidikan Ganesha, Bali, Indonesia
*Corresponding author, email: wisna.ariawan@undiksha.ac.id

https://doi.org/10.17977/um065.v7.i4.2027.11

Article history
Submitted: 15 August 2026
Revised: 18 September 2026
Accepted: 10 October 2026
Published: 12 October 2026

Keywords
AI-assisted Socratic questioning
Formal proof
Mathematics education
Plane geometry

Abstract
Formal proof is at the heart of university-level plane geometry because it requires students to build deductively valid arguments based on definitions, theorems, and known results, rather than visual observation or intuition. However, students often struggle to justify inferential steps, while generative artificial intelligence can provide fluent proofs before learners examine diagrams, check theorem conditions, or construct their own reasoning. This conceptual article proposes a Warrant-Seeking Socratic AI Framework for supporting formal proof learning in plane geometry. Literature was searched in Scopus, Web of Science, ERIC, and Google Scholar from March 15 to May 21, 2026, using terms related to formal proof, geometry, Socratic questioning, and artificial intelligence. Of 72 records identified, 14 duplicates were removed, leaving 58 sources for screening, full-text assessment, and narrative synthesis. Sources were synthesized by identifying recurring pedagogical and epistemic functions related to proof construction, questioning, AI interaction, and validation. Based on the synthesis: problem orientation, conjecture articulation, theorem-condition analysis, inferential construction, proof validation, and metacognitive reflection, a six-phase recursive model is proposed. The framework positions AI as an epistemic interviewer rather than a proof generator, prompting students to distinguish visual conjectures from warranted claims, examine theorem conditions, justify inferential steps, and retain ownership of their proofs. The synthesis distinguishes evidence directly related to geometry and mathematical proving from findings extrapolated from broader educational contexts. The framework is therefore a theoretically grounded proposition rather than an empirically validated model, and its effectiveness remains to be established through future empirical research.

1. Introduction

The background of this study lies in the fundamental role of formal proof in geometry and the persistent difficulty students experience in moving from intuitive and visual reasoning to structured deductive reasoning. Formal proof is the basic element in geometry, because it is the main procedure for discovering mathematical truth by productive reasoning not only by intuition (Rizos & Gkrekas, 2023). In modern geometry learning, truth is not based on the appearance of space but on the deductive-axiomatic system that requires logical precision and valid arguments (Sukestiyarno et al., 2023). The existence of proof in geometry is the highest level of mathematical reasoning, far above visual observation (Suryanto et al., 2022). This is because mathematical proof is the "core" or heart of the mathematical thinking process that enables learners to logically relate definitions and theorems, the truth of which can be universally accounted for (Tristanti & Iffah, 2022). The development of this proof ability is a fundamental goal in the education of geometry because it trains students to shift from intuitive beliefs to structured deductive understanding (Dewi & Dasari, 2023). Based on the recent study, this transition from visual recognition to formal justification is a hard task for students because the construction of a proof involves the selection of valid definitions and theorems, the sequencing of arguments in a logical order and the connection of visual representations to deductive structures (Aisyah et al., 2023; Ariawan et al., 2026; Maifa et al., 2025). Therefore, in the context of modern geometry, proof is not merely the verification of visual correctness but also a structured mathematical reasoning that determines the truth of a geometric statement (Suryanto et al., 2022; Tristanti & Iffah, 2022).

The rationale of this study is based on the need for an instructional approach that actively engages students in examining assumptions, validating mathematical claims, and constructing logically justified arguments rather than merely reproducing existing proofs. In plane geometry, students' difficulties in proof construction are not merely caused by their inability to remember theorems, but by their limited ability to determine when a theorem

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

is applicable, identify the given facts and the target conclusion, and justify each statement by valid definitions, axioms, or theorems (Anwar et al., 2023). Studies on geometry proof also show that students may make incorrect assumptions about congruent triangles, fail to state sufficient reasons for their claims, or draw conclusions without a logically complete argument (Aisyah et al., 2023). Therefore, the proof instruction should not only rely on the worked examples to be imitated but should teach students to ask what is known, what still has to be proved, why each statement is valid, and whether the theorem used really fits the geometrical configuration (Maifa et al., 2025; Suryanto et al., 2022). Such difficulties provide a pedagogical motivation for an instructional approach that engages students actively in examining assumptions, validating mathematical claims, and constructing rather than reproducing deductive arguments.

The literature review of this study is based on two complementary research streams: first, socratic questioning as pedagogical scaffold for mathematical reasoning. Second, artificial intelligence as interactive learning aid that could facilitate learning of proof. Generative chatbots have created a new instructional space for proof learning in plane geometry, as students can get immediate explanations, examples, hints, and even full proofs; but the fluency of such responses might hide invalid reasoning and therefore they need to be carefully verified, especially when dealing with geometry (Wardat et al., 2023). The main pedagogical question is not whether the tool can generate a proof but how to place it so that students are the authors and validators of their own reasoning. The Socratic questioning method has been found to be useful in mathematics classrooms for clarifying ideas, testing assumptions, developing reflective reasoning and consolidating concepts. Moreover, recent research has indicated that questioning strategies do not automatically enhance students' higher-order proof construction unless they are systematically linked to the structure of mathematical argumentation (Bimantara et al., 2025; Rahmawati et al., 2022). Meanwhile, research on the implementation of artificial intelligence for the learning of mathematics provides promising opportunities for dialogic feedback, mediated exploration, and help with proofs, but also warns that these systems may provide wrong explanations, neglect the geometric representations, or encourage passive acceptance where the students do not have criteria to validate mathematical assertions (Pepin et al., 2025; Wardat et al., 2023; Yoon et al., 2024).

Existing studies address important parts of this problem, but from different instructional and epistemic perspectives. Dilling and Herrmann (2024), for example, examined how pre-service teachers interacted with ChatGPT while working on geometric proofs and found considerable variation in the ways students sought proofs, explanations, or feedback. Their study provides important evidence about actual student–AI interaction in geometric proving, but it does not organize that interaction around a deliberately constrained Socratic sequence in which AI is prevented from taking over the proof. Yoon et al. (2024) similarly examined undergraduate students' use of generative AI for proving mathematical statements and developed the SIPE-AI framework to describe students' proving processes and the factors shaping their engagement with AI. Their contribution clarifies how students negotiate AI-generated mathematical arguments, but it does not propose a geometry-specific questioning architecture to guide students through the epistemic demands of proof construction.

Research on Socratic dialogue provides one more piece of the picture. Socratic questioning has been employed to promote explanation, assumption checking, and critical reasoning in mathematics education (Bimantara et al., 2025; Rahmawati et al., 2022), and recent work has explored Socratic dialog via ChatGPT in other disciplinary contexts (Santos-Guevara et al., 2026). Lu and Hu (2025), from a broader human–AI perspective, Lu and Hu (2025) further argue that Socratic interaction may help preserve user autonomy and reduce cognitive deskilling. These studies establish the pedagogical plausibility of dialogic AI, but they do not address several demands that are distinctive to formal proof in plane geometry: separating what is visually suggested from what is mathematically warranted, checking whether all conditions of a theorem have been established, constructing intermediate claims, and validating the inferential relation between successive statements.

The contribution of the present review therefore lies not simply in bringing together three previously separate bodies of literature. More specifically, it translates findings from research on geometric proving, Socratic questioning, and generative AI into a proof-oriented interaction framework in which AI questions are tied to identifiable acts of mathematical proving. The proposed framework organizes these acts into six recursive phases problem orientation, conjecture articulation, theorem-condition analysis, inferential construction, proof validation, and metacognitive reflection and complements them with a taxonomy of questions directed at definitions, diagrammatic assumptions, theorem conditions, intermediate claims, inferential warrants, formalization, and reflection. In this formulation, AI is deliberately positioned as a question-generating scaffold rather than as a source of mathematical authority. The purpose of the study is therefore to develop a traceable conceptual synthesis of how such AI-assisted Socratic questioning could be structured for formal proof learning in plane geometry, while identifying the empirical questions that remain to be tested.

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

2. Method

This study is theoretical and based on conceptual synthesis by adopting a narrative literature review design to develop a conceptual explanation of how artificial intelligence-assisted Socratic questioning can support students' formal proof abilities in plane geometry (Awang et al., 2025; Chigbu et al., 2023; Lessing & Ogbonnaya, 2026; Lu & Hu, 2025; Rahmawati et al., 2022; Sukhera, 2022a, 2022b; Turnbull et al., 2023). This design is appropriate because this study does not aim to test the effectiveness of interventions using quantitative data, but to synthesize, interpret the literature and build a coherent academic argument about geometric proof, deductive reasoning, Socratic pedagogy, artificial intelligence in mathematics education, and responsible use of technology (Bozer Saraç & Korkmaz, 2025; Chigbu et al., 2023; Fock & Siller, 2025; Lessing & Ogbonnaya, 2026; Lu & Hu, 2025; Mohamed et al., 2022; Opesemowo & Adewuyi, 2024; Turnbull et al., 2023) from several literature clusters (Lessing & Ogbonnaya, 2026; Lourie & Mcphail, 2024).

We conducted literature searches in Scopus, Web of Science, ERIC, and Google Scholar from March 15, 2026, to May 21, 2026. The core search strategy combined four concept blocks: ("formal proof" OR proving OR "proof construction") AND (geometry OR "plane geometry") AND ("Socratic questioning" OR "Socratic dialogue" OR questioning) AND ("artificial intelligence" OR "generative AI" OR ChatGPT OR "large language model"). Because relatively few studies were expected to contain all four blocks simultaneously, supplementary searches paired proof/geometry with AI, proof with Socratic questioning, and Socratic dialogue with AI or mathematics education. Search strings were adapted to the syntax of each database, and backward reference checking was used to identify additional relevant sources

Sources were eligible when they were peer-reviewed journal articles or academically credible scholarly sources; addressed at least one of the review domains; provided conceptual or empirical evidence relevant to proof construction, questioning, AI-mediated reasoning, proof validation, or responsible AI use; and were published primarily during 2022–2026, while earlier foundational sources were retained when conceptually necessary. We excluded sources that were duplicates, non-scholarly commentary, lacked substantive relevance to mathematical reasoning or proof, or discussed AI only at a technical level without a pedagogical connection. The search initially identified 72 records. After removing 14 duplicate records, 58 unique sources remained for title and abstract screening. All 58 sources met the preliminary relevance criteria and were subsequently assessed in full text. Following full-text assessment, all 58 sources satisfied the eligibility criteria and were included in the conceptual synthesis. The author conducted screening and eligibility decisions using the predefined inclusion and exclusion criteria.

The final literature corpus comprised 58 sources: 36 empirical studies, nine systematic or scoping reviews, seven conceptual or practice-informed articles, five methodological papers that guided the review procedure, and one narrative review as a broader supporting source. The 36 empirical studies addressed several domains relevant to the proposed framework, including mathematical proving and geometry, Socratic questioning, AI-mediated mathematical reasoning, metacognitive scaffolding, and the responsible use of generative AI in educational contexts. Because the included sources varied in their proximity to the focal context of geometry proof construction, the synthesis explicitly distinguished evidence derived directly from studies of geometry and mathematical proving from evidence extrapolated from adjacent educational settings. This distinction avoided treating findings from broader AI-supported learning or questioning contexts as equivalent to evidence obtained specifically within formal mathematical proof settings.

For synthesis, each included source was examined for the pedagogical or epistemic function attributed to proof activity, questioning, AI interaction, or validation. We coded extracted functions into recurring constructs, including identifying givens and goals, generating intermediate claims, checking theorem applicability, constructing warranted inferences, validating proof steps, and reflecting on unsuccessful strategies. Similar functions expressed with different terminology were grouped into higher-order synthesis categories. The six phase labels are therefore author-generated synthesis categories rather than stages reproduced from any single study. We used a parallel mapping procedure to derive the taxonomy of AI-mediated Socratic questions. The Results section reports representative literature and the scope of evidence for each phase and question category, distinguishing direct geometry/proof evidence from evidence extrapolated from broader mathematics, science, or AI-supported learning contexts.

3. Results and Discussion

3.1. Results

3.1.1. Phase Model of Socratic Questioning for Formal Proof in Plane Geometry

The six-phase model was not treated as an instructional sequence already established in the literature. Instead, we developed it by grouping recurring proof-related functions identified across the reviewed sources.

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Individual studies used different terminology and investigated different parts of the proving process; consequently, the phase labels used here represent higher-order synthesis categories. Table 1 makes this derivation explicit by showing the main function represented by each phase, its literature base, and the scope of evidence from which it was derived.

Table 1. Literature-to-Phase Mapping for the Proposed Six-Phase Framework

Proposed phase	Function identified in the synthesis	Representative literature	Scope of evidence
Problem orientation	Identifying givens, goals, geometric objects, and admissible information	Anwar et al. (2023); Ariawan et al. (2026); Lessing & Ogbonnaya (2026)	Direct geometry/proof evidence
Conjecture articulation	Generating and examining possible intermediate claims or proof directions	Kirsten & Greefrath (2023); Neuhaus-Eckhardt et al. (2025); Lessing & Ogbonnaya (2026)	General mathematical proving; partly geometry-relevant
Theorem-condition analysis	Checking theorem applicability and whether required hypotheses have been established	Braithwaite (2022); Hein & Prediger (2024); Winer & Battista (2022)	Direct proof/geometry evidence
Inferential construction	Organising claims and warrants into a logically connected argument	Hein & Prediger (2024); Kontorovich et al. (2024); Neuhaus-Eckhardt et al. (2025)	Direct proof evidence
Proof validation	Examining validity, unsupported assumptions, and inferential gaps	Kirsten & Greefrath (2023); Neuhaus-Eckhardt et al. (2025); Yoon et al. (2024)	Proof research; includes GenAI-assisted proving
Metacognitive reflection	Reviewing unsuccessful strategies, revisions, and the quality of reasoning	Hoffman et al. (2024); Huda et al. (2025); Kontorovich & Greenwood (2024)	Proof evidence plus extrapolation from mathematics learning

The six phases were not established a priori as a fixed instructional sequence. Rather, they emerged from a comparative synthesis of recurring reasoning functions identified across the reviewed literature on geometric proof, mathematical proving, proof validation, and AI-supported mathematical reasoning. During the synthesis, we compared and grouped activities described under different terminology according to their primary function in the proving process. Activities concerned with interpreting givens, goals, geometric objects, and admissible information were consolidated as problem orientation; those involving the generation and examination of possible intermediate claims were grouped as conjecture articulation; and those concerned with theorem applicability and the verification of required hypotheses were classified as theorem-condition analysis. Similarly, studies emphasizing the coordination of claims and mathematical justifications informed the inferential construction phase, whereas work focusing on logical adequacy, unsupported assumptions, and inferential gaps informed proof validation. Finally, studies emphasizing monitoring, revision, and evaluation of reasoning contributed to the metacognitive reflection phase. Thus, the six-phase framework represents an analytic consolidation of recurring proof-related functions across the literature rather than a sequence derived from a single source or imposed independently by the authors.

The associated question categories were derived using the same functional logic. Questions were classified according to the reasoning function they were intended to elicit within each phase. Problem-orientation questions focus on identifying givens, goals, relevant geometric objects, and admissible information; conjecture-oriented questions elicit and examine possible intermediate claims or proof directions; theorem-condition questions require students to identify potentially relevant mathematical results and verify whether their hypotheses have been established; inferential or warrant-seeking questions require students to make explicit the mathematical justification connecting one claim to another; validation questions direct attention to unsupported assumptions, inferential gaps, and dependence on diagrammatic appearance; and metacognitive questions prompt students to examine unsuccessful strategies, revisions, and the quality of their reasoning. Accordingly, these question categories are treated as operationalizations of the literature-derived reasoning functions represented by the six phases rather than as independently constructed prompt types.

In genuine proof work, students often go back to earlier steps when they realize that a conjecture is not supported, a theorem's conditions are not met, or a diagrammatic assumption is accepted without justification (Kirsten & Greefrath, 2023; Lessing & Ogbonnaya, 2026). This phase-based organization also aligns with empirical geometry-proof instruction, in which problem interpretation, diagram construction, exploration, conjecture formulation, and formal justification are treated as interconnected components of the proof process (Ariawan et al., 2026).

The first step is orienting to the problem. At this step, students identify what is given, what must be demonstrated, the geometric objects involved, and which aspects of the diagram the problem statement

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

guarantees (Lessing & Ogbonnaya, 2026; Wardat et al., 2023). The second stage is speculation articulation. Students express a possible proof as congruence of two triangles, parallelism of two lines, or equality of two angles (Lessing & Ogbonnaya, 2026; Neuhaus-Eckhardt et al., 2025). The third phase is theorem-condition analysis, an essential step in formal proof (Lessing & Ogbonnaya, 2026; Neuhaus-Eckhardt et al., 2025). First, students must decide whether the theorem/property they want to utilize is appropriate in the first place (Kirsten & Greefrath, 2023; Lessing & Ogbonnaya, 2026). At this stage, the role of the theorems shifts from retrieval to conditional reasoning through Socratic questioning (Huda et al., 2025; Lessing & Ogbonnaya, 2026).

The fourth phase is the inferential construction when students weave their reasoning into a cohesive chain of claims and explanations (Lessing & Ogbonnaya, 2026; Neuhaus-Eckhardt et al., 2025). The goal is to help students build proof as an argumentative text, not as a list of unrelated facts (Lessing & Ogbonnaya, 2026; Neuhaus-Eckhardt et al., 2025). The fifth phase is proof validation. In this phase, students do not merely write a proof; they also examine whether the reasoning is mathematically valid. In the context of plane geometry, such validation should also require students to examine whether any conclusion depends only on the appearance of the diagram, whether the required conditions of a theorem have been established, and whether each conclusion is supported by sufficient mathematical justification (Neuhaus-Eckhardt et al., 2025; Wardat et al., 2023). Through this phase, proof is treated not simply as a finished written product, but as a mathematical object that must be checked, evaluated, and justified.

The sixth step is metacognitive reflection. Students discuss what was difficult about their proof, which conjectures failed, which conditions of the theorem were overlooked, and how their strategy changed during the process (Hoffman et al., 2024; Huda et al., 2025). Students learn not simply to solve one problem but to acquire habits of mathematical thinking: to see what data may be used, to follow the chains of inference, to reject conclusions that are not justified, and to amend arguments when the justification is not full (Huda et al., 2025; Lessing & Ogbonnaya, 2026). In this approach, pupils grow progressively able to take charge of their own proof activity (Hoffman et al., 2024; Huda et al., 2025).

3.1.2. Taxonomy of Socratic Questions Used for Proving Plane Geometry

To make the pedagogical use of AI-mediated Socratic questioning more operational in plane geometry proof learning, the conceptual discussion above can be translated into a set of classroom-oriented question categories. These categories do not position AI as a source of final proofs, but as a dialogic tool that prompts students to clarify given information, examine definitions, test diagrammatic assumptions, verify theorem conditions, construct intermediate claims, and justify each inferential step. In this sense, the questions are designed to preserve students' responsibility as the authors of the proof while allowing AI to support reflection, strategic planning, and formalization. Table 2 presents the proposed taxonomy.

Table 2. AI-Mediated Socratic Question Categories in Plane Geometry Proof Construction

Category	Pedagogical Purpose in Proof Construction	Example Socratic Questions	Student Responsibilities	Literature Basis
Orientation	Clarifies the structure of the problem by distinguishing given information, objects, relations, and the statement to be proved.	What is explicitly given in the problem? What exactly must be proved?	Correct the problem structure without adding any unsupported assumptions.	Anwar et al. (2023); Lessing & Ogbonnaya (2026); Winer & Battista (2022)
Reasoning Based on Definitions	Connects formal mathematical meanings with geometric objects and relationships.	What is the official definition of midpoint, angle bisector, parallel line or congruent segment in this problem?	Use definitions as a proper mathematical justification, not as an informal explanation.	Hein & Prediger (2024); Winer & Battista (2022)
Diagrammatic Reasoning	Stops students from just using the way the diagram looks or impressions from taking measurements on it.	What was said: What was logically inferred: What was visually assumed?	Distinguish between information that is stated, logically derived, and visually assumed.	Ariawan et al. (2026); Ramírez-Uclés & Ruiz-Hidalgo (2022); Winer & Battista (2022)
Conditions of Theorem	Ensures that a theorem, property or congruence criterion is applied only when all of its necessary conditions have been verified.	In what circumstances can this theorem or a congruence criterion be employed?	Before applying the theorem, check and justify each of its hypotheses.	Braithwaite (2022); Hein & Prediger (2024); Kontorovich et al. (2024)
Planning	Helps students plan a proof by finding	What intermediate claims can be useful	Select, compare and justify	Kirsten & Greefrath

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Category	Pedagogical Purpose in Proof Construction	Example Socratic Questions	Student Responsibilities	Literature Basis
	intermediate claims that help connect the given information and the conclusion.	to bridge the given information and what is to be proved?	intermediate claims that can help advance the proof.	(2023); Neuhaus-Eckhardt et al. (2025)
Justification of inference	Reinforces deductive reasoning by demanding each statement in the proof have a valid reason.	Why this statement is mathematically sensible? Do we get the conclusion from the above step?	Construct a logical chain of statements and explanations. Each statement is a logical consequence of the previous one.	Hein & Prediger (2024); Kontorovich et al. (2024); Lessing & Ogbonnaya (2026)
Counterexamples for Testing	It tests hidden assumptions and whether the argument extends past a certain diagram.	If we alter the diagram must the conclusion still follow? Any example you know where that isn't the case?	Asses the generality of the argument and qualify any unsupported or overgeneralized claims.	Mahoney et al. (2023); Ramírez-Uclés & Ruiz-Hidalgo (2022)
Formalisation	Translates informal reasoning into written proof.	In what sequence should the statements and reasons be written so that the proof will be clear in a logical sense?	Construct a formal, logical proof of a theorem using definitions, theorems, statements and reasons.	Hein & Prediger (2024); Kontorovich et al. (2024); Winer & Battista (2022)
Reflection	Develops metacognition by having students judge the strength, clarity and validity of their proof.	What is the weakest link in your proof? Why? How do we make it stronger?	Find gaps, weak arguments, or unclear points in the proof, and improve it.	Hoffman et al. (2024); Kirsten & Greefrath (2023); Huda et al. (2025)

The categories in Table 2 should not be interpreted as a taxonomy reproduced from a single prior study. They represent a synthesis of questioning functions and proof-related demands appearing across the reviewed literature. Some categories, such as diagrammatic reasoning, theorem-condition checking, and inferential justification, are grounded directly in research on geometry or mathematical proving. Others, particularly reflection and AI-mediated dialogic prompting, draw partly on findings from broader mathematics or AI-supported learning contexts. Their inclusion in the present taxonomy therefore represents a theoretically informed transfer that still requires empirical examination in plane geometry

This taxonomy supports research showing that the Socratic method improves critical thinking through the process of students expressing their ideas, challenging their beliefs, and defending their arguments. However, in plane geometry, questions must be more domain-specific (Aisyah et al., 2023; Anwar et al., 2023; Winer & Battista, 2022). A general Socratic prompt, such as 'Why do you think so?', is pedagogically stronger when it is translated into a question based on geometry, such as: 'Which theorem justifies this angle relation, and have all the required conditions of the theorem been established? (Aisyah et al., 2023; Anwar et al., 2023; Winer & Battista, 2022). Thus, AI-mediated Socratic questioning invites students to explain their thinking and helps them connect definitions, diagrammatic information, theorem conditions, and deductive justification in a formal proof.

3.2. Discussion

3.2.1. Formal Proof of Plane Geometry

A formal proof in plane geometry is a deductive argument in which a conclusion follows necessarily from given information, definitions, previously established results, and accepted mathematical principles (Kontorovich et al., 2024; Winer & Battista, 2022). What makes an argument a proof is therefore not the appearance of mathematical language, but the explicit inferential structure connecting its premises to its conclusion. Each step must be open to examination and justified in relation to what is already known (Kontorovich et al., 2024; Kontorovich & Greenwood, 2024).

Difficulties in constructing such arguments often arise when students must coordinate different forms of representation, particularly diagrams, verbal statements, and symbolic expressions (Hein & Prediger, 2024; Winer & Battista, 2022; Hultdin & Norqvist, 2024). In many cases, the difficulty is not simply that students do not know an appropriate theorem. Rather, they may recognize a theorem without adequately checking whether its conditions are satisfied in the situation being considered (Braithwaite, 2022; Kontorovich et al., 2024; Winer

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

& Battista, 2022). This distinction is important in geometry because a visually convincing relation is not necessarily a mathematically warranted one.

Proof also has a communicative and social function. It is not merely a private sequence of deductions, but an argument that ca a mathematical community can present, question, refine, and evaluate (Kontorovich & Greenwood, 2024; Planas & Pimm, 2024). Students therefore need to do more than arrive at a correct conclusion. They should be able to explain why a particular step is warranted and recognize why an apparently plausible argument may fail because of an unstated assumption, circular reasoning, or insufficient justification (Braithwaite, 2022; Kontorovich et al., 2024). This need to make assumptions and warrants explicit provides an important pedagogical point of entry for Socratic questioning.

3.2.2. Socratic Questions as a Scaffold of Proof

Socratic questioning rests on the assumption that learners can develop understanding by examining and articulating the basis of their own reasoning rather than simply being given an answer (Mahoney et al., 2023; Santos-Guevara et al., 2026; Tang & Putra, 2025). Such questions can highlight assumptions, prompt reasons for a claim, and encourage students to reconsider the adequacy of their reasoning. In this sense, questioning does not substitute for mathematical reasoning; its pedagogical function is to make that reasoning more deliberate and visible (Mahoney et al., 2023; Santos-Guevara et al., 2026).

In the context of proof, Socratic questioning can serve several closely related purposes (Mahoney et al., 2023; Santos-Guevara et al., 2026). It may first clarify the meaning of formal terms such as parallel, midpoint, or congruent (Hein & Prediger, 2024; Winer & Battista, 2022). It may then challenge the status of information represented in a diagram by asking whether a relation is given, established through reasoning, or merely suggested by the visual representation (Ramírez-Uclés & Ruiz-Hidalgo, 2022; Winer & Battista, 2022). Questions can also focus attention on the use of theorems: whether a selected theorem is relevant and whether its required conditions are met (Braithwaite, 2022; Kontorovich et al., 2024). Finally, students may be encouraged to consider whether their conclusions remain valid when a diagram is redrawn or its visual appearance changes (Ramírez-Uclés & Ruiz-Hidalgo, 2022; Winer & Battista, 2022). These forms of questioning are particularly relevant to formal proof because they direct attention to the grounds on which a mathematical statement is accepted, rather than only to whether the final answer appears correct.

3.2.3. Artificial Intelligence as a Questioner in Dialogue, not as an Evidential Authority

Generative AI can produce mathematical text quickly and often in a form that appears coherent and convincing (Farquhar et al., 2024; Walter, 2024). Fluency, however, should not be equated with mathematical validity, particularly in proof-related activities (Plevris et al., 2023; Yoon et al., 2024). Incorrect arguments may still be expressed using appropriate terminology and seemingly orderly sequences of statements (McGalliard & Otten, 2025; Wardat et al., 2023; Yoon et al., 2024). For this reason, the educational value of generative AI in plane geometry depends less on its capacity to generate answers than on how its role is pedagogically constrained (McGalliard & Otten, 2025; Pepin et al., 2025; Walter, 2024). One possible positioning is to use AI as a dialogical questioner rather than as an authority that determines whether a proof is mathematically valid (Santos-Guevara et al., 2026; Tang & Putra, 2025; Yoon et al., 2024). Under such a design, an AI response is treated as material for examination rather than as a completed proof to be accepted by the learner (Pepin et al., 2025; Walter, 2024; Yoon et al., 2024).

This positioning requires students to continue making mathematical judgments. They still need to ask whether a proposed inference is valid, whether the conditions of a theorem have been met, whether an argument contains circular reasoning, and whether a conclusion depends only on what appears to be true in the diagram (Braithwaite, 2022; Hein & Prediger, 2024; Winer & Battista, 2022). AI literacy is therefore relevant not only to the technical use of a generative system but also to students' ability to recognize the limits of its mathematical authority (Walter, 2024).

In practice, this changes the nature of the interaction with AI. Instead of asking the system to "build a complete proof," students may ask it to question their reasoning, identify points that require justification, or challenge a step without supplying the final argument (Santos-Guevara et al., 2026; Tang & Putra, 2025; Walter, 2024). The intention is to keep the principal intellectual work with the learner while using the technology to sustain reflection and scrutiny of the emerging proof (Pepin et al., 2025; Tang & Putra, 2025; Yoon et al., 2024).

The evidential basis for this proposed use of AI is not uniform across all components of the framework. Research conducted directly in geometry and mathematical proving provides relatively close support for practices such as distinguishing visual information from warranted mathematical claims, checking theorem conditions, constructing inferential chains, and validating proofs (Braithwaite, 2022; Dilling & Herrmann, 2024; Hein & Prediger, 2024; Winer & Battista, 2022; Yoon et al., 2024). Evidence for AI functioning specifically as a

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Socratic interlocutor in plane-geometry proof is less direct. Studies on Socratic dialogue with generative AI, AI-supported metacognition, and learner autonomy offer relevant accounts of dialogic scaffolding, reflection, and student agency, but these mechanisms draw partly from adjacent educational contexts (Huda et al., 2025; Lu & Hu, 2025; Santos-Guevara et al., 2026). The framework proposed in this article therefore combines direct proof-related evidence with theoretically justified extrapolation. Recognizing this distinction is important because the available literature does not yet establish the effectiveness of the complete framework in actual plane-geometry instruction.

3.2.4. Ethical Risks and the Controls

The use of AI-mediated Socratic questioning to construct proofs in plane geometry must be ethically controlled and transparent. But it can also promote dependency and shortcut behavior when students use it to obtain complete proofs without engaging in the reasoning process. This is especially true in formal proof, where the main learning goal is not the final written answer but the capacity to formulate, justify, evaluate, and modify mathematical arguments (Pepin et al., 2025).

AI-generated mathematical responses also require human scrutiny. Although chatbots may support mathematics learning, difficulties remain in interpreting geometric representations and producing consistently reliable mathematical reasoning (Wardat et al., 2023). Concerns have likewise been raised regarding over-reliance, reduced independent problem solving, privacy, and diminished critical engagement when learners treat rapid AI responses as substitutes for their own reasoning (Opesemowo & Ndlovu, 2024).

These risks reinforce the argument for positioning AI as a question-asking partner rather than as the source of mathematical evidence. Such a position does not eliminate the need for ethical regulation; instead, it makes students' responsibility for the argument more explicit. Research on academic integrity similarly cautions against uses of generative AI that obscure authorship or reduce students' accountability for their learning (Ateeq et al., 2024; Gustilo et al., 2024). Recent discussions of assessment in GenAI-enabled environments also emphasize making judgment, justification, transparency, and accountability visible within the task itself (Sharma, 2026). Responsible use, therefore, depends not only on restricting access to AI but also on designing classroom norms and proof activities that preserve student agency and mathematical accountability (Johnston et al., 2024; Yusuf et al., 2024).

3.2.5. Implications

The synthesis supports a conditional pedagogical proposition rather than an effectiveness claim. Specifically, AI-assisted Socratic questioning may provide a plausible scaffold for formal proof learning when AI is deliberately restricted to eliciting, questioning, and challenging students' reasoning instead of supplying a completed proof. The reviewed literature does not establish that the proposed six-phase model improves students' formal proof competence, because the model has not yet been examined as an integrated instructional intervention. Its contribution at this stage is therefore to specify a set of theoretically grounded design conditions that can be tested empirically. The pedagogical plausibility of this proposition depends on several design conditions identified across the reviewed literature (Pepin et al., 2025; Walter, 2024; Yoon et al., 2024).

First, AI must be pedagogically constrained to function as a Socratic questioner rather than as a producer of final proofs (Mcgalliard & Otten, 2025; Santos-Guevara et al., 2026; Tang & Putra, 2025). It promotes students' development of clear definitions of terms, identifying relevant information in diagrams, checking theorem conditions, and expressing the logical relations between inferences (Mcgalliard & Otten, 2025; Yoon et al., 2024). This positioning aligns with technology-supported proof environments in which technological outputs function as feedback for validation while students still construct the structure of the mathematical argument (Ariawan et al., 2026).

Second, students need to make their mathematical reasoning visible in their answers by indicating what is known, what is to be proved, which theorem they are applying, and why each conclusion makes sense (Hein & Prediger, 2024; Winer & Battista, 2022; Yoon et al., 2024). Third, students need to have epistemic agency in the proving process, selecting, refining, and justifying each step of the proof rather than accepting only AI-generated arguments (Mcgalliard & Otten, 2025; Stylianides et al., 2024; Yoon et al., 2024).

Fourth, every AI suggestion must be mathematically verified against definitions, theorem requirements, diagrammatic assumptions, and the proof's deductive structure so students do not treat AI-generated suggestions as authoritative mathematical conclusions (Mcgalliard & Otten, 2025; Stylianides et al., 2024; Yoon et al., 2024). Fifth, the activity should be accompanied by teacher orchestration, AI literacy, ethical rules, and intentional constraints on the use of technology; teachers should check proof norms, control the extent of technological assistance, and ensure that AI is a scaffold for reasoning rather than a replacement for mathematical judgment (Pepin et al., 2025; Tan & Yuan, 2024; Walter, 2024).

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Sixth, assessment should value the process of constructing proof, including students' efforts to make conjectures, justify intermediate steps, revise invalid reasoning, and reflect on the structure of the final argument (Hoffman et al., 2024; Kirsten & Greefrath, 2023; Laina, 2025). Under these conditions, AI-assisted Socratic questioning can support formal proof learning because it scaffolds students' reasoning without transferring ownership of the proof from the learner to the technology (Mcgalliard & Otten, 2025; Santos-Guevara et al., 2026; Tang & Putra, 2025; Yoon et al., 2024).

Taken together, these conditions define how the proposed framework is intended to operate, but they should not be read as evidence that the framework is effective. Whether the combination of these conditions produces measurable gains in students' proof construction, validation, or independence remains an empirical question.

3.2.6. Limitations

This article is theoretical and does not provide empirical evidence that the six-phase model will be effective in all plane geometry classes (Pepin et al., 2025; Santos-Guevara et al., 2026; Yoon et al., 2024). Its efficacy has yet to be empirically validated through experimental, quasi-experimental, design-based research, or mixed methods research (Hein & Prediger, 2024; Santos-Guevara et al., 2026; Yoon et al., 2024).

Not all questions are helpful (Santos-Guevara et al., 2026). Overly general questions can confuse students, while overly specific questions can reveal solutions (Hein & Prediger, 2024; Santos-Guevara et al., 2026). More research is needed to determine which question types work best for different skill levels (Hein & Prediger, 2024; Santos-Guevara et al., 2026). Next, we will develop a dedicated assessment rubric for AI-assisted proofs (Hoffman et al., 2024; Kaldaras et al., 2024; Xia et al., 2024). The rubric should consider mathematical validity, quality of interaction with technology, the possibility of verifying suggestions, and depth of reflection (Hoffman et al., 2024; Kirsten & Greefrath, 2023; Xia et al., 2024; Yoon et al., 2024).

4. Conclusion

This review proposes AI-assisted Socratic questioning as a conceptual approach for organizing student–AI interaction during formal proof construction in plane geometry. By synthesizing literature on geometric proving, Socratic questioning, and AI-supported learning, we organized six recurring proof-related functions into a recursive model comprising problem orientation, conjecture articulation, theorem-condition analysis, inferential construction, proof validation, and metacognitive reflection. The accompanying taxonomy translates these functions into question types intended to keep students responsible for identifying assumptions, selecting and checking theorems, justifying inferential steps, and evaluating the validity of their own arguments. The framework's contribution is therefore conceptual and pedagogical rather than evidential: the present review does not show that AI-assisted Socratic questioning improves formal proof competence, nor does it establish the effectiveness of the six-phase model as an instructional intervention. Some components are supported directly by research on geometry and mathematical proving, whereas others are theoretically extrapolated from studies of Socratic dialogue, AI-supported reasoning, and learner autonomy in adjacent contexts. Future research should consequently examine the feasibility, implementation, and effectiveness of the complete framework in plane geometry classrooms, including whether different question types produce observable changes in proof construction, proof validation, student independence, and the quality of mathematical justification.

Author Contributions

I Putu Wisna Ariawan: Conceptualization, writing - original draft, editing; I Made Candiasa: Writing - review & editing, formal analysis, and methodology; Ni Made Sri Mertasari: Writing - review & editing and Formal analysis. Dewa Gede Hendra Divayana: Writing - review & editing, Validation and Data Curation; I Made Sugiarta: Writing - review & editing, Formal analysis and Validation.

Funding

This research was supported by the Institute for Research and Community Service at Universitas Pendidikan Ganesha under research contract number: 501/UN48.16/PT.01.03/2026 dated April 28, 2026.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/ or publication of this article.

Ethics and Consent Declarations

The study was performed as a narrative literature review. There was no direct involvement of human subjects, recruitment of participants, collection of primary data or access to identifiable personal information. The information used and synthesized in this study was gathered from scholarly sources that were previously

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

published and available in public domain or academic databases. This study did not require approval from an institutional review board (IRB) or ethics committee because no intervention or interaction with human subjects was involved and no private and sensitive personal data were collected.

Informed consent was not applicable as there were no participants involved directly in the study. Participant confidentiality, privacy protection and data anonymization procedures were also not applicable, as the study did not collect or process identifiable data at the individual level. The review was carried out in accordance with the accepted standards of academic integrity, including proper citation and acknowledgment of all original sources. Publication consent was also not required as the manuscript contains no identifiable personal information, images or other material derived directly from human participants.

Data Availability

This research was done using a narrative literature review, and no primary datasets were generated or collected in this study. The data supporting the findings of this study are available from information extracted and synthesized from previously published articles and other relevant scholarly sources cited in the manuscript. The publications reviewed in this study can be accessed through their corresponding journals, publishers, and academic databases using the bibliographic information in the reference list, including the Digital Object Identifier (DOI) for each source when available. The DOI provides a permanent link to the location of the article on the internet and is a good way to access the literature supporting the review. Additional information related to the literature selection and synthesis is available from the corresponding author upon reasonable request. No personal or identifiable participant data were collected or processed in this study.

Declaration on AI Use

During the preparation of this manuscript, the authors used ChatGPT (OpenAI) to assist in the organization and refinement of the narrative synthesis of the literature. The AI-assisted outputs were critically reviewed, checked against original sources, and extensively edited by the authors. The authors accept full responsibility for the accuracy, interpretation and content of the manuscript.

Acknowledgment

We extend our gratitude to the Institute for Research and Community Service at Universitas Pendidikan Ganesha for funding this research through the Research Group Development Scheme for the 2026 Fiscal Year.

References

Aisyah, N., Susanti, E., Meryansumayeka, Siswono, T. Y. E., & Maat, S. M. (2023). Proving geometry theorems: Student prospective teachers' perseverance and mathematical reasoning. *Infinity: Journal of Mathematics Education, 12*(2), 377–392. https://doi.org/10.22460/infinity.v12i2.p377-392

Anwar, L., Goedhart, M. J., & Mali, A. (2023). Learning trajectory of geometry proof construction: Studying the emerging understanding of the structure of Euclidean proof. *Eurasia Journal of Mathematics, Science and Technology Education, 19*(5), Article em2266. https://doi.org/10.29333/ejmste/13160

Ariawan, I. P. W., Ardana, I. M., Sugiarta, I. M., Sariyasa, Yanisari, R., Adiarta, A., & Divayana, D. G. H. (2026). Bridging verbal geometry problems and formal proofs: A GeoGebra-supported element–relation detection framework. *Online Learning in Educational Research, 6*(1), 227–243. https://doi.org/10.58524/oler.v6i1.1225

Ateeq, A., Alzoraiki, M., Milhem, M., & Ateeq, R. A. (2024). Artificial intelligence in education: Implications for academic integrity and the shift toward holistic assessment. *Frontiers in Education, 9*, Article 1470979. https://doi.org/10.3389/feduc.2024.1470979

Awang, L. A., Yusop, F. D., & Danaee, M. (2025). Current practices and future direction of artificial intelligence in mathematics education: A systematic review. *International Electronic Journal of Mathematics Education, 20*(2), Article em0823. https://doi.org/10.29333/iejme/16006

Bimantara, A. R., Sugiatno, & Kusumastuti, N. (2025). Cognitive structure of students through Socratic questioning in algebra. *Al-Jabar: Jurnal Pendidikan Matematika, 16*(1), 71–82. https://doi.org/10.24042/ajpm.v16i1.25657

Bozer Saraç, E. N., & Korkmaz, İ. (2025). The effect of Socratic questioning on secondary school students' higher-order thinking. *International Journal of Modern Education Studies, 9*(2), 482–501. https://doi.org/10.51383/ijonmes.2025.424

Braithwaite, D. W. (2022). Relations between geometric proof justification and probabilistic reasoning. *Learning and Individual Differences, 98*, Article 102201. https://doi.org/10.1016/j.lindif.2022.102201

Chigbu, U. E., Atiku, S. O., & Du Plessis, C. C. (2023). The science of literature reviews: Searching, identifying, selecting, and synthesising. *Publications, 11*(1), Article 2. https://doi.org/10.3390/publications11010002

Dewi, N. S., & Dasari, D. (2023). Systematic literature review: Kemampuan pembuktian matematis. *Jurnal Cendekia: Jurnal Pendidikan Matematika, 7*(1), 240–254. https://doi.org/10.31004/cendekia.v7i1.1987

Dilling, F., & Herrmann, M. (2024). Using large language models to support pre-service teachers mathematical reasoning—An exploratory study on ChatGPT as an instrument for creating mathematical proofs in geometry. *Frontiers in Artificial Intelligence, 7*, Article 1460337. https://doi.org/10.3389/frai.2024.1460337

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Farquhar, S., Kossen, J., Kuhn, L., & Gal, Y. (2024). Detecting hallucinations in large language models using semantic entropy. *Nature, 630*(8017), 625–630. https://doi.org/10.1038/s41586-024-07421-0 Nature

Fock, A., & Siller, H.-S. (2025). Generative artificial intelligence in secondary STEM education in the light of human flourishing: A scoping literature review. *International Journal of STEM Education, 12*, Article 67. https://doi.org/10.1186/s40594-025-00589-5

Gustilo, L., Ong, E., & Lapinid, M. R. (2024). Algorithmically-driven writing and academic integrity: Exploring educators' practices, perceptions, and policies in AI era. *International Journal for Educational Integrity, 20*, Article 3. https://doi.org/10.1007/s40979-024-00153-8

Hein, K., & Prediger, S. (2024). Scaffolds for seeing, using, and articulating logical structures in proofs: Design research study with high school students. *The Journal of Mathematical Behavior, 74*, Article 101123. https://doi.org/10.1016/j.jmathb.2023.101123

Hoffman, K., Williams, T. H., & Kephart, K. (2024). The use of guided reflections in learning proof writing. *Education Sciences, 14*(10), Article 1084. https://doi.org/10.3390/educsci14101084

Huda, N., Anwar, K., Novferma, & Kurniawan, W. (2025). ChatGPT scaffolding in supporting metacognition for limit concepts in guided inquiry mathematics learning. *Journal of Information Technology Education: Research, 24*, Article 038. https://doi.org/10.28945/5645

Hultdin, U. W., & Norqvist, M. (2024). Students' reception of two alternative arrangements of mathematical symbols and words: Differences in focus and text navigation. *The Journal of Mathematical Behavior, 75*, Article 101159. https://doi.org/10.1016/j.jmathb.2024.101159

Johnston, H., Wells, R. F., Shanks, E. M., Boey, T., & Parsons, B. N. (2024). Student perspectives on the use of generative artificial intelligence technologies in higher education. *International Journal for Educational Integrity, 20*, Article 2. https://doi.org/10.1007/s40979-024-00149-4

Kaldaras, L., Akaeze, H. O., & Reckase, M. D. (2024). Developing valid assessments in the era of generative artificial intelligence. *Frontiers in Education, 9*, Article 1399377. https://doi.org/10.3389/feduc.2024.1399377

Kirsten, K., & Greefrath, G. (2023). Proof construction and in-process validation—Validation activities of undergraduates in constructing mathematical proofs. *The Journal of Mathematical Behavior, 70*, Article 101064. https://doi.org/10.1016/j.jmathb.2023.101064

Kontorovich, I., & Greenwood, S. (2024). From collaborative construction, through whole-class presentation, to a posteriori reflection: Proof progression in a topology classroom. *International Journal of Research in Undergraduate Mathematics Education, 10*, 516–546. https://doi.org/10.1007/s40753-023-00217-z

Kontorovich, I., Liu, N. Q., & Kang, S. (2024). Transitioning to proof via writing scripts on the rules of a new discourse. *Educational Studies in Mathematics, 117*, 143–162. https://doi.org/10.1007/s10649-024-10324-6

Laina, V. (2025). Employing digital tools for collaborative and flexible mathematical proving. *Digital Experiences in Mathematics Education, 11*, 247–261. https://doi.org/10.1007/s40751-024-00157-6

Lessing, M., & Ogbonnaya, U. I. (2026). Reasoning and proof in mathematics education: A systematic literature review. *Discover Education, 5*, Article 31. https://doi.org/10.1007/s44217-025-01016-1

Lestari, I. L., Sari, M., Uripno, G., Suprihatiningsih, S., Hariyanti, F., & Bonyah, E. (2025). Analysis of artificial intelligence assisted proof process through principle of mathematical induction in real analysis course. *Journal of Mathematical Pedagogy, 6*(2), 94–102. https://doi.org/10.26740/jomp.v6n2.p94-102

Lourie, M., & McPhail, G. (2024). A realist conceptual methodology for qualitative educational research: A modest proposal. *New Zealand Journal of Educational Studies, 59*, 393–407. https://doi.org/10.1007/s40841-024-00344-4

Lu, W., & Hu, Z. (2025). Addressing autonomy risks in generative chatbots with the Socratic method. *Science and Engineering Ethics, 31*, Article 41. https://doi.org/10.1007/s11948-025-00567-8

Mahoney, B. B., Oostdam, R. R., Nieuwelink, H. H., & Schuitema, J. J. (2023). Learning to think critically through Socratic dialogue: Evaluating a series of lessons designed for secondary vocational education. *Thinking Skills and Creativity, 50*, Article 101422. https://doi.org/10.1016/j.tsc.2023.101422

Maifa, T. S., Suryadi, D., & Fatimah, S. (2025). Identifying learning obstacles in proof construction for geometric transformations: Conceptual, procedural, and visualization errors. *Infinity: Journal of Mathematics Education, 14*(3), 673–694. https://doi.org/10.22460/infinity.v14i3.p673-694

McGalliard, W., & Otten, S. (2025). AI responses to challenging problems and educator responses to AI availability. *Digital Experiences in Mathematics Education, 11*(2), 319–332. https://doi.org/10.1007/s40751-024-00167-4

Mohamed, M. Z. bin, Hidayat, R., Suhaizi, N. N. binti, Mat Sabri, N. binti, Mahmud, M. K. H. bin, & Baharuddin, S. N. binti. (2022). Artificial intelligence in mathematics education: A systematic literature review. *International Electronic Journal of Mathematics Education, 17*(3), Article em0694. https://doi.org/10.29333/iejme/12132

Neuhaus-Eckhardt, S., Sommerhof, D., Rach, S., & Ufer, S. (2025). Proof comprehension, proof validation, & proof construction: All the same or different skills? An answer based on an empirical analysis. *International Journal of Science and Mathematics Education, 23*(8), 4153–4178. https://doi.org/10.1007/s10763-025-10621-3

Opesemowo, O. A. G., & Adewuyi, H. O. (2024). A systematic review of artificial intelligence in mathematics education: The emergence of 4IR. *Eurasia Journal of Mathematics, Science and Technology Education, 20*(7), Article em2478. https://doi.org/10.29333/ejmste/14762

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Opesemowo, O. A. G., & Ndlovu, M. (2024). Artificial intelligence in mathematics education: The good, the bad, and the ugly. *Journal of Pedagogical Research, 8*(3), 333–346. https://doi.org/10.33902/JPR.202426428

Pepin, B., Buchholtz, N., & Salinas-Hernández, U. (2025). A scoping survey of ChatGPT in mathematics education. *Digital Experiences in Mathematics Education, 11*, 9–41. https://doi.org/10.1007/s40751-025-00172-1

Planas, N., & Pimm, D. (2024). Mathematics education research on language and on communication including some distinctions: Where are we now? *ZDM–Mathematics Education, 56*(1), 127–139. https://doi.org/10.1007/s11858-023-01497-0

Plevris, V., Papazafeiropoulos, G., & Rios, A. J. (2023). Chatbots put to the test in math and logic problems: A comparison and assessment of ChatGPT-3.5, ChatGPT-4, and Google Bard. *AI, 4*(4), 949–969. https://doi.org/10.3390/ai4040048

Rahmawati, D., Vahlia, I., Mustika, Yunarti, T., & Nurhanurawati. (2022). Socrates questions to optimize students critical thinking. *Al-Jabar: Jurnal Pendidikan Matematika, 13*(2), 373–382. https://doi.org/10.24042/ajpm.v13i2.13728

Ramírez-Uclés, R., & Ruiz-Hidalgo, J. F. (2022). Reasoning, representing, and generalizing in geometric proof problems among 8th grade talented students. *Mathematics, 10*(5), Article 789. https://doi.org/10.3390/math10050789

Rizos, I., & Gkrekas, N. (2023). Is there room for conjectures in mathematics? The role of dynamic geometry environments. *European Journal of Science and Mathematics Education, 11*(4), 589–598. https://doi.org/10.30935/scimath/13204

Santos-Guevara, A., Aquines-Gutiérrez, O., Martínez-Huerta, H., Chavarría-Garza, W. X., & Azuela, J. A. (2026). Fostering student engagement and learning perception through Socratic dialogue with ChatGPT: A case study in physics education. *Education Sciences, 16*(2), Article 184. https://doi.org/10.3390/educsci16020184

Sharma, S. (2026). Educational integrity in GenAI-augmented assessment: Making judgement visible. *International Journal for Educational Integrity, 22*, Article 6. https://doi.org/10.1007/s40979-026-00216-y

Stylianides, G. J., Stylianides, A. J., & Moutsios-Rentzos, A. (2024). Proof and proving in school and university mathematics education research: A systematic review. *ZDM–Mathematics Education, 56*(1), 47–59. https://doi.org/10.1007/s11858-023-01518-y

Sukestiyarno, Y. L., Nugroho, K. U. Z., Sugiman, & Waluya, B. (2023). Learning trajectory of non-Euclidean geometry through ethnomathematics learning approaches to improve spatial ability. *Eurasia Journal of Mathematics, Science and Technology Education, 19*(6), Article em2285. https://doi.org/10.29333/ejmste/13269

Sukhera, J. (2022a). Narrative reviews: Flexible, rigorous, and practical. *Journal of Graduate Medical Education, 14*(4), 414–417. https://doi.org/10.4300/JGME-D-22-00480.1 pubmed.ncbi.nlm.nih.gov

Sukhera, J. (2022b). Narrative reviews in medical education: Key steps for researchers. *Journal of Graduate Medical Education, 14*(4), 418–419. https://doi.org/10.4300/JGME-D-22-00481.1 Journal of Graduate Medical Education

Suryanto, D., Aisyah, N., & Susanti, E. (2022). Pengembangan bahan ajar geometri modern dengan strategi syntactic with two-column proof untuk mahasiswa calon guru. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika, 11*(4), 2929–2940. https://doi.org/10.24127/ajpm.v11i4.6162

Tan, Q., & Yuan, Z. (2024). A professional development course inviting changes in preservice mathematics teachers' integration of technology into teaching: The lens of instrumental orchestration. *Humanities and Social Sciences Communications, 11*, Article 934. https://doi.org/10.1057/s41599-024-03408-4

Tang, K.-S., & Putra, G. B. S. (2026). Generative AI as a dialogic partner: Enhancing multiple perspectives, reasoning, and argumentation in science education with customized chatbots. *Journal of Science Education and Technology, 35*(1), 128–140. https://doi.org/10.1007/s10956-025-10240-1 Springer Nature Link

Tristanti, L. B., & Iffah, J. D. N. (2022). Pengembangan media pembelajaran geometri ruang berbasis android berbantuan smart apps creator dalam meningkatkan kemampuan pembuktian. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika, 11*(3), 1716–1728. https://doi.org/10.24127/ajpm.v11i3.5103

Turnbull, D., Chugh, R., & Luck, J. (2023). Systematic-narrative hybrid literature review: A strategy for integrating a concise methodology into a manuscript. *Social Sciences & Humanities Open, 7*(1), Article 100381. https://doi.org/10.1016/j.ssaho.2022.100381

Walter, Y. (2024). Embracing the future of artificial intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. *International Journal of Educational Technology in Higher Education, 21*, Article 15. https://doi.org/10.1186/s41239-024-00448-3

Wardat, Y., Tashtoush, M. A., AlAli, R., & Jarrah, A. M. (2023). ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia Journal of Mathematics, Science and Technology Education, 19*(7), Article em2286. https://doi.org/10.29333/ejmste/13272

Winer, M. L., & Battista, M. T. (2022). Investigating students' proof reasoning: Analyzing students' oral proof explanations and their written proofs in high school geometry. *International Electronic Journal of Mathematics Education, 17*(2), 1–21. https://doi.org/10.29333/iejme/11713

Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education, 21*, Article 40. https://doi.org/10.1186/s41239-024-00468-z

Yoon, H., Hwang, J., Lee, K., Roh, K. H., & Kwon, O. N. (2024). Students' use of generative artificial intelligence for proving mathematical statements. *ZDM–Mathematics Education, 56*(7), 1531–1551. https://doi.org/10.1007/s11858-024-01629-0

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174

Yusuf, A., Pervin, N., & González, M. R. (2024). Generative AI and the future of higher education: A threat to academic integrity or reformation? Evidence from multicultural perspectives. *International Journal of Educational Technology in Higher Education, 21*, Article 21. https://doi.org/10.1186/s41239-024-00453-6

© 2026 The Author(s)
This is an Open Access article under the Creative Commons Attribution-ShareAlike 4.0 International License
ISSN: 2797-3174