Indonesian Technology and Education Journal
E-ISSN: 2985-8216; P-ISSN: 2985-8208
https://journal.diginus.id/index.php/ITEJ

Utilization of Artificial Intelligence in Mathematics Education: Perspectives from Professors and Doctoral Students in Indonesia

Fadhil Zil Ikram[1*], R. H. Yanti Silitonga[2], Rosidah[3], Mei Radia Putri[4]
[1]Universitas Sulawesi Barat, Indonesia
[2] Universitas Pattimura, Indonesia
[3] Universitas Negeri Makassar, Indonesia
[4] Universitas Negeri Malang, Indonesia

Corresponding e-mail : fadhilzikram@unsulbar.ac.id

ARTICLE INFO

Keywords:

Artificial intelligence;
Doctoral students;
Mathematics education;
Professors;
Qualitative study.

Article History

Received: March 06, 2026
Revised : April 29. 2026
Accepted : May 10, 2026

ABSTRACT

Artificial intelligence (AI) is increasingly integrated into higher education, yet comparative studies of how doctoral students and professors view AI in mathematics education remain scarce, particularly in Indonesia. This study explores the perspectives of both groups through a qualitative design involving focus group discussions and semi-structured interviews with eight doctoral students and two mathematics education professors at several public universities. Data were analysed through condensation, presentation, and verification, with triangulation and member-checking ensuring credibility. Five interpretative patterns emerged: (1) heterogeneous AI adoption among doctoral students, ranging from extensive to restrictive-reflective; (2) shared epistemological grounding that AI functions as a tool rather than a substitute for conceptual mastery; (3) a proposal for differential course regulation, distinguishing a 'no AI' category for conceptual courses from a 'maximum AI' category for technology-based courses; (4) asymmetric control strategies that can be bridged through a disclosure-and-verification paradigm; and (5) complementary ethical concerns across roles. Participants also showed limited awareness of AI's potential as conversational, pedagogical, and teachable agents. While the small sample limits generalisability, these findings offer preliminary insights into how AI use is negotiated within supervisor–student relationships and suggest directions for further inquiry into AI adoption in Indonesian doctoral programmes in mathematics education.

This is an open access article under the CC BY-SA license

To cite this article: Ikram, F. Z., Silitonga, R. H. Y., Rosidah, Putri, M. R. (20xx). Utilization of Artificial Intelligence in Mathematics Education: Perspectives from Professors and Doctoral Students in Indonesia. Information Technology Education Journal, 4(1), 113-128. Doi. 10.61255/itej.v4i1.1162

INTRODUCTION

Mathematics is a foundational discipline whose value for developing problem-solving skills across disciplines and professional life has been recognised since ancient times (Maass & Engeln, 2019; Wulandari et al., 2024). In recent years, artificial intelligence (AI) has become increasingly embedded in mathematics education and is now often framed less as a supplementary tool than as an integral partner in teaching and learning (Li, 2025). AI applications have shown promise for supporting personalised and adaptive instruction and for widening access to mathematics learning content for students with diverse learning needs (Koedinger & Aleven, 2007; Lyu et al., 2025).

Research on AI in mathematics education has expanded rapidly over the past two decades. Bibliometric and systematic reviews have identified three dominant roles of AI in the field intelligent tutoring systems, profiling and prediction, and adaptive personalisation Hwang & Tu, (2021) and have grouped AI approaches into broader categories such as robotics, systems, tools, teachable agents, and autonomous agents (Mohamed et al., 2022). Since the release of ChatGPT in late 2022, research attention has shifted markedly towards generative AI. Empirical work suggests that ChatGPT can provide adequate explanations across many mathematical topics, yet remains limited in domains that require deep conceptual understanding, such as geometry, and is often ineffective at correcting misconceptions (Wardat et al., 2023). Experimental evidence in statistics and applied mathematics is mixed: ChatGPT-based learning has supported statistical reasoning and positive attitudes (Wahba et al., 2024), but students using ChatGPT performed better only on general assessments rather than on contextually designed ones (Said et al., 2025). A recent meta-analysis of 69 experimental studies further indicates that ChatGPT improves academic achievement, affective–motivational conditions, and higher-order thinking, and reduces mental load, while showing no significant effect on self-efficacy (Deng et al., 2025). Read together, this literature suggests that the educational value of AI in mathematics is genuine but conditional on task design and the depth of conceptual demand.

A parallel strand of research has examined user perceptions of generative AI in higher education more broadly. Large-scale surveys report broadly positive but qualified attitudes among students, who value AI for personalised support, writing assistance, and brainstorming while raising concerns about accuracy, ethics, privacy, and academic integrity (Chan & Hu, 2023; Ravšelj et al., 2025). At the doctoral level, AI use is intensive: most doctoral students draw on AI for tasks such as brainstorming and outlining (Aich et al., 2025), and attitude has been shown to consistently predict the intention to use ChatGPT for academic writing (Zou & Huang, 2023). On the faculty side, the integration of ChatGPT into pedagogical practice has been associated with improved subjective well-being and reduced stress among lecturers (Cambra-Fierro et al., 2025).

Studies that directly compare student and faculty perspectives reveal an emerging perception gap. Krumsvik (2024) found that around 60% of Norwegian doctoral supervisors remained sceptical of AI use in dissertation writing, even when candidates valued it as a sparring partner. Espartinez (2024) identified three contrasting profiles among Filipino students and professors Ethical Tech Guardians, Balanced Pedagogy Integrators, and Convenience-Embracing AI Enthusiasts while Verboom et al. (2025) mapped seven attitude profiles among AI experts and professors in Portugal, the Netherlands, and the United States, underscoring the need for clear regulation and continuous professional development. Comparable work has also shown that generative AI can enrich writing, feedback, and argumentation in doctoral study while raising concerns about critical thinking and over-reliance (Rafi & Amjad, 2025; Lesh & Lancaster, 2024). The recurring pattern across these studies is cautious positivity tempered by ethical and pedagogical concerns that are unevenly distributed across roles, which suggests that supervisor student dynamics around AI cannot be inferred from single-group studies alone.

Despite this growing body of work, several limitations remain. Studies on AI perceptions in higher education have largely focused on general academic contexts (Chan & Hu, 2023; Ravšelj et al., 2025), the social sciences and humanities (Rafi & Amjad, 2025), or the undergraduate level (Said et al., 2025; Wahba et al., 2024). Within mathematics education specifically, the available evidence has concentrated on secondary students Egara et al., (2025) and school teachers Al Darayseh & Mersin, (2025) rather than on doctoral-level academics. This is a meaningful gap because mathematics has distinctive epistemic demands abstraction, logical deduction, and procedural precision that interact in particular ways with generative AI's known limitations in advanced mathematical reasoning (Wardat et al., 2023), so adoption patterns in mathematics education cannot be assumed to mirror those in other disciplines (Hwang & Tu, 2021; Mohamed et al., 2022). The Indonesian context warrants particular attention: Indonesia hosts one of the largest higher education systems in Southeast Asia and is currently undergoing successive curriculum reforms in which AI adoption is outpacing the development of institutional guidance, yet remains under-represented in the international literature on AI in higher education. To date, no published study has compared the perspectives of doctoral students and mathematics education professors on AI use within Indonesian doctoral programs. Addressing this intersection discipline (mathematics education), academic level (doctoral), role comparison (students and professors), and national setting (Indonesia) constitutes the novel contribution of the present study.

In light of the gaps outlined above, this study aims to explore the perspectives of doctoral students and mathematics education professors on the use of AI in mathematics education in Indonesia. It is guided by the following research question: How do doctoral students and mathematics education professors in Indonesian doctoral programs perceive and negotiate the use of AI in mathematics education and mathematics learning, and in what respects do their perspectives converge or diverge?

METHOD

Research Design

This study employed a qualitative research design to explore the perspectives of academic doctoral students and professors specialising in mathematics education on the use of artificial intelligence (AI) in mathematics education. A qualitative approach was selected because the study sought rich, context-sensitive accounts of how AI is perceived, used, and negotiated in mathematics education and learning, rather than generalisable measurements of attitudes. Data were collected over a one-month period in May 2025.

Research Setting

The study was conducted within the doctoral programs in mathematics education at several public universities in Indonesia. The setting was selected because it brings together active doctoral candidates and senior faculty who regularly engage with both the pedagogical and research applications of AI, thereby providing access to both the user and supervisor perspectives examined in this study.

Participants and Sampling

Participants were recruited through purposive sampling with criterion-based inclusion. The inclusion criteria were: (a) current enrolment in, or affiliation as faculty with, the doctoral programs in mathematics education; and (b) self-reported use of at least one generative AI tool (e.g., ChatGPT, Gemini, Copilot, Claude) for academic purposes during the twelve months preceding data collection. Potential participants were identified through program records and invited by email; participation was voluntary, uncompensated, and could be withdrawn at any point without consequence.

The final sample comprised ten participants: eight doctoral students (four male, four female) and two professors (one male, one female). The sample size was considered appropriate for an in-depth qualitative study aiming for analytic rather than statistical generalisation, and recruitment continued until both groups (doctoral students and faculty) were represented and preliminary thematic saturation was observed in the FGD and interview data. Participant characteristics are summarised in Table 1; to protect anonymity, participants are referred to by code throughout the manuscript.

Table 1. Participant Background and Code

Code	Role	Gender	Academic stage/experience	AI-use background
M1	Doctoral Student	M	1 year of study	ChatGPT or other AI-applications for academic purposes
M2	Doctoral Student	M	1 year of study	ChatGPT or other AI-applications for academic purposes
M3	Doctoral Student	M	1 year of study	ChatGPT or other AI-applications for academic purposes
M4	Doctoral Student	F	1 year of study	ChatGPT or other AI-applications for academic purposes
M5	Doctoral Student	F	1 year of study	ChatGPT or other AI-applications for academic purposes
M6	Doctoral Student	F	1 year of study	ChatGPT or other AI-applications for academic purposes
M7	Doctoral Student	M	1 year of study	ChatGPT or other AI-applications for academic purposes
M8	Doctoral Student	F	1 year of study	ChatGPT or other AI-applications for academic purposes
P1	Professor	M	40 years of academic experience	ChatGPT or other AI-applications for academic purposes
P2	Professor	F	37 years of academic experience	ChatGPT or other AI-applications for academic purposes

Data Collection

Data were collected through focus group discussions (FGDs) and semi-structured interviews with doctoral students and professors. The FGD was conducted in two sessions: the first session lasted 45 minutes and focused on doctoral students, whilst the second session focused on both the doctoral students and the professors and lasted 30 minutes. Interviews were conducted individually using a semi-structured method and lasted approximately 15–30 minutes. The interview sessions were conducted in two formats, depending on whether participants chose to attend in person or online. Both the interviews and the FGD were conducted in Indonesian to maximise the responses from the research participants. Seven doctoral students participated in the first session of the FGD. As for the second session, the seven doctoral students and one professor participated. For the interview, all research participant participated and they were individually interviewed. As for the two participants who were not in the FGD, they did not attend due to the personal reasons. Therefore, researchers conducted the interview directly.

The interview and FGD guides covered four thematic areas derived from the research question and the reviewed literature: (i) actual use and perceived benefits of AI in mathematics teaching, learning, and research; (ii) obstacles, limitations, and challenges encountered; (iii) ethical considerations and academic integrity; and (iv) implications for doctoral supervision, curriculum, and assessment. Representative open-ended prompts included: "Could you describe specific tasks for which you have used AI in your mathematics education work?", "What benefits or improvements have you experienced?", "What concerns or risks, if any, do you associate with AI use?", and "How, in your view, should AI use be regulated within doctoral mathematics education?" Follow-up probes were used flexibly to explore unanticipated themes. The guide was developed by the research team and reviewed for content relevance and clarity by two researchers external to the study who held expertise in mathematics education and in qualitative methods; minor wording changes were made based on their feedback prior to data collection.

All FGD and interview sessions were audio-recorded with participants' explicit, prior consent. Recordings were transcribed verbatim by the research team, checked against the audio for accuracy, and de-identified before analysis. Field notes were also taken during and immediately after each session to capture non-verbal cues, group dynamics, and the researchers' initial impressions, and were used to inform subsequent analysis. The data from the FGD and the interview will be integrated to look for patterns or consistencies. Interview was employed to have a better understanding or to explore more about participants' responses during FGD.

Data Analysis

Data were analysed following the three-stage interactive model proposed by Miles et al. (2020) data condensation, data display, and conclusion drawing and verification and operationalised through a hybrid inductive–deductive thematic coding strategy. The deductive component drew an initial coding frame from the four thematic areas of the interview guide; the inductive component allowed additional codes to emerge from close reading of the transcripts and was used to refine, expand, or revise the initial frame.

The analytic process proceeded in five steps. First, familiarisation: each transcript was read in full at least twice by the lead researcher to develop an overall sense of the data and to generate preliminary analytic memos. Second, initial coding: meaningful segments of text phrases, sentences, or short passages expressing a single idea were assigned descriptive codes capturing participants' statements about AI use, benefits, challenges, ethical concerns, and supervisory implications. Third, category formation: related codes were clustered into sub-categories, which were then grouped into broader overarching categories through constant comparison across transcripts. Fourth, theme development: categories were compared across participants and across the two role groups (doctoral students and professors) to identify recurring patterns, convergences, and divergences, leading to the interpretive themes reported in the Findings. Fifth, verification: themes were tested against the full dataset using negative case analysis; cases that did not fit the emerging patterns were retained and used to refine, qualify, or extend the themes rather than excluded, ensuring that the final account reflected the full range of the data.

To strengthen the rigour of coding, two members of the research team independently coded an initial subset of approximately 25% of the transcripts. Coding discrepancies were discussed in regular debriefing meetings until consensus was reached, and the agreed coding frame was then applied to the remaining data. An audit trail was maintained throughout the analysis, documenting code definitions, category revisions, decision rules, reflexive memos, and meeting notes, to support the transparency and traceability of analytic decisions.

Trustworthiness

The credibility and trustworthiness of the findings were supported through four strategies. First, methodological triangulation was achieved by comparing accounts elicited through the two data collection methods (FGDs and individual interviews) and across the two participant groups (doctoral students and professors), allowing convergent and divergent perspectives to be identified. Second, member-checking was conducted by returning concise summaries of preliminary interpretations to participants for verification and clarification (Creswell & Guetterman, 2019). Third, peer debriefing was carried out through regular meetings of the research team, in which interpretations, coding decisions, and potential researcher biases were openly discussed. Fourth, an audit trail comprising raw data, coding records, category development logs, and reflexive memos was retained throughout the study to support dependability and confirmability.

Ethical Considerations

The study was conducted in accordance with the ethical principles of voluntary participation, informed consent, confidentiality, and the protection of personal data. Before any data were collected, each participant received a written information sheet describing the study's purpose, procedures, anticipated duration, potential benefits and risks, and their rights as participants, including the right to withdraw at any time without consequence.

To protect participant anonymity, all identifying information (names, institutional affiliations, course titles, and other identifiers) was removed from transcripts and replaced with alphanumeric codes (M1–M8 for doctoral students; P1–P2 for professors). Audio recordings, transcripts, and analytic files were stored on a password-protected, institutionally managed device accessible only to the research team. No personally identifiable information appears in any publication arising from the study.

RESULTS AND DISCUSSION

The findings most directly relevant to the research question are organised in Table 2 under seven themes that emerged from the FGD and interview data: (i) age-related guidance for AI use; (ii) AI use across doctoral study activities; (iii) perceived benefits of AI; (iv) challenges and risks of AI; (v) strategies and regulations for AI use in higher education; (vi) the role of AI in mathematics education; and (vii) future outlook. Within each theme, sub-themes capture more specific patterns, and one illustrative excerpt per sub-theme is shown to ground the sub-theme in participant data. Themes that emerged in the discussions but did not bear directly on AI in mathematics education notably general technology-restriction policies and the monitoring of school inspectors have been excluded from Table 2 to keep the presentation focused on the research question.

Table 2. Themes, sub-themes, and illustrative excerpts from the focus group discussions and interviews.

Theme	Sub-Theme	Illustrative excerpt (English Translation)	Participant
Perception on Age-related guidance for AI use	Elementary level	*"If they are still children, they need parental guidance; they may use AI, but only under supervision — they should not access it on their own."*	M4
	Secondary level	*"In lower and upper secondary school, they may use AI independently, but there should still be supervision from teachers and parents or guardians."*	M4
	Higher education and adults	*"Once you are an adult, it is fine to use AI independently. However, ethical considerations, ethical issues, and data privacy are crucial to raise awareness about."*	M4
AI use across doctoral study activities	Data analysis	*"I use AI from data analysis through to literature review… SPSS can be combined with NVivo for data analysis."*	M5
	Literature review	*"For literature reviews, we can use ChatGPT and DeepSeek alongside SkySpace."*	M5
	Working through mathematical problems	*"For problems like ODEs, I usually use MathGPT to see how to solve them, and then I make a video of it."*	M5
	Source of inspiration and discussion partner	*"There are things I had not thought of but AI had… we take those ideas and adapt them."* *"AI can help to provide inspiration or suggest improvements to our lesson …"*	M6 P2
Perceived benefits of AI	Time efficiency	*"The benefit is in the time saved… where people once had to read many articles, now we can use SkySpace."* *"…can save time when designing the lessons to*	M8 P2

Theme	Sub-Theme	Illustrative excerpt (English Translation)	Participant
		be delivered…"	
	Creative stimulation	*"I wanted to research something but was unsure how to phrase the title… after typing it into ChatGPT, the idea came to me."*	M6
Challenges and risks of AI	Cognitive dependence	*"I am afraid that if I get involved with AI, I will become dependent on it, and my own cognitive processing will decline."*	M7
		"The challenge with using AI is that students must always approach the output with a critical eye; otherwise, it can lead to dependency, particularly in cognitive terms"	P2
	Limitations of AI logic	*"If there is a division by zero, it gets confused… the loop keeps going round and round, so it ends up giving an error."*	P1
	Bias and academic integrity	*"We must use AI wisely… it should not be used to deceive lecturers who are not yet familiar with AI."*	P1
Strategies and regulations for AI use in higher education	Course grouping	*"Courses that cannot use AI — such as Real Analysis and Algebraic Structures — should be balanced by IT-based learning-resources and media courses, where AI use is maximised."*	M8
	Creative assessment design	*"As teachers, we need to be more creative in setting tasks that incorporate AI while still engaging students' individual thinking processes."*	M8
	Institutional support	*"Ideally, the university should provide tools like that [Wolfram Alpha Pro]… because it is part of the lab."*	P1
Role of AI in mathematics education	AI as a tool, not a substitute	*"ChatGPT serves the same purpose as GeoGebra, as does Cabri… there is plenty of software out there to help."*	P1
		"*AI should serve as a tool, not as the primary means of conducting teaching or research*"	P2
	Verification by the lecturer	*"I would say: "I used ChatGPT and my analysis is like this. Please confirm, Professor, whether it is correct or not." That would actually be excellent."*	P1
Future outlook	Continued, unavoidable use	*"Of course we cannot shut ourselves off; if we do, we will only fall behind."*	Several participants

Note. M denotes a doctoral student participant; P denotes a professor. Excerpts have been translated from Bahasa Indonesia; original transcripts in the source language are available from the corresponding author on request. The participant code in the rightmost column identifies the speaker of the excerpt shown, not necessarily the only participant who articulated the sub-theme.

Table 3 provides a comparison of the responses of the doctoral students and professors in the field of mathematics education.

Table 3. Comparisons of Doctoral Students' and Professors' Perspectives on AI in Mathematics Education

Aspect	Doctoral Students' Perspective	Professors' Perspective
Position of AI	Tool to support academic productivity (research, writing, idea generation)	Tool for knowledge exploration, equivalent to a textbook or mathematical software

E-ISSN: 2985-8216
P-ISSN: 2985-8208

Aspect	Doctoral Students' Perspective	Professors' Perspective
Main focus	Practical applications across doctoral study activities	Mastery of fundamentals remains the priority; AI is a supporting tool
Control strategy	Course grouping, creative assessment, self-monitoring	Academic verification through direct questioning; expectation of transparent disclosure
Key risks identified	Dependence, bias, over-reliance on AI outputs	Limitations of AI logic; potential misuse to "deceive" supervisors
Stance on AI guidance	Need for creative pedagogical responses from lecturers without waiting for lengthy institutional processes	Aligned with international trends balancing AI with conventional teaching
Attitude toward the future	Continued use; "cannot shut ourselves off"	Optimistic; sees AI advances as the realisation of long-standing academic aspirations

Across the two tables, five interpretive patterns recur, together with a further observation about an area that participants did not address. These are taken up in the Discussion section below.

Pattern 1: Heterogeneity in AI adoption among doctoral students

Among the eight doctoral students, AI engagement ranged along a spectrum from extensive to restrictive-reflective (Table 2, rows 4–10). At one end, M5 reported integrating AI into nearly every phase of doctoral work data analysis (combining NVivo and SPSS with AI), literature search (ChatGPT, DeepSeek, SkySpace), and problem-solving (MathGPT). In the middle, M6 used AI as a generator of ideas and a discussion partner while emphasising that the originality of the work remained their own. At the other end, M7 deliberately limited engagement with AI out of concern that habitual use would erode their own cognitive processing. Within this study, then, adoption appeared to be shaped not by access alone but by an interplay of perceived benefit, anticipated cognitive risk, and personal disposition.

The range observed here is broadly consistent with quantitative work showing high but variable AI engagement among doctoral students elsewhere. Aich et al. (2025) found that 76.3% of a sample of doctoral students used AI for dissertation-related activities, particularly brainstorming and outlining. Zou and Huang (2023) applying the Technology Acceptance Model to 242 doctoral students at a Chinese technology university, found that attitude was a significant predictor of intention to use ChatGPT for academic writing and mediated the influence of perceived benefits and perceived ease of use. Rafi and Amjad (2025) further documented a spectrum of perceived benefits improving the mechanical quality of writing, providing efficient feedback, enriching thinking, structuring arguments, and supporting data analysis closely matching the range articulated by M5.

M7's restrictive stance is more distinctive but is not without antecedent in the literature. It resonates with what Rafi and Amjad (2025) describe as a "technological singularity syndrome" a concern that AI may surpass academics' capacity for independent thought and with Deng et al.'s (2025) meta-analysis of 69 experimental studies, which found that ChatGPT reduces students' mental load. As Deng and colleagues caution, however, a short-term reduction in mental load could, if not balanced by independent practice, lead to longer-term cognitive atrophy. Read alongside this evidence, M7's position is plausibly interpreted not as conservatism but as a reflective cognitive-protection strategy.

We are cautious about reading too much into this pattern given the small sample. What can be said with reasonable confidence is that, among the doctoral students in this study, AI adoption was not uniform, and a one-size-fits-all assumption about doctoral AI use would not capture the observed variation.

Pattern 2: Convergence on AI as a tool, not a substitute

Despite the heterogeneity among doctoral students, students and professors converged on a single epistemological position: AI is a tool for exploration, not a substitute for conceptual mastery (Table 3, "Position of AI"). P1 articulated this explicitly by equating ChatGPT with GeoGebra, Cabri, and Wolfram Alpha software that supports mathematical exploration without replacing the conceptual work of the user (Table 2).

This converged position aligns with how AI has been conceptualised in syntheses of the field. Hwang and Tu (2021) identified three principal roles of AI in mathematics education through bibliometric analysis intelligent tutoring systems (45.24%), profiling and prediction (28.57%), and adaptive systems and personalisation (21.43%) all instrumental rather than substitutive. Mohamed et al. (2022) extended this with six categories: robotics, systems, tools, teachable agents, autonomous agents, and comprehensive approaches. The professors' framing locates ChatGPT within the "tools" and "adaptive systems" categories of these typologies.

Experimental evidence on where AI's strengths and weaknesses lie supports the same instrumental framing. Wardat et al. (2023) found that ChatGPT can provide adequate explanations across many mathematical topics but remains limited in domains requiring deep conceptual reasoning such as advanced geometry, and is ineffective at correcting misconceptions. Wahba et al. (2024) reported that ChatGPT-based learning supported the development of statistical reasoning and a positive attitude towards statistics. P1's direct observations about AI's failure on operations such as division by zero and the rounding of series approximations (Table 2) are consistent with these documented limitations. Within this study, then, the "AI as tool" position is grounded both in participants' explicit statements and in their reported experiences of where AI succeeds and fails.

Pattern 3: A proposal for differential regulation by course type

A distinctive proposal emerging from the FGD articulated most clearly by M8 (Table 2, "Strategies and regulations") was that AI policy should not be uniform across the curriculum but should be differentiated by the cognitive demands of each course. Courses requiring deep conceptual mastery (Real Analysis, Algebraic Structures, and Graph and Number Theory) would be designated AI-minimal, while technology-based courses (such as Learning Resources and Media) would maximise AI integration. M8 added that within-course design creatively constructing assessments so that AI provides only basic ideas while students develop them could further protect conceptual learning.

Said et al.'s (2025) quasi-experimental study with 54 Business Mathematics students at Sultan Qaboos University provides a strong external warrant for the assessment-design half of this proposal: students using ChatGPT performed better on general assessments but not on contextually designed ones, indicating that contextual assessment naturally limits the displacement of student work by AI. Beyond assessment, broader work supports the view that uniform AI policies are unlikely to be effective. Al Darayseh and Mersin's (2025) extended-TAM survey of 448 mathematics and science teachers in Turkey indicated that a positive attitude towards AI does not translate automatically into unfiltered adoption targeted interventions remain necessary and the existence of multiple distinct attitude profiles among students and faculty (three in Espartinez's (2024) Q-Methodology study with Filipino students and professors, seven in Verboom et al.'s (2025) cross-national study of AI experts and professors in Portugal, the Netherlands, and the United States) suggests that policy designed for an average user is unlikely to fit the actual distribution of perspectives.

We frame this as a participant-generated proposal that the data can support as a useful starting point for further inquiry, rather than as a recommendation derived from sufficient evidence for institutional adoption. The course-grouping component, in particular, is novel in this literature and merits investigation in other mathematics programmes where similar epistemic distinctions exist between proof-based and application-based courses.

Pattern 4: From asymmetric control strategies to a disclosure-and-verification paradigm

Across the two role groups, control strategies were asymmetric (Table 3, "Control strategy"). Doctoral students emphasised internal regulation self-monitoring, course-level choice about whether to engage AI, and creative assessment design. Professors, while sharing the goal of responsible AI use, emphasised external verification: students should bring AI-assisted analyses back to their supervisor for confirmation rather than retain them privately. P1's framing of this expectation that a student saying "I used ChatGPT, here is my analysis; please confirm whether it is correct" would be regarded as commendable rather than problematic (Table 2) reframes AI use from something to be concealed into something to be openly negotiated within the supervisory relationship.

Building on this participant proposal, we describe the expectation as a disclosure-and-verification paradigm. Under this paradigm, students retain the right to use AI as appropriate to the task; in exchange, they disclose specific instances of AI use to their supervisor, who verifies the quality of AI-supported outputs and incorporates this verification into ongoing supervision. The paradigm has three components: (1) Disclosure where students proactively report concrete instances of AI use in their work which tasks, which tools, what was generated, and how the AI output was used or modified; (2) Verification where supervisors examine AI-supported outputs and evaluate their accuracy, depth, and contribution against the standards of the field, drawing on their own disciplinary expertise rather than on rules about AI per se; and (3) Iteration where the disclosure and verification cycle becomes a regular feature of supervisory meetings, building shared understanding of where AI helps and where it does not.

This paradigm is, in our reading, the most actionable contribution of the present study. It addresses precisely the perception gap documented elsewhere in supervisor–student relationships around AI for example, Krumsvik's (2024) finding that around 60% of Norwegian doctoral supervisors remained sceptical of AI in dissertation writing even as their candidates valued it as a sparring partner by neither prohibiting AI use nor leaving it unsupervised. It is also consistent with broader evidence that doctoral AI engagement is intensive (Aich et al., 2025; Zou & Huang, 2023) and that supervisors who themselves engage with ChatGPT in pedagogical practice report improved subjective well-being and reduced stress (Cambra-Fierro et al., 2025); supervisors equipped to engage in disclosure–verification dialogue may therefore benefit personally as well as fulfil their academic role more effectively.

We are explicit that the paradigm, as described here, is grounded in the present study primarily in the perspective of P1 (with concurrence from P2) and in concordant statements from selected doctoral students. Its broader feasibility, the conditions under which disclosure can be encouraged rather than concealed, the specific verification practices that work in mathematics doctoral programs, and any unintended consequences (such as increased supervisory workload) remain open questions that this study cannot settle. We offer the paradigm as a hypothesis worth structured empirical testing, not as an established practice.

Pattern 5: Complementary, role-specific ethical concerns

Ethical concerns were shared across the two groups but distributed asymmetrically (Table 3, "Key risks identified"). Doctoral students focused on threats to themselves cognitive dependence, bias absorbed from AI outputs, and over-reliance (M7; Table 2). Professors focused on threats to the educational relationship the misuse of AI to misrepresent one's work, and the strain this place on academic supervision (P1; Table 2). Read together, the two concern profiles are complementary: one looks inward at the impact of AI on the individual learner; the other looks outward at the impact of AI on the supervisor–student relationship.

This complementarity is consistent with the broader literature, which shows persistent concerns about accuracy, ethics, privacy, and academic integrity across both undergraduate and postgraduate populations. Chan and Hu's (2023) survey of 399 students in Hong Kong and Ravšelj et al.'s (2025) global study of 23,218 students across 109 countries and territories both report broadly positive but qualified attitudes, with regulation seen as necessary to prevent academic dishonesty, plagiarism, and social isolation. What this study contributes in addition is the observation that the locus of concern differs by role: among these participants, students worried about cognition, while professors worried about integrity in supervision. Both concerns are addressed, in part, by the disclosure-and-verification paradigm discussed above disclosure addresses the integrity concern from the professor's side, while open negotiation of AI use can support more reflective student engagement with their own cognition.

How the findings address the research question

The research question asked how doctoral students and mathematics education professors at Indonesian doctoral programs perceive and negotiate the use of AI in mathematics education and mathematics learning, and where their perspectives converge or diverge. The five patterns provide a structured answer.

On convergence, both groups locate AI as a tool rather than a substitute for conceptual mastery (Pattern 2); both express ethical concern, albeit with different emphases (Pattern 5); and both anticipate continued AI use (Table 3). On divergence, doctoral students display heterogeneous adoption ranging from extensive to restrictive-reflective (Pattern 1), while professors articulate a more uniform position; and doctoral students emphasise internal regulation, while professors emphasise external verification through supervisory dialogue (Pattern 4).

On how the two groups negotiate AI use, the most distinctive observation from this study is the asymmetry in control strategies (Pattern 4), which can be bridged by the disclosure-and-verification paradigm articulated by P1 and developed above as the study's principal substantive contribution. The differential-regulation proposal (Pattern 3) addresses negotiation at the curricular rather than the supervisory level. The absence of awareness of pedagogical and teachable agents (the noticeable-absence finding) identifies a possibility for future enrichment of AI use that neither group currently appears to view as available.

Preliminary implications

The implications of these findings are preliminary given the small sample. With that caveat, three directions appear worth further consideration. First, the heterogeneity of doctoral AI adoption suggests that guidance at the doctoral level should accommodate variation rather than assume uniform adoption; reflective restrictive use is a legitimate stance that need not be discouraged. Second, the disclosure-and-verification paradigm even at this exploratory stage offers a model for supervisor–student dialogue around AI that is neither prohibitive nor permissive, and that converts AI use from a hidden practice into a supervised one; whether and how it can be implemented in doctoral program deserves explicit attention in further work. Third, the absence of pedagogical and teachable-agent uses of AI in participants' accounts suggests that doctoral mathematics education may benefit from professional development that surfaces these richer roles, particularly for the conceptual courses where current participants would otherwise exclude AI altogether. We deliberately refrain from making broader institutional policy recommendations, which would require evidence beyond what a ten-participant qualitative study can provide.

Future Research Implications

A further observation emerges not from what participants said but from what they did not say. Across both Tables 2 and 3, the AI uses that participants described were dominated by assistive applications generating ideas, summarising literature, suggesting solutions, performing computations. Richer pedagogical roles that AI can play in mathematics education did not appear in any participant's account, even from the professors.

Three such roles are worth noting briefly, since their absence in the data is itself a finding. As conversational agents, AI systems can engage learners in multi-turn, mixed-initiative dialogue that supports the elaboration and explanation of concepts (Graesser et al., 2001; Kuhail et al., 2023; Nye et al., 2014). As pedagogical agents, AI can act as a tutor or teaching assistant that adapts to learners' needs, difficulties, and characteristics, providing affective as well as cognitive support (Johnson & Lester, 2018; Schlimbach et al., 2022; Tao et al., 2022). As teachable agents, AI takes the role of a learner that the student instructs, with documented benefits including a low-pressure learning environment, increased engagement, and the consolidation of conceptual knowledge through the act of teaching (Debbané et al., 2023; Lyu et al., 2025b; Xing et al., 2025).

We note the absence of these roles in the data rather than treating it as a critique of participant practice. The implication is that pedagogically richer uses of AI may already be available to doctoral mathematics education, but are not currently in view for the participants in this study. In particular, the very conceptual courses that participants would designate AI-minimal under the differential-regulation proposal (Pattern 3) might benefit from pedagogical or teachable agents, which scaffold rather than supply answers, provided that such agents are deployed within an appropriate pedagogical frame. This is offered as a direction for further exploration rather than a conclusion drawn from the present data.

Limitations

Several limitations should be made explicit. The sample is small (n = 10) and drawn from several doctoral programs in mathematics education in Indonesia, so the findings should be read as analytic descriptions of these participants' perspectives rather than as generalisable claims about Indonesian mathematics education or doctoral programs elsewhere. The asymmetric representation of the two professors in the quoted material with P1's perspectives more extensively captured than P2's should be borne in mind; while P2 contributed to the discussion, her views are reflected in the synthesised summary in Table 3 rather than in additional verbatim excerpts. Self-reported perceptions are subject to social desirability bias, and the interviews were conducted at a single point in time during a period in which AI adoption in higher education is evolving rapidly. Further research with a broader, multi-institution sample, using mixed methods and incorporating supervisor perspectives in a more structured way, is needed to validate or revise the disclosure-and-verification paradigm and the differential-regulation proposal developed here.

CONCLUSIONS

Key findings

This study set out to explore how doctoral students and mathematics education professors at Indonesian doctoral programs perceive and negotiate the use of AI in mathematics education and learning. Focus group discussions and interviews with eight doctoral students and two professors yielded five interpretive patterns. Doctoral students' AI adoption was heterogeneous, ranging from extensive integration to deliberately restrictive use. Both groups converged on the view that AI is a tool for exploration rather than a substitute for conceptual mastery. Participants proposed differentiating AI policy by course type rather than applying a single rule across the curriculum. Control strategies differed by role, with doctoral students emphasising internal regulation and professors emphasising verification through supervisory dialogue. Ethical concerns were shared but asymmetrically distributed: students focused on threats to their own cognition, while professors focused on threats to the integrity of the supervisory relationship. An additional observation, drawn from what participants did not say, was the limited awareness of AI's potential roles as conversational, pedagogical, and teachable agents.

Preliminary implications

Given the small sample, the implications of this study are preliminary and should be read as directions for further inquiry rather than as ready-to-implement policy. With that caveat, the findings point to three directions worth exploring. First, guidance for AI use at the doctoral level appears to require accommodation for heterogeneity in adoption rather than a uniform stance. Second, the disclosure-and-verification paradigm developed in the Discussion offers a candidate model for supervisor–student dialogue around AI that is neither prohibitive nor permissive, and warrants explicit empirical testing. Third, the differential-regulation proposal distinguishing AI-minimal conceptual courses from AI-maximal technology-based courses, supported by contextual assessment design merits investigation in other mathematics program where comparable epistemic distinctions exist between proof-based and application-based content.

Limitations and future research

Several limitations should be acknowledged. The sample is small (n = 10) and drawn from several doctoral programs in mathematics education in Indonesia; the findings should be read as analytic descriptions of these participants' perspectives rather than as generalisable claims. The asymmetric representation of the two professors in the quoted data with P1 more extensively captured than P2 and the self-reported nature of the perceptions are further reasons for cautious reading. Further research with a broader, multi-institution sample using mixed methods, and incorporating supervisor perspectives in a more structured way, is needed to validate or revise the disclosure-and-verification paradigm and the differential-regulation proposal explored here. Experimental studies could also test the effectiveness of differentiating AI policy by course type, particularly in the context of doctoral-level mathematics education in Indonesia.

ACKNOWLEDGMENT

The researchers would like to express their deepest gratitude to Universitas Sulawesi Barat for assisting in the completion of this work and for providing Turnitin for plagiarism checking purposes.

AUTHOR CONTRIBUTION STATEMENT

The research team contributed to several aspects of the study. FI contributed to the development of the research ideas and concepts, the drafting of the manuscript, and data collection. RS contributed to data collection and data transcription. R was responsible for data analysis and triangulation, whilst MI was responsible for data collection and the drafting of the manuscript, including the compilation of references.

AI DISCLOSURE STATEMENT

The author used Claude Pro and DeepL during the preparation of this work for translation and readability purposes. After using the tool/service, the author thoroughly reviewed and edited the content as needed and takes full responsibility for the content of the publication.

REFERENCES

Aich, F. Z., Boumahdy, L., Benfilali, I., & Sandy, K. (2025). AI usage in academic writing among Moroccan doctoral researchers at Sidi Mohamed Ben Abdellah University. *Arab World English Journal (AWEJ) Special Issue on CALL, 11*, 192–209. https://doi.org/10.24093/awej/call11.12

Al Darayseh, A., & Mersin, N. (2025). Integrating generative AI into STEM education: Insights from science and mathematics teachers. *International Electronic Journal of Mathematics Education, 20*(3), em0832. https://doi.org/10.29333/iejme/16232

Cambra-Fierro, J. J., Fuentes Blasco, M., López-Pérez, M.-E., & Trifu, A. (2025). ChatGPT adoption and its influence on faculty well-being: An empirical research in higher education. *Education and Information Technologies, 30*(2), 1517–1538. https://doi.org/10.1007/s10639-024-12871-0

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education, 20*(1), Article 43. https://doi.org/10.1186/s41239-023-00411-8

Creswell, J. W., & Guetterman, T. C. (2019). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (6th ed.). Pearson.

Debbané, A., Lee, K. J., Tse, J., & Law, E. (2023). Learning by teaching: Key challenges and design implications. *Proceedings of the ACM on Human-Computer Interaction, 7*(CSCW1), Article 34. https://doi.org/10.1145/3579501

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education, 227*, Article 105224. https://doi.org/10.1016/j.compedu.2024.105224

Egara, F. O., Mosimege, M., & Mosia, M. (2025). Secondary school students' perceptions of their usage of artificial intelligence-based ChatGPT in mathematics learning. *Journal of Education (University of KwaZulu-Natal)*, (98), 124–146. https://doi.org/10.17159/2520-9868/i98a07

Espartinez, A. S. (2024). Exploring student and teacher perceptions of ChatGPT use in higher education: A Q-methodology study. *Computers and Education: Artificial Intelligence, 7*, Article 100264. https://doi.org/10.1016/j.caeai.2024.100264

Graesser, A. C., VanLehn, K., Rosé, C. P., Jordan, P. W., & Harter, D. (2001). Intelligent tutoring systems with conversational dialogue. *AI Magazine, 22*(4), 39–51. https://doi.org/10.1609/aimag.v22i4.1591

Hwang, G.-J., & Tu, Y.-F. (2021). Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review. *Mathematics, 9*(6), Article 584. https://doi.org/10.3390/math9060584

Johnson, W. L., & Lester, J. C. (2018). Pedagogical agents: Back to the future. *AI Magazine, 39*(2), 33–44. https://doi.org/10.1609/aimag.v39i2.2793

Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review, 19*(3), 239–264. https://doi.org/10.1007/s10648-007-9049-0

Krumsvik, R. J. (2024). Chatbots and academic writing for doctoral students. *Education and Information Technologies, 30*, 9427–9461. https://doi.org/10.1007/s10639-024-13177-x

Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies, 28*(1), 973–1018. https://doi.org/10.1007/s10639-022-11177-3

Lesh, J. J., & Lancaster, J. J. (2024). AI-assisted academia: Unveiling doctoral students' perspectives on dissertation in practice innovation. *Bulletin of Educational Studies, 3*(2), 113–126.

Li, M. (2025). Integrating artificial intelligence in primary mathematics education: Investigating internal and external influences on teacher adoption. *International Journal of Science and Mathematics Education, 23*(5), 1283–1308. https://doi.org/10.1007/s10763-024-10515-w

Lyu, B., Li, C., Li, H., Oh, H., Song, Y., Zhu, W., & Xing, W. (2025a). Exploring the role of teachable AI agents' personality traits in shaping student interaction and learning in mathematics education. In *Proceedings of the 15th International Learning Analytics and Knowledge Conference* (pp. 887–894). Association for Computing Machinery. https://doi.org/10.1145/3706468.3706532

Lyu, B., Li, C., Li, H., Oh, H., Song, Y., Zhu, W., & Xing, W. (2025b). The role of teachable agents' personality traits on student-AI interactions and math learning. *Computers & Education, 234*, Article 105314. https://doi.org/10.1016/j.compedu.2025.105314

Maass, K., & Engeln, K. (2019). Professional development on connections to the world of work in mathematics and science education. *ZDM, 51*(6), 967–978. https://doi.org/10.1007/s11858-019-01047-7

Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). *Qualitative data analysis: A methods sourcebook* (4th ed.). SAGE.

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

Nye, B. D., Graesser, A. C., & Hu, X. (2014). AutoTutor and family: A review of 17 years of natural language tutoring. *International Journal of Artificial Intelligence in Education, 24*(4), 427–469. https://doi.org/10.1007/s40593-014-0029-5

Rafi, M. S., & Amjad, I. (2025). The role of generative AI in writing doctoral dissertation: Perceived opportunities, challenges, and facilitating strategies to promote human agency. *Discover Education, 4*(1), Article 165. https://doi.org/10.1007/s44217-025-00503-9

Ravšelj, D., Keržič, D., Tomaževič, N., Umek, L., Brezovar, N., Iahad, N. A., … Aristovnik, A. (2025). Higher education students' perceptions of ChatGPT: A global study of early reactions. *PLOS ONE, 20*(2), Article e0315011. https://doi.org/10.1371/journal.pone.0315011

Said, B., Abdullah, S., & AlGabshi, B. (2025). Evaluating the role of ChatGPT in mathematics education: Impact on student performance, academic integrity, and perceptions in higher education. *International Journal for Multidisciplinary Research, 7*(5), Article IJFMR250556338.

Schlimbach, R., Rinn, H., Markgraf, D., & Robra-Bissantz, S. (2022). A literature review on pedagogical conversational agent adaptation. In *Proceedings of the Pacific Asia Conference on Information Systems* (Vol. 1, pp. 1–14). Association for Information Systems. https://aisel.aisnet.org/pacis2022/

Tao, Y., Zhang, G., Zhang, D., Wang, F., Zhou, Y., & Xu, T. (2022). Exploring persona characteristics in learning: A review study of pedagogical agents. *Procedia Computer Science, 201*, 87–94. https://doi.org/10.1016/j.procs.2022.03.014

Verboom, A. D. P. R., dos Santos, N. R., Pais, L., Mónico, L., & van Echtelt, P. (2025). Perceptions of artificial intelligence in academic teaching and research: A qualitative study from AI experts and professors' perspectives. *International Journal of Educational Technology in Higher Education, 22*(1), Article 46. https://doi.org/10.1186/s41239-025-00546-w

Wahba, F., Ajlouni, A. O., & Abumosa, M. A. (2024). The impact of ChatGPT-based learning statistics on undergraduates' statistical reasoning and attitudes toward statistics. *Eurasia Journal of Mathematics, Science and Technology Education, 20*(7), em2468. https://doi.org/10.29333/ejmste/14726

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), em2286. https://doi.org/10.29333/ejmste/13272

Wulandari, T., Hasanah, R. U., & Fadillah, D. T. (2024). Kemampuan pembuktian matematis mahasiswa: Systematic literature review [Mathematical proof ability of students: A systematic literature review]. *Journal of Multidisciplinary Inquiry in Science, Technology and Educational Research, 1*(3), 350–359. https://doi.org/10.47709/mda.v1i1.3820

Xing, W., Song, Y., Li, C., Liu, Z., Zhu, W., & Oh, H. (2025). Development of a generative AI-powered teachable agent for middle school mathematics learning: A design-based research study. *British Journal of Educational Technology, 56*(5), 2043–2077. https://doi.org/10.1111/bjet.13586

Zou, M., & Huang, L. (2023). To use or not to use? Understanding doctoral students' acceptance of ChatGPT in writing through technology acceptance model. *Frontiers in Psychology, 14*, Article 1259531. https://doi.org/10.3389/fpsyg.2023.1259531