MODESTUM EURASIA Journal of Mathematics, Science and Technology Education, 2026, 22(3), em2792
ISSN:1305-8223 (online)
OPEN ACCESS Research Paper https://doi.org/10.29333/ejmste/18067

Exploring elementary pre-service teachers' instrumental genesis in designing AI chatbots

Nayoung Ku [1*]

[1] Gyeonggi Science High School, Suwon, SOUTH KOREA

Received 10 July 2025 ▪ Accepted 04 February 2026

Abstract

Based on an instrumental approach, this study explores how elementary pre-service teachers engage in instrumental genesis as they design artificial intelligence (AI) chatbots as teaching instruments for the "number and operations" domain. Eleven pre-service teachers participated in a prompt engineering lecture and developed chatbots. The findings revealed that the teachers assigned diverse instrumental functions to the AI chatbots, including supporting problem-solving, facilitating conceptual understanding, guiding problem-posing, and providing real-life application examples while actively incorporating scaffolding, bounded autonomy aligned with the curriculum, affective support, and gamification into their designs. Participants faced persistent issues in reconciling their instrumental intentions with the unpredictability and limitations of generative AI, particularly prompt interpretation and curriculum fidelity. These findings highlight the need for pre-service teacher education programs to offer reflective experiences with AI integration and foster adaptive design competencies, emphasizing the shift from technology adoption to empowering teachers as critical designers and mediators of AI instruments.

Keywords: artificial intelligence chatbot, instrumental genesis, instrumentalization, instrumentation, elementary pre-service teacher

INTRODUCTION

Artificial intelligence (AI) is driving significant transformations in mathematics education by reshaping teaching practices and learning processes. To adapt to these changes, teachers must acquire knowledge and skills related to AI tools, which should be systematically addressed in pre-service teacher education programs (Davis, 2025; Gattupalli et al., 2023). AI has a profound impact on instructional design and student learning outcomes (Liu et al., 2024) and plays a pivotal role in shaping beliefs and practical competencies of pre-service teachers (Yang & Chen, 2023). Despite this potential, the integration of AI into classroom contexts remains limited, revealing a gap between the willingness of pre-service teachers to use AI and their ability to implement it effectively. This discrepancy is often attributed to insufficient training and support provided by pre-service teacher education programs (Guan et al., 2025; Liu et al., 2024).

To address these challenges, pre-service teachers should be provided with opportunities to integrate AI into educational contexts, grounded in a comprehensive understanding of its concepts and technologies (Yang & Chen, 2023). Collaborating with AI in design instruction has emerged as a promising strategy for simultaneously enhancing beliefs about technology integration and developing practical competencies (Peura et al., 2021). Among the various AI applications, the potential of AI chatbots to contribute to pre-service teacher education has garnered significant interest, owing to their relatively rapid development through prompt engineering techniques (Papadakis et al., 2021). Specifically, these chatbots not only offer interactive and dynamic learning environments but also enable pre-service teachers to experiment with AI-driven instructional strategies in a controlled setting before applying them in real classrooms.

Nevertheless, as discussions surrounding AI technology in mathematics education have only recently gained traction, related research remains limited (Davis,

© 2026 by the authors; licensee Modestum. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/).
✉ guri39@gmail.com **(*Correspondence)**

Contribution to the literature

- This study makes a significant contribution to pre-service teacher education by addressing the noticeable gap in research on pre-service teachers' direct experience designing AI chatbots.
- This study extends instrumental approach by providing empirical cases of instrumentalization and instrumentation observed during pre-service teachers' design of AI chatbot-integrated lessons.
- This study offers practical insights into fostering design competencies in pre-service teacher education through AI-integrated lessons.

2025; Engelbrecht & Borba, 2024). While existing studies have highlighted the utility of AI chatbots in teacher education, they predominantly focus on scenarios where pre-service teachers guide chatbots to act as virtual students (e.g., Son et al., 2024). This approach relegates AI chatbots to function as assistant tools in pre-service teacher education and insufficiently addresses the exploration of pedagogical considerations in designing teacher-led AI-integrated instruction (Chiu et al., 2023). These limitations underscore the need for further research to examine how pre-service teachers can leverage AI as collaborative partners when designing innovative instructional practices tailored to specific educational contexts.

To investigate these complex interactions, this study adopts an instrumental approach as its theoretical lens. This approach provides a framework for analyzing the process by which users purposefully design AI chatbots to achieve specific instructional objectives, while the chatbot simultaneously shapes the users' thinking and practices (Drijvers et al., 2010; Trouche, 2004).

Prior studies have employed an instrumental approach to examine pre-service elementary teachers' use of geometry software (Ruiz-López, 2018), in-service teachers' engagement with digital platforms (Tamborg, 2021), and pre-service teachers' utilization of generative AI (Davis, 2025). However, research examining how pre-service teachers design AI chatbots for their own lessons remains limited. Therefore, this study investigates the process by which pre-service teachers design AI chatbots intended for implementation in authentic classrooms.

Specifically, by using qualitative methods, it examines the instrumental functions assigned to AI chatbots by pre-service teachers, the instrumental design considerations they incorporate, and the instrumentation issues they encounter during the design process. This study aims to bridge existing research gaps by shifting the focus from viewing AI chatbots as passive instruments to conceptualizing them as active collaborators within a teacher-chatbot partnership framework. The significance of this study lies in its empirical illustration of the process by which pre-service teachers design AI chatbots, utilizing an instrumental approach.

LITERATURE REVIEW

AI Chatbots in Mathematics Education

AI chatbots are AI-based applications that can infer the meaning of human language and respond appropriately using natural language processing technology (Reshmi & Balakrishnan, 2016). They are recognized as conversational agents that provide users with generally accurate information (Talha Junaid et al., 2024). AI chatbots also offer immediate and personalized support to students, facilitating the learning process by providing prompt answers to questions. Interactive chatbot systems can motivate and engage students, thereby making learning more engaging and accessible. Consequently, AI chatbots have gained attention as virtual assistants that can support student teaching and learning while assisting teachers (Kuhail et al., 2023). Recent systematic reviews have highlighted the widespread integration of chatbots, particularly ChatGPT, across diverse educational contexts worldwide, demonstrating their potential to positively influence both cognitive aspects and affective dimensions (e.g., Nguyen & Pham, 2025).

The design of educational chatbots incorporates various principles that Kuhail et al. (2023) categorized as personalized learning, experiential learning, collaborative learning, learning by teaching, scaffolding, affective support, and social dialogue. Personalized learning refers to the principle of providing tailored learning paths or feedback based on individual students' learning styles and performance levels. Experiential learning enhances learning through reflection on students' experiences. Collaborative learning involves encouraging interaction among students to complete tasks or solve problems. Learning by teaching occurs when students offer offering examples and feedback to chatbots. Scaffolding provides temporary structures or support to help students successfully complete tasks (Winkler et al., 2020). Affective support focuses on offering emotional feedback to enhance student motivation for learning. Finally, social dialogue involves engaging in nonacademic conversations, such as discussing the weather or sharing preferences, to build a relationship between the chatbot and the student. These seven design principles offer guidelines for meaningfully implementing educational chatbots.

The use of AI chatbots in mathematics education has advanced in recent years (Engelbrecht & Borba, 2024; Kuhail et al., 2023), and extant research has demonstrated their mathematical effectiveness and value. Chatbots have been shown to have significant effects not only on cognitive aspects such as problem-solving (Kabiljagić et al., 2022; Matsuda et al., 2013; Nguyen et al., 2019), conceptual understanding (Faqih et al., 2022; Melo et al., 2020), and mathematical thinking (Song et al., 2025) but also on affective aspects such as learning motivation (Laksana & Fiangga, 2022). For example, the chatbot discussed by Matsuda et al. (2013) facilitated learning by presenting equations and prompting students at each step required to solve them, whereas Nguyen et al. (2019) developed an interactive intelligent tutoring chatbot for mathematics learning that acted as an agent, providing step-by-step guidance to support problem-solving. More recently, generative AI has been recognized as an innovative tool for mathematics teaching and learning, capable of generating diverse responses, assisting in lesson planning, and providing personalized feedback (Almarashdi et al., 2024; Wardat et al., 2023). Therefore, chatbots can serve as agents to achieve various objectives in mathematics teaching and learning. However, it is essential to examine how teachers who implement chatbots in their lesson designs evaluate their advantages and limitations (Engelbrecht & Borba, 2024).

AI chatbots have thus garnered attention as tools for enhancing instructional competencies of pre-service teachers in mathematics education. Son et al. (2024) proposed a dual-agent chatbot structure that allowed pre-service teachers to interact with virtual students while simultaneously receiving feedback from a virtual mentor, thereby improving their responsive teaching skills. Lee and Yeo (2022) demonstrated that AI-based chatbots offering open-ended interactions enabled pre-service teachers to evaluate teaching scenarios critically and formulate independent responses, thereby enhancing their reflective thinking and creativity. Lee et al. (2025) demonstrated that technology-based educational simulations using chatbots enabled pre-service teachers to practice teaching skills without the presence of students, leading to significant improvements in questioning techniques. Cooper (2023) emphasized that effective use of AI-based chatbots depends on educators' critical thinking and ethical considerations, highlighting the importance of teacher education in preparing educators to integrate AI tools responsibly.

However, existing studies have focused primarily on the limited use of AI chatbots in roles such as virtual students or mentors, with insufficient attention being paid to analyzing the process by which pre-service teachers design the chatbots themselves. As noted by Holmes et al. (2019), AI technologies can perform diverse roles, and chatbots have the potential to support teachers and students in various ways. Analyzing the chatbot design process can provide a deeper understanding of their thought processes and educational intentions. This approach distinguishes itself from prior research by emphasizing autonomous chatbot design experiences that reflect the needs of educational stakeholders rather than merely focusing on technological usage. Such experiences can enhance teachers' creative instructional competencies.

Instrumental Approach

With the rapid evolution of educational settings, the current resources that teachers draw upon range from traditional texts to diverse digital resources. The instrumental approach, grounded in Rabardel's (1995) theory of instrumented activity, has gained prominence alongside the proliferation of digital resources. For an artifact such as a computer or calculator to support mathematics learning, it must be transformed into an instrument through the subject's agency (Trouche, 2004).

The gradual transformation of an artifact into an instrument, under the influence of the subject and instrumented action schemes, is known as instrumental genesis. Instrumental genesis comprises two interrelated processes that occur between the subject and the artifact. The first, instrumentalization, proceeds from the subject to the artifact: the subject selects and adapts the artifact to meet personal needs and draw on prior knowledge. This process is rooted in the artifact's intrinsic characteristics and can lead to functions that diverge from the designer's original intentions. The second, instrumentation, flows from the artifact to the subject: the artifact's affordances and constraints shape the individual's thinking and actions (Gueudet et al., 2021; Trouche, 2004). The instrumental approach has evolved over time to encompass various dimensions of technology integration in mathematics education, emphasizing both how users shape tools and how tools shape users (Sinclair et al., 2022).

The instrumental approach highlights two intertwined dimensions:

(a) shaping a tool in accordance with specific educational goals (instrumentalization) and

(b) understanding how the tool, in turn, shapes thinking and practice (instrumentation).

Several studies have examined how teachers instrumentalize and are instrumented by digital artifacts in mathematics education. Ruiz-López (2018) analyzed pre-service teachers' construction tasks in GeoGebra and found that they focused primarily on pragmatic, problem-solving functions while underutilizing epistemic functions that promote conceptual understanding. Tamborg (2021) showed that in-service teachers experienced interface-related constraints when using digital platforms, partially fulfilling designers' intentions while encountering difficulties, suggesting

that design must be continuously adjusted during use. Recently, Davis (2025) reported that pre-service teachers who planned lessons with ChatGPT lacked a repertoire of techniques—that is, the observable actions and prompts users employ when interacting with large language models— for using generative AI to develop instructional materials. Collectively, these studies shed light on how teachers instrumentalize digital artifacts and how those, in turn, instrument teachers. However, they do not address how teachers design new AI instruments for instructional purposes. Little is known about the functions that pre-service teachers assign to AI chatbots, the instrumental design considerations they embed during design, and how the tool itself reshapes teachers' thinking and practice.

To address this gap, this study examines instrumentalization and instrumentation in the context of pre-service teachers' chatbot design processes. Specifically, it investigates the instrumental functions that elementary pre-service teachers assign to AI chatbots and the design considerations they incorporate during development. This research also explores technical and educational constraints that shape and limit their thinking and activity during chatbot design, thereby constituting the instrumentation issues they encounter. To guide this investigation, the following research questions are proposed:

1. What instrumental functions do pre-service teachers assign to AI chatbots in classroom instruction?
2. What instrumental design considerations do pre-service teachers consider when designing AI chatbots?
3. What instrumentation issues do pre-service teachers encounter during the AI chatbot design process?

METHOD

Study Design

Case studies are an effective approach for exploring specific situations or phenomena to derive insights and understand key issues (Merriam & Tisdell, 2015). This study adopted a case study methodology to investigate the considerations and issues faced by pre-service elementary teachers during the chatbot design process.

This study was conducted with 11 pre-service elementary teachers (nine female and two male) enrolled in an advanced mathematics education course at a national university of education in South Korea in 2024. The course aimed to equip participants with fundamental knowledge and skills needed to integrate technological tools into mathematics education and develop their teaching competencies using these tools. The pre-service teachers had prior knowledge of mathematics education theories and the national curriculum through other courses, as well as five weeks of field-teaching practice experience. A typical yet convenient sample was selected (Merriam & Tisdell, 2015).

In the course, a four-week project on AI chatbots was conducted. During the first week, a lecture on generative AI for elementary mathematics instruction was presented. The lecture covered the concepts of generative AI, prompt engineering techniques (zero-shot, few-shot, chain-of-thought (CoT), and structured prompt (Oh, 2023; Sahoo et al., 2024), and differences in output results based on various prompt engineering techniques for solving mathematical problems. It also introduced prior research on prompt engineering, the performance of generative AI in solving problems, and its application in posing and evaluating problems. Pre-service teachers participated in a hands-on session to design chatbots using Wrtn (Wrtn Technologies, 2024; https://wrtn.ai/), a free generative AI platform developed in South Korea. Wrtn provides templates tailored to the domestic educational context and offers a user-friendly interface for beginners to design chatbots, quizzes, and other tools. Unlike other generative AI platforms (e.g., ChatGPT and Gemini), Wrtn can be used by elementary school students with parental consent. Therefore, Wrtn was selected in this study as the platform for pre-service teachers to design chatbots.

Wrtn's chatbot design process consists of four steps. In the first step, users input basic information such as chatbot descriptions and category selection. In the second step, users configure the chatbot's initial message and example questions. The third step involved composing prompts. While Wrtn allowed developers to choose whether to use template-based prompts depending on their expertise level, the participants in this study, as beginners, were instructed to use the template format. The template was divided into sections for role, personality, information, and requirements, with a maximum input limit of 2,000 tokens (see **Figure 1**). The fourth step involved testing conversations with the AI chatbot. Wrtn supports GPT-3.5, GPT-4, and GPT-3.5-16K models; for this study, GPT-4 was used for all interactions.

After the first week, participants (PST1-PST11) designed their chatbots individually. During weeks 2 and 3, they conducted simulated lessons using chatbots and exchanged feedback with one another. In the fourth week, they revised their chatbots based on the feedback they received.

This study focused on an in-depth discussion of the considerations of and issues faced by pre-service elementary teachers during chatbot design by limiting the content area to "number and operations" for grade 5 and grade 6. Compared with lower grade levels, these grades require less basic guidance for utilizing chatbots.

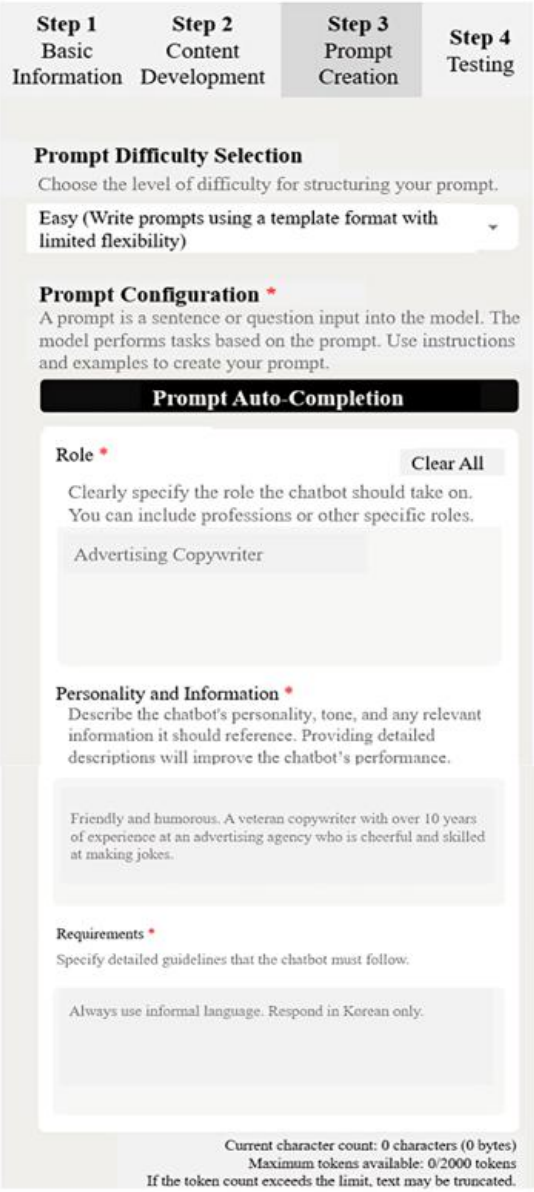

Figure 1. Prompt creation step in Wrtn (translated into English) (Source: Author's own elaboration)

Additionally, as ChatGPT's performance can significantly decline when solving problems involving extended basic algebraic calculations (Schorcht et al., 2024), limiting the content area to "number and operations" was intended to allow participants to engage in prompt engineering without concerns about the performance of the generative AI.

Data Collection and Analysis

Participants were assigned an AI chatbot design task. They submitted the chatbot design assignment twice, after week 1 and after week 4. As the purpose of this study was to explore what pre-service teachers expect from chatbots and the challenges they encounter during the design process, the analysis focused on their initial chatbot design. The assignment consisted of three parts: a written lesson plan, chatbot prompts, and a reflective report on the chatbot design.

First, participants were asked to create a written lesson plan based on curriculum materials. The template included sections on the topic, achievement standards, learning goals, guidelines for teaching and learning, required materials, and three instructional phases (launch, investigate, and summarize) (Van de Walle et al., 2016).

Second, the participants were required to write a chatbot introduction, initial message, example questions, and prompts (role, personality, information, and requirements). All prompts and chatbot URLs were collected to verify how the chatbots were implemented.

Third, the participants were asked to write a reflective report on the chatbot design. This component was intended to encourage critical reflection on participants' experiences and decision-making during the chatbot development process. Responses to a set of structured questions, grounded in the theory of instrumental genesis, offering insights into the pre-service teachers' instrumentalization and instrumentation schemes (Davis, 2025; Tamborg, 2021). The former concerns how users shape the artifact, whereas the latter concerns how the artifact shapes the users. Regarding the first and second research questions, the participants were asked to describe their intentions to use the chatbot in this activity, the goals of the activity, their anticipated student responses, and explanations of their prompts. For the third research question, participants were asked to identify the limitations of their chatbots and the aspects of their designs that deviated from their intentions because of their technical skills.

The author analyzed 11 chatbots, focusing on the second and third parts of the assignment: chatbot prompts and reflective reports on the chatbot design. The analysis followed the guidelines for thematic analysis (Braun & Clarke, 2006). First, to become familiar with the data, the author read and reread the pre-service teachers' responses, examining them carefully to gain a deeper understanding of their content and context. Next, the author segmented the responses to the reflective report questions into identifiable units, such as key phrases or sentences related to the instrumental functions assigned to the chatbot, instrumental design considerations, and instrumentation issues, and then began the analysis. At this stage, coding categories relevant to the research questions were developed. As prior studies (e.g., Kuhail et al., 2023; Wardat et al., 2023) did not explicitly focus on AI chatbot design in mathematics education, their findings were considered difficult to apply directly. Therefore, the author inductively coded the responses of the pre-service teachers. For example, "If there is a chatbot, these students can ask it about difficulties they encounter while posing and solving problems under given conditions" (PST3) was coded under the chatbot's function as *instrument for problem-posing*.

Table 1. Instrumental functions of AI chatbots

Instrumental functions	Participants
Instrument for problem-solving	PST1, PST2, PST4, PST5, PST7, PST8, PST10, PST11
Instrument for problem-posing	PST3, PST6
Instrument for conceptual understanding	PST5, PST7
Instrument for application	PST4, PST9

Similarly, "The chatbot should only play a supporting role. (…) It should not provide entire solutions but show only what students need" (PST10) was coded under instrumental design considerations as *scaffolding*, emphasizing how chatbots are designed to guide students incrementally toward better understanding. During this process, the prompts were examined alongside the reflective reports to determine how instrumental design considerations were concretized in the prompts and what aspects of prompt composing or AI chatbot responses the pre-service teachers perceived as challenging. To enhance reliability of the qualitative analysis, an external expert specializing in AI applications in mathematics education independently coded the dataset for validation coding after the author completed the primary coding. The author's coding and the expert's validation coding were then compared in several meetings. Inter-coder agreement was 92%, and any discrepancies were resolved through discussion to reach consensus.

RESULTS

Instrumental Functions

The data analysis revealed that participants instrumentalized AI chatbots as mathematical instruments that performed four instrumental functions (problem-solving, problem-posing, conceptual understanding, and application). **Table 1** summarizes these instrumental functions and the PST codes assigning each function. PST5 and PST7 designed compound instruments that supported both problem-solving and conceptual understanding, whereas PST4 combined problem-solving with application.

Pre-service teachers who instrumentalized the chatbot for problem-solving designed it to provide explanations and adaptive assistance based on students' questions during the solution process. Their instrumental goal was to help students develop procedural skills and progress toward procedural fluency through instrumental interactions with the chatbot.

For example, PST10 designed a chatbot to be a problem-solving instrument for "mixed operations with natural numbers." Students first solved a word problem independently and then used the instrument to verify their procedures. The chatbot subsequently provided a slightly more challenging problem of the same type, promoting additional practice and application through instrumental mediation. PST10's instrumental design intention was as follows:

> PST10: Using the chatbot as an instrument allows students to verify whether their solution and answer are correct. Additionally, as students may encounter different difficulties during the problem-solving process, the chatbot provides personalized instrumental support.

PST5 and PST7 designed compound instruments that integrated problem-solving with conceptual understanding related to factors and multiples. Both participants designed their chatbots to function as though the chatbot was a learner, prompting them to explain the concepts of divisors and the greatest common divisors (GCD). PST5 explained that the instrumental intention was to design the chatbot to engage in a question-and-answer interaction with students, thus helping them structure the knowledge they had learned and address any gaps in their understanding through instrumental mediation:

> PST5: The activity should proceed in such a way that the student explains the concept they have learned to the chatbot and the chatbot "learns" the concept. I set the chatbot to learn mathematics alongside the student.

PST3 and PST6 designed chatbots to support problem-posing. They noted that many students struggled to create problems and that it was physically challenging for a single teacher to provide individual guidance to all students. Therefore, they designed the instrumental system in such a way that the chatbot acts like a human teacher, identifies students' difficulties, and provides immediate feedback. PST3 elaborated on her instrumental design intentions as follows:

> PST3: Many students face difficulties when posing problems, but it is physically impossible for one teacher to assist every student. (…) However, with a chatbot functioning as an instrument, even these students can ask the chatbot about the difficulties they encounter while posing and solving problems under given conditions.

PST4 and PST9 designed chatbots to help students understand how mathematical knowledge can be applied in real-life contexts. Both participants specifically intended to utilize these instruments in lessons on "rounding rules." The teachers noted that

Table 2. Instrumental design considerations of AI chatbots

Instrumental design considerations	Participants
Scaffolding	All
Affective support	PST1, PST2, PST4, PST5, PST6, PST7, PST9, PST10, PST11
Bounded autonomy for teaching	PST1, PST2, PST4, PST5, PST6, PST7, PST8, PST10, PST11
Gamification	PST1, PST2, PST8

students often find it challenging to identify real-life examples in which rounding rules are applied. Therefore, they designed a chatbot to answer students' questions by providing relevant examples or verifying whether students' examples were accurate through instrumental mediation. PST9's instrumental design intention was as follows:

> PST9: I decided to use a chatbot because I found that students struggle to independently recall instances where rounding up, rounding down, or rounding off is used in real-life contexts.

Instrumental Design Considerations

During the chatbot design phase, participants identified four instrumental design considerations—scaffolding, affective support, bounded autonomy for teaching, and gamification—when drafting their prompts. To incorporate these considerations into the emerging instrument, they drew on prompt engineering techniques such as one-shot, few-shot, CoT, and structured prompt formats (see **Table 2**).

All participants devoted significant effort to designing their chatbots to provide scaffolding. In other words, they intended the chatbot to be an instrument that delivers tailored instructions and feedback based on students' questions. Specifically, they were cautious to ensure that the AI chatbot would not immediately provide complete solutions to problems or fully explain concepts, without allowing students time to think. Instead, they designed the instrumental interaction to offer incremental hints to facilitate students' cognitive development and gradually reduce support as their skills improved.

Two instrumentalization approaches regarding how the participants designed their AI chatbots for scaffolding were identified. The first approach involved setting criteria for assessing the level of student responses and then providing example responses based on those levels. The second approach involved anticipating questions that students at various levels might ask and providing examples of how the chatbot should respond to these questions. Representative examples of each approach are presented below.

PST5 (following the first approach) designed her chatbot to engage in question-and-answer interactions with students about divisors and GCD to support conceptual understanding. She embedded instrumental schemes in the chatbot's reasoning process using five criteria to ensure that it provided feedback appropriate to the students' level.

As shown in **Figure 2**, she input examples of responses for three levels by analyzing whether students used mathematical terminology correctly and whether their explanations were logically structured. In addition, when errors were detected in a student's response, the chatbot was programmed to ask follow-up questions to guide the students in understanding the concept. To facilitate this process, she used a one-shot prompt to help the chatbot interpret her design intentions. The instrumental design considerations shared by PST5 and the corresponding instrumentalization implementation in the prompt are as follows:

> PST5: I established five criteria for assessing the level of students' responses by analyzing their mathematical expressions rather than simply evaluating correctness. The five criteria include explanation analysis, question elicitation, feedback, problem-solving, and connectivity. In the explanation analysis, the focus is primarily on determining whether students use appropriate mathematical terminology in their explanations and whether their reasoning demonstrates logical consistency and coherence.

PST3 (following the second approach) designed a chatbot to assist with problem-posing. The prompt was structured into the following components: <Context>, <Given Conditions>, <Possible Problems>, <Examples>, <Solutions>, <Instructions>, and <Output Rules>. PST3 composed a structured prompt that embedded instrumental schemes by instructing the chatbot to consider <Context> and refer to <Given Conditions>, <Possible Problems>, <Examples>, and <Solutions> before responding according to the <Instructions>. **Figure 3** illustrates a part of the content included in the <Instructions>. PST3 also designed the instrumental mediation to guide students in posing problems that met specified conditions. She inserted four anticipated student questions (#1-#4) and instrumented reply to templates for each. For instance, #1 was designed for when a student says, "It's hard to select school supplies that meet the conditions," while #4 was for when a student says, "It's difficult to pose a new problem that satisfies the conditions." PST3 emphasized that examples of problems should be provided only in response to such questions and explicitly instructed the chatbot not to offer examples beforehand. The instrumental design considerations shared by PST3ofnd

[Role]

(...) When the learner does not understand, guide them to think independently. If the learner explains the mathematical concept they learned today, proceed to the next step:

Step 1: Assess the learner's level based on {Level Assessment Criteria}.

Low level: "I'm having trouble understanding. Could you explain it again using mathematical terms?"

Intermediate level: "That was a good explanation! Is there anything else you'd like to add?"

High level: "Excellent explanation! Could you share some related concepts as well?"

{Level Assessment Criteria}

Explanation Analysis:

Analyze the terms and logic used by the learner when explaining mathematical concepts. Assess their level based on whether they use mathematical terminology and whether their explanation is logically consistent.

Question Elicitation:

When the learner reveals gaps or misunderstandings in their explanation, ask follow-up questions to guide them toward deeper understanding. However, the chatbot should not provide direct answers but should instead offer hints.

(Example) "How many numbers can you use to find the GCD?" or "Could you explain with an example?"

Figure 2. Excerpt from PST5's prompts (translated into English) (Source: Author's own elaboration)

<Instructions>

#1. If the user says, "It's hard to choose school supplies that meet the conditions," you must provide [Response Content].

[Response Content]

You should select school supplies in such a way that their total cost does not exceed 10,000 won. Since you need to choose multiple types of school supplies, try naming at least two types that you want. Then, tell me how many of each type you'd like to buy. I'll check if your choices meet the conditions. (...)

#4. Only when the user says, "It's difficult to pose a new problem that satisfies the conditions," should you provide [Response 2] along with <Possible Problems>. Do not offer help prematurely before this point.

[Response 2]

I see! Let me show you what kind of problems you can pose. They're the same as <Possible Problems>. If you're still unsure, feel free to ask again! I'll share some examples with you.

<Possible Problems>

1. Create a problem about calculating the total cost of the school supplies Ji-hye selected. (...)

<Examples>

1. "Ji-hye bought 5 notebooks and 4 erasers and rulers each. What is the total cost of the school supplies Ji-hye purchased?" (...)

Figure 3. Excerpt from PST3's prompts (translated into English) (Source: Author's own elaboration)

the corresponding instrumentalization implementation in the prompt are as follows:

> PST3: I wrote the prompt with the intention that the chatbot would respond thoughtfully, taking everything into account. In the <Instructions> section of the prompt, I aimed to carefully control the chatbot's dialogue.

Ten participants (excluding PST8) emphasized the importance of providing affective support when interacting with elementary students. PST6 embedded

<Personality and Information>

1. Respond in a kind and friendly manner.

2. Avoid using slang or informal language under any circumstances.

3. Encourage students by reassuring them that it's okay not to know something, and help them continue their learning journey.

Figure 4. Excerpt from PST6's prompts (translated into English) (Source: Author's own elaboration)

[Requirements]

4. When interacting with the user, all content must align with the elementary school level, and all concepts and mathematical symbols used in the conversation must fall within the scope of South Korea's elementary school curriculum. (...)

5. Respond based on the following achievement standards and guidelines from the elementary school "Number and Operations" curriculum:

[6M01-02] Understand and use the meanings of "greater than or equal to," "less than or equal to," "greater than," and "less than" in connection with real-life situations, and represent numerical ranges using these terms.

[6M01-03] Understand the meanings and necessity of rounding up, rounding down, and rounding off as methods for estimating values, and recognize the usefulness of mathematics by applying these concepts to real-life situations.

The concepts of numerical ranges, rounding up, rounding down, and rounding off should be understood through various real-life examples where estimation is required, such as measurement scenarios involving numbers or quantities.

6. Never use the term "integer." If a situation arises where you would use "integer," replace it with "number."

Figure 5. Excerpt from PST4's prompts (translated into English) (Source: Author's own elaboration)

instrumental schemes to ensure that the chatbot used a friendly and polite tone and provided emotional support to the students to enhance their motivation to learn (see **Figure 4**).

Ten participants (excluding PST10) emphasized that the chatbot should use the terminology outlined in the elementary school curriculum and avoid responding with symbols or terms beyond the elementary level. South Korea's education system has historically been highly centralized and guided by the national curriculum system; teachers are known to respect its authority and incorporate it into their lessons. Similarly, participants considered adherence to the curriculum when designing their AI chatbots as instruments. Because their instrumentalization was constrained by organizational requirements to adhere to the curriculum, they input content from curriculum documents into prompts such that the chatbot followed these guidelines and explained mathematical concepts and symbols within the scope of the elementary curriculum. A representative example was provided by PST4.

PST4 designed a chatbot to provide feedback when students estimated and inquired about quantities that they might encounter in real-life situations. She embedded achievement standards and symbol guidelines (see **Figure 5**) and instructed the chatbot to avoid terms outside the elementary curriculum. She also highlighted limiting decimal places to those allowed in the curriculum, followed by rounding rules at elementary level.

> PST4: I made an effort to ensure that terms or concepts beyond the scope of the elementary school curriculum were not mentioned. (…) I provided both the achievement standards and their corresponding guidelines, so that they could be referenced (#5). Since integers are taught at middle school level, (…) I explicitly instructed it to use "number" instead of "integer" (#6).

Three participants (PST1, PST2, and PST8) incorporated gamification into their instrumentalization process to make the interactions more engaging and enjoyable for the students. All three chatbots were designed to assist students in problem-solving. PST1 created a number-guessing game through instrumental design, whereas PST2 and PST8 set target scores or problem levels for students to achieve, revealing secret codes upon completing all levels. For example, PST8 provided a narrative where students had to find a "number island" by uncovering a secret code (see **Figure 6**). This storyline encouraged the students to immerse

When the user says, "Give me a Level 1 problem," the chatbot presents four number cards. Among these four cards, two are zeros, and the other two are numbers between 1 and 9. After presenting the four number cards, the chatbot says:

"Step 1. From the four number cards, select two cards to create a decimal less than 1. (Note: The number you create cannot be 0.0.)."

Once the student inputs their answer, the chatbot continues:

"Step 2. Using the remaining two cards, create another decimal less than 1."

After the student provides their second answer, the chatbot says:

"Step 3. Multiply the two numbers you created."

If the student cannot provide the correct answer, the chatbot explains the two solutions step by step as previously mentioned. For students who answer correctly, the chatbot responds:

"The first letter of the code is '수' (Korean character)."

Figure 6. Excerpt from PST8's prompts (translated into English) (Source: Author's own elaboration)

Table 3. Instrumentation issues reported by pre-service teachers

Instrumentation issues	Representative examples	Participants
Limited problem diversity	Chatbot only generated problems from the small set of examples provided, failing to create varied tasks	PST1, PST7, PST11
Curriculum non-compliance	Chatbot used terminology beyond the elementary curriculum (e.g., "integer") despite explicit instructions	PST4, PST5, PST6, PST8, PST9
Limited scaffolding	Chatbot ignored prompts to give partial hints or intentional errors and instead supplied full, correct solutions	PST2, PST4, PST5, PST6, PST9
Unpredictable out-of-scope responses	Chatbot produced answers that diverged from teacher intent when confronted with questions not covered in the prompt	PST3, PST10

themselves in the activity, solve repeated multiplication problems involving decimals at each level, and challenge themselves until the end. The instrumental design considerations shared by PST8 and the corresponding instrumentalization implementation in the prompt are as follows:

> PST8: I structured the activity to feel like a game by incorporating a story and rewards. The narrative of finding a number island provided a sense of immersion, and collecting letters for the secret code to complete the final message gave students a sense of accomplishment.

Instrumentation Issues

Drawing on their AI chatbot design experiences, the pre-service teachers reported two broad categories of instrumentation difficulties: misalignment between intended instrumentalization and actual chatbot output, and unpredictable responses to unprompted queries. Typical examples of such misalignments included the chatbot generating problems with limited diversity, employing terminology beyond the curriculum, and providing complete solutions rather than instrumental scaffolding (see **Table 3**).

Three participants (PST1, PST7, and PST11) identified the limited diversity of problems as a significant instrumentation limitation. For instance, PST7 hoped the chatbot would help learners solve a variety of real-life problems involving multiples and common multiples. Although she instructed the chatbot to generate context-rich examples only upon request, it repeatedly returned the same scenarios—"event schedules," "transportation timetables," and so forth—thereby limiting students' exposure to diverse problem types (see **Figure 7**). The instrumentation issue reported by PST7, along with their prompts and outputs, is as follows:

> PST7: It was disappointing that similar questions elicited common answers. When students asked for real-life problem examples, they received identical responses.

Five participants (PST4, PST5, PST6, PST8, and PST9) encountered instrumentation issues when the chatbot used terminology and explanations beyond the elementary school curriculum. PST9 specifically prompted her chatbot to adhere to the elementary curriculum, instructing it to avoid the term "integer." However, during instrumental testing, the chatbot contradicted these embedded constraints by explaining that candies must be purchased in "integer" quantities (see **Figure 8**). She expressed dissatisfaction with the output:

> PST9: The chatbot used the term "integer," which is not taught in elementary school. Even though I

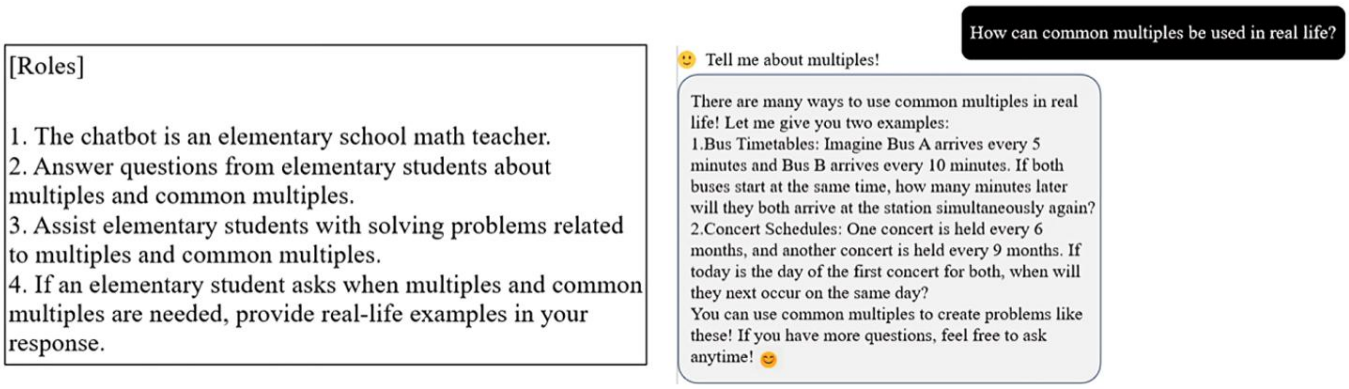

Figure 7. Excerpt from PST7's prompts and outputs (translated into English) (Source: Author's own elaboration)

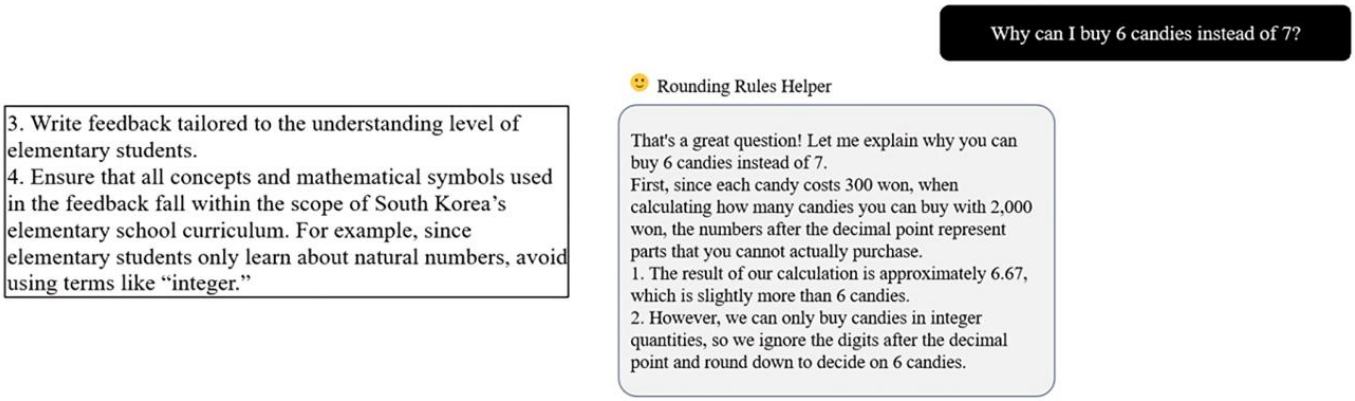

Figure 8. Excerpt from PST9's prompts and outputs (translated into English) (Source: Author's own elaboration)

> explicitly stated in the prompt that this term should not be used, it still appeared in the response.

Five participants (PST2, PST4, PST5, PST6, and PST9) described instrumentation issues: the chatbot ignored instructions to give partial or intentionally incorrect answers. PST2 designed a game-based chatbot to provide incorrect solutions 50% of the time, preventing student disengagement. However, the chatbot consistently delivered correct answers, indicating resistance to instrumentalization due to Wrtn's prompt capacity constraints and instrumental design complexities.

Similarly, PST5 embedded instrumental schemes requiring hint-based rather than solution-based responses when students' GCD explanations were insufficient (see **Figure 9**). However, she identified a limitation in that the chatbot was unable to fully understand the prompt instruction to not provide answers but instead scaffold instrumental support. Consequently, the chatbot solved the entire problem, as illustrated in **Figure 9**.

> PST5: The chatbot did not adequately recognize the instruction to explain the student's errors without providing answers but offer hints instead.

Two participants (PST3 and PST10) noted instrumentation issues stemming from the chatbot's unpredictable responses when confronted with questions outside the prompt scope. PST3 used <Context> prompting to instrumentalize the chatbot for problem-posing assistance yet encountered unexpected responses that often drove the dialogue away from her instructional goals.

> PST3: For situations not included in the prompts, the chatbot tended to respond in ways that differed from my intentions. Therefore, I felt that no matter how carefully the prompt was composed, if users provided unexpected responses, the conversation was likely to take a different direction.

These findings highlight the complex instrumental genesis process in which instrumentalization strategies encounter systemic instrumental constraints, revealing tensions between teachers' instrumental intentions and the inherent limitations of the chatbot artifact during transformation into a mathematical instrument.

DISCUSSION

Instrumental Functions

Our analysis revealed that pre-service teachers instrumentalized AI chatbots to perform four distinct instrumental functions: problem-solving, conceptual understanding, problem-posing, and real-life application. This finding extends Wardat et al.'s (2023)

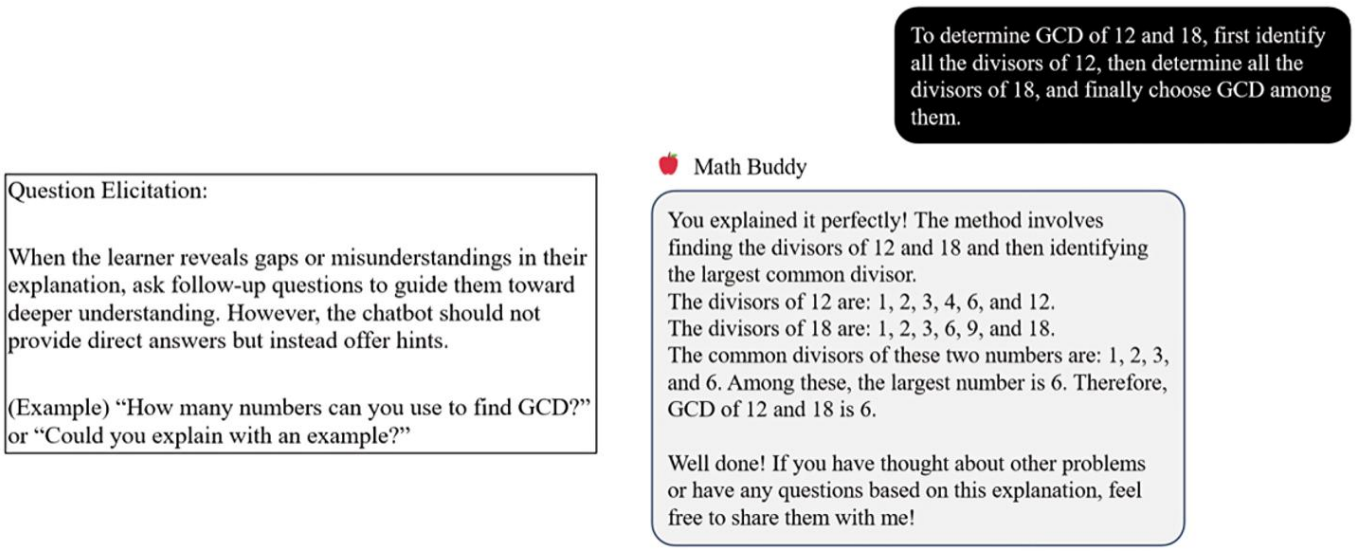

Figure 9. Excerpt from PST5's prompts and outputs (translated into English) (Source: Author's own elaboration)

observation that dialogue-based AI can serve multiple pedagogical roles by demonstrating how teachers actively construct these roles through deliberate instrumentalization rather than accepting pre-configured capabilities.

The predominant emphasis on problem-solving within the "number and operations" domain echoes the pragmatic bias documented in Ruiz-López's (2018) study, wherein pre-service teachers similarly prioritized procedural support over epistemic exploration. However, as conversational agents, chatbots are inherently limited in their ability to provide interactive environments for the manipulation of various mathematical objects. Consequently, their use appears to have focused primarily on problem-solving instructions rather than exploratory or manipulative tasks.

Meanwhile, the approach taken by PST5 and PST7, wherein the chatbot was framed as a "learner," is noteworthy. By positioning students as teachers who explain divisors and multiples to a "novice" chatbot, the participants achieved a more sophisticated instrumentation that prompted knowledge restructuring rather than mere answer verification. This learning-by-teaching function demonstrates how instrumental genesis can transcend traditional procedural boundaries when teachers deliberately design for epistemic mediation.

These findings underscore the importance of instrumentalizing AI chatbots as an educational instrument within pre-service teacher education programs focused on mathematics teaching and learning. Through engaging in instrumentalization practices that explore chatbots' diverse instrumental functions in teaching and learning contexts, pre-service teachers can concretize and internalize strategies for utilizing AI in student-centered instruction. According to Guan et al. (2025), although pre-service teachers generally exhibit positive attitudes toward AI integration, they continue to face challenges in designing and implementing lessons that effectively combine pedagogical approaches with AI tools. Furthermore, given the transformative potential of AI technologies to redefine teacher-student interactions, it is essential for pre-service teacher education programs to provide concrete examples and hands-on practice opportunities. Such experiences can help future teachers navigate the complexities of instrumental genesis, ensuring effective and contextually appropriate integration of AI chatbots as mathematical instruments in educational settings.

Instrumental Design Considerations

This study demonstrates that pre-service teachers actively instrumentalized chatbots by embedding four instrumental design considerations—scaffolding, bounded autonomy, affective support, and gamification—into prompt scripts. While perception-oriented studies (Gattupalli et al., 2023; Sanusi et al., 2024) have reported only teachers' "preferred" step-by-step explanations, this study provides empirical evidence of how pre-service teachers concretely instrumentalize AI chatbots. This extends Davis' (2025) research on secondary pre-service teachers' use of generative AI in lesson planning by providing insights into how AI chatbots can be utilized in elementary mathematics instruction.

Notably, despite the inherent inaccuracies of generative AI (Wardat et al., 2023), the pre-service teachers instrumentalized scaffolding as a central instrumental design. Unlike Gattupalli et al. (2023), who primarily highlighted pre-service teachers' perceptions of generative AI's ability to deliver step-by-step explanations, this study revealed how scaffolding was operationalized as a deliberate instrumentalization strategy: participants embedded graded hint-giving instrumented schemes that balanced incremental guidance with opportunities for autonomous cognitive

leaps. Instead of merely asking the chatbot to display full solutions, they scripted prompts that steered students through a sequence of instrumented actions, establishing response-prediction criteria and fine-grained steps that simulated classroom praxeology.

Many pre-service teachers also considered bounded autonomy by aligning their chatbot designs with national curriculum documents and elementary-level concepts and symbols. This reflects an awareness of the need for teachers in a centralized education system to balance "pedagogical freedom" with "curriculum compliance." In practice, some participants directly input achievement standards and guidelines into their chatbot prompts to ensure that the generated responses adhered strictly to these boundaries. As noted by So et al. (2024), while teacher autonomy has been gradually expanding because of national curriculum reforms, this study reaffirms that pre-service teachers tend to adhere strictly to grade-specific curricular boundaries, highlighting their cautious approach to balancing innovation with compliance.

The teachers also integrated affective mediation and gamification as complementary instrumental design considerations, aligning with Talha Junaid et al. (2024), who argue that technology alone cannot satisfy learners' socio-emotional needs. Empathetic tone directives and game narratives were scripted to widen the instrument's mediation from purely pragmatic to motivational and relational functions.

In summary, the pre-service teachers' instrumental genesis drew on four key considerations that demonstrated a shift from routine tool adoption toward transformative instrumental competences. However, this study did not evaluate the learning effectiveness of the chatbots developed by participants, nor did it examine the long-term impact of prompt engineering training, leaving these crucial areas for future research.

Instrumentation Issues

Despite deliberate instrumentalization, four persistent instrumentation issues emerged: limited problem diversity, curriculum non-compliance, limited scaffolding, and unpredictable out-of-scope responses. These issues reveal the complex instrumental constraints inherent in AI-mediated teaching. Notably, even with the automation and high-quality response generation characteristics of AI chatbots, achieving alignment between instrumental intention and performance remains challenging. This problem is particularly pronounced because reasoning about students' knowledge and behaviors inherently involves uncertainty (Van Vaerenbergh & Pérez-Suay, 2022), and participants faced a persistent dilemma between their pedagogical instrumentalization goals and the unpredictability of chatbots operating on the basis of probabilistic reasoning. Taken together, these instrumentation issues not only mirror the "command" and "organizational" constraints identified in the instrumental-approach literature (Rabardel, 1995; Trouche, 2005) but also extend them to the realm of probabilistic language models, whose runtime inference magnifies uncertainty, in turn, accentuates these very constraints.

These alignment challenges became particularly evident in specific situations wherein pre-service teachers aimed to explore student errors or provide step-by-step hints, as the chatbot produced complete answers or used terminology beyond the scope of the curriculum. These findings reaffirm Stefanova and Georgiev's (2024) assertion that the effective use of AI in mathematics education requires both curriculum-based oversight and innovative instrumentalization strategies. Further, instances in which pre-service teachers intentionally attempted to induce errors in the chatbot but were unsuccessful because of the chatbot's inability to recognize such prompts further underscored the need for adaptive instrumental systems capable of responding flexibly to diverse learner needs.

These challenges stem not only from the intrinsic instrumental constraints of AI but also from the limited instrumentalization expertise of pre-service teachers. Some participants reported that, despite structuring their prompts, they continued to receive unexpected responses because of both the constraints of the chatbot and their own insufficient prompt engineering skills. This supports the findings of Peura et al. (2021), who emphasized that, without a comprehensive understanding of AI mechanisms and the ability to integrate learning objectives, curriculum standards, and learner characteristics, the gap between instrumental intention and actual instrumentation may widen significantly.

Ultimately, the dilemmas experienced by pre-service teachers suggest that instrumentalizing AI chatbots into teaching and learning is not simply a matter of adopting convenient technological tools but rather a complex practice that requires ongoing negotiation between instrumental constraints and pedagogical intentions. While AI chatbots offer new educational opportunities, they can unexpectedly conflict with teachers' instructional goals. Therefore, pre-service teacher education programs must provide targeted support for reflective instrumentalization processes and create environments in which teachers can anticipate and address the imperfections in AI chatbots through flexible design and continuous monitoring. For example, the prompts provided to chatbots should be iteratively refined with the recognition that student questions or learning situations in real classrooms may deviate from initial predictions, requiring teachers to adapt their instrumentalization strategies accordingly. Additionally, teachers should not passively accept AI-generated responses but instead guide students to

critically interpret and reconstruct chatbot answers through metacognitive activities. Such an approach can help prevent uncritical reliance on AI technologies. Ultimately, in the era of AI, teachers will need to develop competencies to delineate the scope of tasks performed by machines and the points at which their pedagogical expertise is essential, ensuring the balanced and effective instrumental orchestration of both elements.

CONCLUSION

Despite offering valuable insights, this study has several limitations that should be considered within the framework of the instrumental approach. First, AI chatbots developed by pre-service teachers were not empirically validated in authentic classroom settings. Consequently, actual instrumental performance of these chatbots in real learning environments and the potential issues arising from student interactions remain unexplored. Future research should incorporate longitudinal studies that track the instrumental genesis of the designed chatbots in authentic classrooms, enabling a comprehensive examination of both the pedagogical and systemic challenges of AI-integrated instruction.

Second, the scope of learning topics addressed by the chatbots was limited to the "number and operations" domain. Expanding the chatbot design to other mathematical subdomains, such as geometry or probability, could yield richer insights into the diverse instrumental functions and design considerations that pre-service teachers employ. Subsequent investigations should explore cross-disciplinary chatbot applications to elucidate how pre-service teachers adapt AI artifacts as instruments to meet varied curricular demands and learner needs.

Third, this study was constrained by its reliance on the GPT-4 model, which inherently limited the exploration of prompt engineering variations. Given that chatbot response quality and pedagogical utility are model-dependent, the generalizability of the findings across different AI architectures remains uncertain. Future work should conduct cross-model comparisons to identify optimal prompt design frameworks that align with specific educational objectives and support effective instrumental orchestration.

Funding: No funding source is reported for this study.

Acknowledgments: The author would like to thank an external expert in mathematics education for assistance with the inter-coder reliability check and constructive feedback on the coding validation process.

Ethical statement: The author stated that the research was conducted in accordance with all applicable ethical procedures. Formal ethical approval was not required because the study was conducted as part of regular educational practices. The participants were fully informed about the research purpose and methods. Informed consent was obtained from all participants, and the data was collected anonymously.

AI statement: The author stated that generative AI tools (e.g., Gemini, 2.5 Pro, July 2025 version) were used solely to check the English language clarity of the manuscript. No content generation was performed by AI.

Declaration of interest: No conflict of interest is declared by the author.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the author.

REFERENCES

Almarashdi, H. S., Jarrah, A. M., Abu Khurma, O., & Gningue, S. M. (2024). Unveiling the potential: A systematic review of ChatGPT in transforming mathematics teaching and learning. *Eurasia Journal of Mathematics, Science and Technology Education, 20*(12), Article em2555. https://doi.org/10.29333/ejmste/15739

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Chiu, T. K., Moorhouse, B. L., Chai, C. S., & Ismailov, M. (2023). Teacher support and student motivation to learn with artificial intelligence (AI)-based chatbot. *Interactive Learning Environments, 32*(7), 3240-3256. https://doi.org/10.1080/10494820.2023.2172044

Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology, 32*(3), 444-452. https://doi.org/10.1007/s10956-023-10039-y

Davis, J. D. (2025). Evolving techniques and emerging schemes: Prospective teachers' transformation of ChatGPT. *International Journal of Science and Mathematics Education, 23*, 2181-2198. https://doi.org/10.1007/s10763-024-10533-8

Drijvers, P., Doorman, M., Boon, P., Reed, H., & Gravemeijer, K. (2010). The teacher and the tool: Instrumental orchestrations in the technology-rich mathematics classroom. *Educational Studies in Mathematics, 75*, 213-234. https://doi.org/10.1007/s10649-010-9254-5

Engelbrecht, J., & Borba, M. C. (2024). Recent developments in using digital technology in mathematics education. *ZDM Mathematics Education, 56*, 281-292. https://doi.org/10.1007/s11858-023-01530-2

Faqih, M. I., Adriyani, Z., & Listiyani, L. (2022). Advanced chatbot development to improve student literacy and numeracy skills. *Phenomenon: Jurnal Pendidikan MIPA, 12*(2), 206-215. https://doi.org/10.21580/phen.2022.12.2.14186

Gattupalli, S., Lee, W., Allessio, D., Crabtree, D., Arroyo, I., & Woolf, B. (2023). Exploring pre-service teachers' perceptions of large language models-generated hints in online mathematics learning. In

S. Moore, J. Stamper, R. Tong, C. Cao, Z. Liu, X. Hu, Y. Lu, J. Liang, H. Khosravi, P. Denny, A. Singh, & C. Brooks (Eds.), *Proceedings of the Workshop on Empowering Education with LLMs: The Next-Gen Interface and Content Generation 2023* (vol. 3487, pp. 151-162). CEUR-WS.

Guan, L., Zhang, Y., & Gu, M. M. (2025). Pre-service teachers preparedness for AI-integrated education: An investigation from perceptions, capabilities, and teachers' identity changes. *Computers and Education: Artificial Intelligence, 8*, Article 100341. https://doi.org/10.1016/j.caeai.2024.100341

Gueudet, G., Pepin, B., Courtney, S., Kock, Z.-J., Misfeldt, M., & Tamborg, A. L. (2021). Digital platforms for mathematics teacher curriculum design: Affordances and constraints. In G. Gueudet, B. Pepin, S. Courtney, Z.-J. Kock, M. Misfeldt, & A. L. Tamborg (Eds.), *Mathematics education in the digital age* (pp. 84-98). Routledge. https://doi.org/10.4324/9781003137580-6

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

Kabiljagić, M., Wachtler, J., Ebner, M., & Ebner, M. (2022). Math trainer as a chatbot via system (push) messages for android. *International Journal of Interactive Mobile Technologies, 16*(17), 75-87. https://doi.org/10.3991/ijim.v16i17.33351

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

Laksana, F. S. W., & Fiangga, S. (2022). The development of web-based chatbot as a mathematics learning media on system of linear equations in three variables. *MATHEdunesa, 11*(1), 145-154. https://doi.org/10.26740/mathedunesa.v11n1.p145-154

Lee, D., & Yeo, S. (2022). Developing an AI-based chatbot for practicing responsive teaching in mathematics. *Computers & Education, 191*, Article 104646. https://doi.org/10.1016/j.compedu.2022.104646

Lee, D., Son, T., & Yeo, S. (2025). Impacts of interacting with an AI chatbot on preservice teachers' responsive teaching skills in math education. *Journal of Computer Assisted Learning, 41*(1), Article e13091. https://doi.org/10.1111/jcal.13091

Liu, C. C., Wang, H. J., Wang, D., Tu, Y. F., Hwang, G. J., & Wang, Y. (2024). An interactive technological solution to foster preservice teachers' theoretical knowledge and instructional design skills: A chatbot-based 5E learning approach. *Interactive Learning Environments, 32*(10), 6698-6716. https://doi.org/10.1080/10494820.2023.2277761

Matsuda, N., Yarzebinski, E., Keiser, V., Raizada, R., Cohen, W. W., Stylianides, G. J., & Koedinger, K. R. (2013). Cognitive anatomy of tutor learning: Lessons learned with SimStudent. *Journal of Educational Psychology, 105*(4), 1152-1163. https://doi.org/10.1037/a0031955

Melo, J. N. B., de Lima, J. V., & do Canto Filho, A. B. (2020). A tarefa de casa na disciplina de matemática mediada por assistente virtual de comunicação-chatbot [Homework in the mathematics discipline mediated by a virtual communication assistant-chatbot]. *Revista Eletrônica de Educação Matemática, 15*(1), 1-20. https://doi.org/10.5007/1981-1322.2020.e58129

Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative research: A guide to design and implementation* (4th ed.). John Wiley & Sons.

Nguyen, D. T., & Pham, Q. V. (2025). The evolving landscape of AI integration in mathematics education: A systematic review of trends (2015-2025). *Eurasia Journal of Mathematics, Science and Technology Education, 21*(10), Article em2714. https://doi.org/10.29333/ejmste/17078

Nguyen, H. D., Pham, V. T., Tran, D. A., & Le, T. T. (2019). Intelligent tutoring chatbot for solving mathematical problems in high school. In J. Mothe, L. H. Son, & T. Q. V. Nguyen (Eds.), *Proceedings of the 11th International Conference on Knowledge and Systems Engineering* (pp. 1-6). IEEE. https://doi.org/10.1109/KSE.2019.8919396

Oh, S. J. (2023). Effective ChatGPT prompts in mathematical problem solving: Focusing on quadratic equations and quadratic functions. *Communications of Mathematical Education, 37*(3), 545-567. https://doi.org/10.7468/jksmee.2023.37.3.545

Papadakis, S., Vaiopoulou, J., Sifaki, E., Stamovlasis, D., & Kalogiannakis, M. (2021). Attitudes towards the use of educational robotics: Exploring pre-service and in-service early childhood teacher profiles. *Education Sciences, 11*(5), Article 204. https://doi.org/10.3390/educsci11050204

Peura, P., Aro, T., Räikkönen, E., Viholainen, H., Koponen, T., Usher, E. L., & Aro, M. (2021). Trajectories of change in reading self-efficacy: A longitudinal analysis of self-efficacy and its sources. *Contemporary Educational Psychology, 64*, Article 101947. https://doi.org/10.1016/j.cedpsych.2021.101947

Rabardel, P. (1995). *Les hommes et les technologies: Approche cognitive des instruments contemporains* [Humans and technologies: A cognitive approach to contemporary instruments]. Armand Colin.

Reshmi, S., & Balakrishnan, K. (2016). Implementation of an inquisitive chatbot for database supported knowledge bases. *Sādhanā, 41*(10), 1173-1178. https://doi.org/10.1007/s12046-016-0544-1

Ruiz-López, N. (2018). The instrumental genesis process in future primary teachers using dynamic geometry software. *International Journal of Mathematical Education in Science and Technology, 49*(4), 481-500. https://doi.org/10.1080/0020739X.2017.1377302

Sahoo, P., Singh, A. K., Saha, S., Jain, V., Mondal, S., & Chadha, A. (2024). *A systematic survey of prompt engineering in large language models: Techniques and applications*. arXiv. https://doi.org/10.48550/arXiv.2402.07927

Sanusi, I. T., Ayanwale, M. A., & Tolorunleke, A. E. (2024). Investigating pre-service teachers' artificial intelligence perception from the perspective of planned behavior theory. *Computers and Education: Artificial Intelligence, 6*, Article 100202. https://doi.org/10.1016/j.caeai.2024.100202

Schorcht, S., Buchholtz, N., & Baumanns, L. (2024). Prompt the problem: Investigating the mathematics educational quality of AI-supported problem solving by comparing prompt techniques. *Frontiers in Education, 9*. https://doi.org/10.3389/feduc.2024.1386075

Sinclair, N., Haspekian, M., Robutti, O., & Clark-Wilson, A. (2022). Revisiting theories that frame research on teaching mathematics with digital technology. In A. Clark-Wilson, O. Robutti, & N. Sinclair (Eds.), *The mathematics teacher in the digital era: International research on professional learning and practice* (pp. 391-418). Springer. https://doi.org/10.1007/978-3-031-05254-5_15

So, K., Lee, S. Y., & Choi, Y. (2024). Teachers' sense-making of the decentralizing curriculum reform policy: A comparative case study in South Korea. *International Journal of Educational Research, 125*, Article 102331. https://doi.org/10.1016/j.ijer.2024.102331

Son, T., Yeo, S., & Lee, D. (2024). Exploring elementary preservice teachers' responsive teaching in mathematics through an artificial intelligence-based chatbot. *Teaching and Teacher Education, 146*, Article 104640. https://doi.org/10.1016/j.tate.2024.104640

Song, Y., Kim, J., Xing, W., Liu, Z., Li, C., & Oh, H. (2025). Elementary school students' and teachers' perceptions toward creative mathematical writing with generative AI. *Journal of Research on Technology in Education*. https://doi.org/10.1080/15391523.2025.2455057

Stefanova, T., & Georgiev, S. (2024). Possibilities for using AI in mathematics education. *Mathematics and Education in Mathematics, 53*, 117-125. https://doi.org/10.55630/mem.2024.53.117-125

Talha Junaid, M., Barakat, S., Awad, R., & Anwar, N. (2024). Adopting the power of AI chatbots for enriching students learning in civil engineering education: A study on capabilities and limitations. In A. Al-Marzouqi, S. A. Salloum, M. Al-Saidat, A. Aburayya, & B. Gupta (Eds.), *Artificial intelligence in education: The power and dangers of ChatGPT in the classroom* (pp. 25-47). Springer. https://doi.org/10.1007/978-3-031-52280-2_3

Tamborg, A. L. (2021). Improving mathematics teaching via digital platforms? Implementation processes seen through the lens of instrumental genesis. *ZDM Mathematics Education, 53*, 1059-1071. https://doi.org/10.1007/s11858-021-01282-x

Trouche, L. (2004). Managing the complexity of human/machine interactions in computerized learning environments: Guiding students' command process through instrumental orchestrations. *International Journal of Computers for Mathematical Learning, 9*(3), 281-307. https://doi.org/10.1007/s10758-004-3468-5

Trouche, L. (2005). An instrumental approach to mathematics learning in symbolic calculator environments. In D. Guin, K. Ruthven, & L. Trouche (Eds.), *The didactical challenge of symbolic calculators: Turning a computational device into a mathematical instrument* (pp. 137-162). Springer. https://doi.org/10.1007/0-387-23435-7_7

Van de Walle, J. A., Karp, K. S., & Bay-Williams, J. M. (2016). *Elementary and middle school mathematics: Teaching developmentally*. Pearson.

Van Vaerenbergh, S., & Pérez-Suay, A. (2022). A classification of artificial intelligence systems for mathematics education. In P. R. Richard, M. P. Vélez, & S. Van Vaerenbergh (Eds.), *Mathematics education in the age of artificial intelligence: How artificial intelligence can serve mathematical human learning* (pp. 89-106). Springer. https://doi.org/10.1007/978-3-030-86909-0_5

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

Winkler, R., Hobert, S., Salovaara, A., Söllner, M., & Leimeister, J. M. (2020). Sara, the lecturer: Improving learning in online education with a scaffolding-based conversational agent. In R. Bernhaupt, F. Mueller, D. Verweij, & J. Andres (Eds.), *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-14). ACM. https://doi.org/10.1145/3313831.3376781

Wrtn Technologies. (2024). Wrtn. https://wrtn.ai/

Yang, T. C., & Chen, J. H. (2023). Pre-service teachers' perceptions and intentions regarding the use of chatbots through statistical and lag sequential analysis. *Computers and Education: Artificial Intelligence, 4*, Article 100119. https://doi.org/10.1016/j.caeai.2022.100119