frontiers | Frontiers in Education

TYPE Original Research
PUBLISHED 09 September 2025
DOI 10.3389/feduc.2025.1605921

OPEN ACCESS

EDITED BY
Jasmin Cowin,
Touro University, United States

REVIEWED BY
Widodo Winarso,
Universitas Islam Negeri Siber Syekh Nurjati Cirebon, Indonesia
Annalisa Cusi,
Sapienza University of Rome, Italy

*CORRESPONDENCE
Gabriele Kaiser
gabriele.kaiser@uni-hamburg.de

RECEIVED 07 May 2025
ACCEPTED 25 August 2025
PUBLISHED 09 September 2025

CITATION
Bastian A, Buchholtz N and Kaiser G (2025) Using AI chatbots to facilitate mathematics preservice teachers' noticing skills. *Front. Educ.* 10:1605921.
doi: 10.3389/feduc.2025.1605921

Using AI chatbots to facilitate mathematics preservice teachers' noticing skills

Anton Bastian[1], Nils Buchholtz[1] and Gabriele Kaiser[1,2]*

[1]Department of Mathematics Education, Faculty of Education, University of Hamburg, Hamburg, Germany, [2]Faculty of Education and Arts, Nord University, Bodø, Norway

The ability to notice and interpret relevant aspects of classroom interactions is central to effective teaching and plays a critical role in teachers' professional development. This study investigated the development and testing of an AI chatbot (NiCo) designed to support preservice teachers' noticing skills. The chatbot was built using the GPT-4o model and included a meta-prompt that focused on structured support and providing feedback based. The chatbot was tested with 25 mathematics preservice teachers analyzing classroom videos in the field of secondary algebra education. Data collection involved questionnaires on prior experience with the content, self-assessed AI competencies, and user perception of the chatbot. Results indicate that preservice teachers experienced the chatbot as motivating and helpful for developing their noticing skills. They reported support in perceiving classroom events, interpreting them, and generating alternative actions. Usability was rated positively, though improvements were suggested regarding input options, accessibility, and subject-specific guidance. The study demonstrates the potential of AI chatbots to provide individualized support for the development of teacher noticing skills in teacher education. With the empirically validated meta-prompt structure of NiCo, we propose a novel, accessible and scalable format for AI-enhanced teacher education and even professional development. By demonstrating both feasibility and positive user perceptions, this research contributes concretely to the academic discourse on AI-driven interventions for preservice teachers' noticing skill development.

KEYWORDS

teacher noticing, initial teacher education, artificial intelligence, ChatGPT, teachers' professional competence, preservice teachers, mathematics secondary education

1 Introduction

To transform teachers' knowledge into successful performance in the classroom and to teach in an inclusive and student-centered way, teachers require teacher noticing skills that enable them to identify important events in the classroom and process them further (Dindyal et al., 2021; Metsäpelto et al., 2021). Therefore, teacher noticing has been a growing research field in education and, particularly, mathematics education research (König et al., 2022; Wei et al., 2023). Besides the structure of the construct and its development, one main research focus has been the facilitation of teacher noticing skills, especially for preservice teachers (PSTs) (Amador et al., 2021; König et al., 2022). Researchers have developed several intervention formats, such as video clubs, professional development, and university courses, to promote (preservice) teachers' skills to notice classroom situations and make appropriate pedagogical decisions (Fernández et al., 2020; Kleinknecht and Gröschner, 2016; van Es and Sherin, 2002). Empirical investigations of these interventions reveal ambiguous effects on teacher noticing dependent on the specific program design, although the discussion of classroom videos, the linkage of theories to specific situations, lesson-follow up, and feedback appear to be promising approaches to facilitate teacher noticing skills (Amador et al., 2021; Bastian et al., 2024; Santagata et al., 2021; Stürmer et al., 2013; Weber et al., 2018). However, the existing interventions are often time-consuming and demanding for both participants and facilitators. As a result, only a limited number of teachers are able to benefit from these interventions, and teachers find it difficult to fit them into their already busy work schedules (Hill, 2009; OECD, 2019). On the other hand, there is a significant need for professional development in teachers' noticing skills, as indicated by low scores in empirical evaluations and teachers' reported difficulties (Bastian et al., 2022, 2025; Jacobs et al., 2010; Jong et al., 2021). This requires an accessible, low-threshold, and scalable professional development format to support teacher noticing skills.

At the same time, artificial intelligence (AI) and the latest breakthroughs in generative AI (GenAI) models have impacted and disrupted education and, of course, also teacher education worldwide, offering new opportunities and approaches (Mosher et al., 2024; Saúde et al., 2024; Pepin et al., 2025). Prompted AI chatbots provide teachers with easy access to a dialogue and discussion partner, with the potential to facilitate reflection and to provide feedback (Schorcht et al., 2024, Buchholtz et al., 2024; Tarantini, 2023). Yet, this potential has rarely been utilized to support teachers' noticing skills. To address this research gap, in this study, we present the development and initial results of the testing of a prompted AI chatbot (NiCo) based on the GPT-4o model that was used with PSTs after viewing two classroom videos. In particular, we examine (1) the development and structure of the chatbot, (2) the prior experiences and skills that PSTs brought to working with the AI chatbot, and, most importantly, (3) their feedback on the AI chatbot and its use.

In the broader context of teacher education, GenAI-powered tools like NiCo have the potential to fundamentally transform professional development by delivering scalable, personalized support. By offering automated, individualized feedback and sustained reflective dialogue, such chatbots enable larger groups of (preservice) teachers to hone their noticing skills anytime and anywhere. Additionally, this allows professional learning to be tailored dynamically to each teacher's needs. Therefore, integrating AI chatbots into teacher education may not only advance research on teacher noticing but also opens new pathways for sustainable, technology-enhanced professional growth.

2 Literature review, theoretical framework, and research questions

2.1 Teacher noticing and its facilitation

Originally introduced more than two decades ago to promote student-centered teaching, the construct of teacher noticing can be understood as a professionalized way of perceiving and processing classroom situations that is characteristic of the teaching profession, meaning to navigate the "maelstrom" of sensory data in classroom situations (Sherin et al., 2011; van Es and Sherin, 2002). In recent years, teacher noticing–its structure, characteristics, and development–has received particular attention in educational research, and especially in mathematics education research, as evidenced by a wealth of empirical studies and theoretical considerations (Dindyal et al., 2021; König et al., 2022; Wei et al., 2023). Teacher noticing has been incorporated as an integral component in recent conceptualizations of teacher competence as well as expertise (Kaiser et al., 2017; Lachner et al., 2016; Metsäpelto et al., 2021). There, it acts as a mediator between teachers' dispositions, such as knowledge and beliefs, and their actual performance in the classroom (Blömeke et al., 2015). Further, empirical research has demonstrated a positive influence on student performance and achievement as well as on instructional quality (Blömeke et al., 2022; Gheyssens et al., 2021; Mischo et al., 2023). Thus, its development constitutes an important aim of teacher education and professional development (Amador et al., 2021; Bastian et al., 2022).

2.1.1 Discourse on teacher noticing

Despite, or perhaps because of, the growing body of research, the conceptualization of the construct has been heterogeneous and shaped by four theoretical perspectives (König et al., 2022; Weyers et al., 2023). In the cognitive-psychological perspective, which can be considered the dominant one in the current discussion, teacher noticing is understood as a set of cognitive processes that occur during noticing in the classroom (König et al., 2022). Influential approaches from this perspective are the *Learning to notice* framework by van Es (2011) and the *Professional noticing of children's mathematical thinking* framework by Jacobs et al. (2010). Another perspective, the expertise paradigm, emphasizes the differences between novice and expert teachers in noticing classroom events and originates from the research on teaching expertise (Berliner, 2001). A sociocultural perspective, based on the approach of professional vision by Goodwin (1994), focuses on socially constructed ways of perceiving and understanding that are characteristic of teachers as a community of practice. Mason (2002) established a discipline-specific perspective on teacher noticing, highlighting the role of the individual teachers and their development in achieving sensitized and methodical awareness in the classroom. In the current discourse, conceptualizations are often influenced by several of the four perspectives (Weyers et al., 2023).

Following the cognitive-psychological perspective, most studies define teacher noticing, sometimes also referred to as professional vision (Sherin et al., 2025), as comprising the processes of perceiving or selectively attending to classroom events and interpreting or making sense of the perceived events (Sherin et al., 2011; Weyers et al., 2023). A growing number of studies also include the subsequent decision-making in their understanding of noticing (Jacobs et al., 2010; Kaiser et al., 2015), while only few studies focus on the enactment of the decisions (Thomas et al., 2020).

2.1.2 Our own theoretical framework on teacher noticing

In this study, we understand teacher noticing as an integral component of teacher competence that contributes to the transformation of teachers' dispositions, including the knowledge they acquired in teacher education, into their classroom performance (Bastian et al., 2022; Kaiser et al., 2015). Following the cognitive-psychological perspective and an analytical understanding of teacher noticing, we conceptualize teacher noticing as consisting of three facets: (1) the perception of important events in instructional settings, (2) the interpretation of the perceived events based on one's dispositions, such as professional knowledge, and (3) the subsequent decision-making on how to act, e.g., to respond to student action or to develop instructional strategies on how to proceed in the perceived moment (Kaiser et al., 2015). This conceptualization includes teaching situations relevant for qualitative teaching of all kinds, combining a mathematics pedagogical and a general pedagogical perspective and incorporates decision-making as an integral part of teacher noticing (Bastian et al., 2022).

2.1.3 Professional development of teacher noticing skills

To investigate factors that promote teacher noticing skills, several studies have examined the change in teacher noticing over the course of parts of (initial) teacher education, particularly for included practical experiences and phases of teaching internships, and factors that may have caused the change. Overall, teacher experience and especially reflection on it have been shown to have a major impact on the development of teacher noticing (Stürmer et al., 2013; Weber et al., 2018). Bastian et al. (2024) demonstrated the positive impact of activities that focused on linking theories to specific situations and reflection on the development of teacher noticing for a half year long teaching internship.

While the previous studies investigated parts of teacher education that are not solely focused on developing teacher noticing skills, other studies created and examined professional development programs specifically addressing the improvement of teacher noticing skills. In their seminal study, van Es and Sherin (2002) presented the use of video analysis and how several teachers benefited from its use in their noticing over the course of 3 months. As a scaffold for noticing classroom situations, prompts and frameworks, often derived directly from research on teacher noticing, were applied in several studies to facilitate the development of teacher noticing skills (van Es et al., 2017; van Es and Sherin, 2002; Weber et al., 2018). Kleinknecht and Gröschner (2016) emphasized the positive influences of expert and peer feedback on PSTs' noticing of their own teaching in a one-semester university course, while Fernández et al. (2020) showed how narrative description of instructional events, reasoning about them, and receiving feedback improved preservice teacher noticing.

However, existing interventions are often time-intensive, challenging, and require intensive facilitator support, making them suitable for only a small number of participants. This poses a challenge in scaling them to larger groups of teachers, and in getting teachers to participate in related professional development despite an urgent need.

2.2 Generative artificial intelligence in teacher education

While the use of AI in teaching with a focus on students has been a field of research for some time (Labadze et al., 2023; Zawacki-Richter et al., 2019), research targeting at the use of AI in teacher education and professional development is rather scarce. However, the use of GenAI in the daily life of teachers offers great potential to support and improve teaching, but also new risks and responsibilities for teachers (Saúde et al., 2024; Pepin et al., 2025). GenAI may be able to support teachers in diagnosing student needs and assessing student performance as well as enabling teachers to individualize their teaching and meet individual students' conditions in order to teach inclusively (Buchholtz et al., 2024; Mosher et al., 2024). Moreover, GenAI may be particularly useful for teacher education, as it could help PSTs learn to plan, analyze and follow up on instruction as an informed dialog partner (Al-Shammari and Al-Enezi, 2024; Buchholtz et al., 2024; Neumann et al., 2024). As this area of research is relatively new, there are only few empirical studies beyond theoretical discussions of the uses and challenges of GenAI in teacher education.

Some studies examine, teachers attitudes and beliefs toward the use of AI. Thararattanasuwan and Prachagool (2024) demonstrated a high willingness among PSTs to use AI technology in instructional settings. This was corroborated by Zhang et al. (2023), who also illustrated the influences of perceived usefulness, ease of use, and also perceived enjoyment on the intention to use AI in the classroom. Further, research suggests that feedback from an AI chatbot is understandable and useful to teachers and can influence their intended actions (Kasepalu et al., 2022). Other research focuses on the testing of AI chatbots and their usefulness in supporting (preservice) teachers. Jacobsen and Weber (2023) compared AI-driven feedback to expert and novice feedback for learning goal generation and found that AI-driven feedback was superior in terms of specificity and explanation, but only when AI prompts were of high quality. To promote lesson planning and task adaptation skills, Buchholtz and Huget (2025) explored the use of ChatGPT and certain prompts in a setting, where PSTs chatted with an AI chatbot. They illustrated the potential of the AI preservice teacher chats in activating and developing teacher competence, but also emphasized the limits of GenAI in providing advice in line with modern, innovative teaching and in fine-tuning lesson plans and task adaptions with detailed subject-specific knowledge.

In the domain of teacher noticing, studies that address or incorporate GenAI and AI in general are limited. Exceptions include the studies by Zhang et al. (2024) and Lee et al. (2025), both of which used AI chatbots as virtual students to simulate teacher-student interaction, thereby facilitating PSTs' noticing and responsive teaching skills. They illustrated possibilities of AI-supported simulation to provide PSTs with training opportunities but also highlighted some challenges, especially in terms of authenticity, the lack of personal connections with students, and limited knowledge of PSTs, which hindered the interaction.

Several recent studies have explored the potential of AI chatbots to foster noticing more directly. For instance, Galiç et al. (2025) developed a teachable-agent-based chatbot that provided real-time feedback on mathematics teachers questions in mathematics simulations. The chatbot's responses prompted teachers to refine their questioning techniques and engage more deeply with student ideas. Similarly, Lee and Yeo (2022) and Son et al. (2024) implemented chatbots with common student misconceptions in risk-free environments to allow preservice teachers to practice interpreting and responding to student thinking. These studies suggest that chatbot-supported interactions can encourage more strategic questioning and foster adaptive teaching behavior.

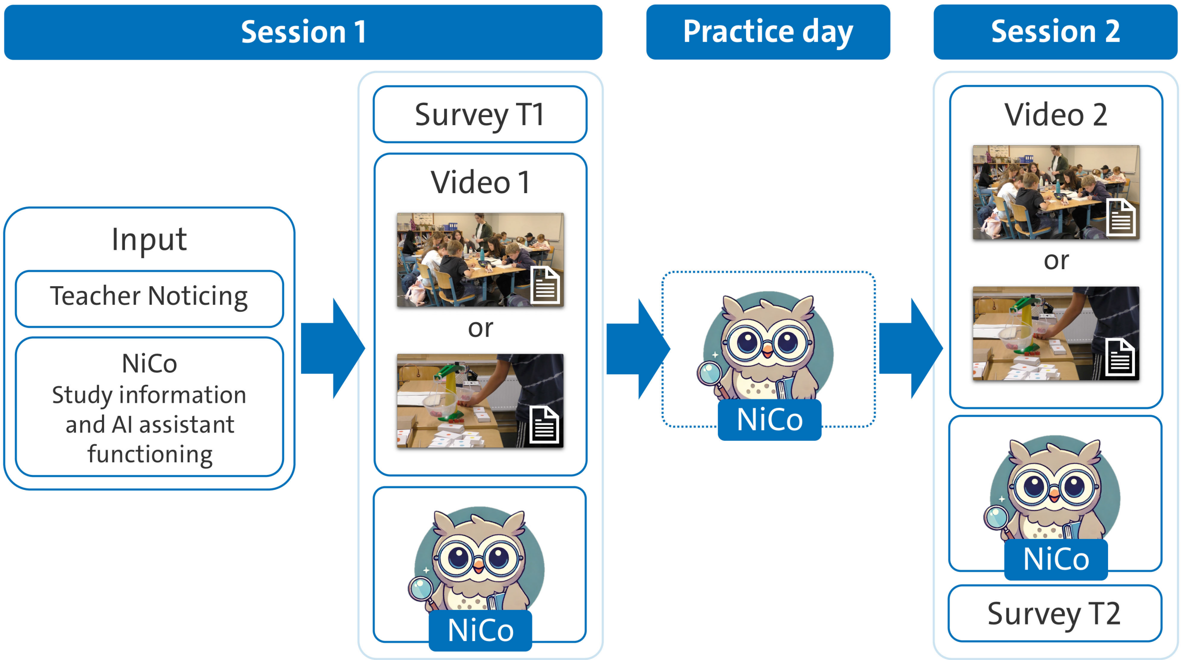

FIGURE 1
Study design.

Notably, Lee et al. (2025) found that participants interacting with a chatbot demonstrated a significant improvement in their ability to attend to, interpret, and respond to classroom situations compared to a control group.

While these approaches show how AI chatbots can be used to simulate classroom dynamics, they primarily focus on enabling responsive teaching in simulated dialogue. Often, so called Persona prompting is used, which focuses on the dialogic framing of the chatbot. Here, the chatbot is assigned a specific role, for example, as a supportive mathematics teacher or reflective mentor (Zheng et al., 2024; Olea et al., 2024). However, recent studies in general show inconclusive results on the effects of Persona, when it comes to increased GenAI performance (Zheng et al., 2024; Olea et al., 2024). The potential of AI chatbots to actively scaffold PSTs' noticing processes during the analysis of real classroom situations–such as video-based noticing–remains largely unexplored. Further, although attitudes and beliefs of PSTs toward AI and its use have been investigated, the examination of the usability of the chatbots and respective feedback from the PSTs as their users has yet to be considered as a core process in a user-centered design that facilitates the chat bots use of the end-user (International Organization for Standardization, 2019).

2.3 Research questions

Based on the identified need for professional development in teacher noticing, the research gaps identified in the literature review, and our theoretical framework, this study aims to develop and test an AI chatbot that leverages reflection, feedback, and narration to support mathematics PSTs in their noticing of videos of classroom interactions. Further, we aim to provide insight into PSTs perception of the chatbot and its usability. Specifically, we address the following research questions:

(1) How can an AI chatbot be designed and prompted to scaffold and facilitate PSTs' noticing skills–(a) perception of salient classroom events, (b) interpretation of those events in light of professional knowledge, and (c) decision-making about next instructional moves?
(2) Which prior experience and skills (e.g., opportunities to learn, content knowledge, AI-experience) shape PSTs' engagement with and responses to the noticing facets when interacting with the chatbot?
(3) How do PSTs evaluate the chatbot's support for (a) perceiving, (b) interpreting, and (c) deciding upon classroom events, and what usability factors influence its effectiveness?

3 Methodological approach

3.1 Study design

In the present study, we constructed and explored the use of an AI chatbot, called Noticing Companion (NiCo), to facilitate teacher noticing skills of mathematics PSTs in a course that accompanies a teaching internship during the master's program as a part of initial teacher education. In this teaching internship, the PSTs spend one practice day a week in in a school for 3 months during the semester, followed by a block period of 5–6 weeks. The PSTs attend an accompanying seminar at the university that provides them with mathematics pedagogical knowledge and skills for their internship, as well as space to reflect on their experiences and relate theory to practice. In two of these seminars, we used NiCo for two sessions in consecutive weeks conducted by the first author (see Figure 1).

In the first session, we provided the PSTs with a short input on teacher noticing, i.e., the construct's structure and ways to develop it, as well as information about the study and instructions for the use of NiCo. Participants then completed an initial survey to collect demographic information, such as gender, teaching experience, and teaching type, previous opportunities to learn (OTL) for the content covered in the videos and previous use of digital media and AI in instruction, and their beliefs about their own AI competencies. Afterward, they received contextual information about the video vignettes depicting some excerpts of a secondary mathematics lesson that they were going to watch, such as the grade level and the lesson's objective. We chose to use videos as stimulus material for PSTs' chats with NiCo instead of their own classroom experiences to control for the situation and to enable us to analyze the chat in a controlled way.[1] PSTs then watched the video once with the whole seminar group. Subsequently, the PSTs chatted with NiCo about the seen video vignette for 20–30 min, while they were allowed to revisit the video vignette and access the video's transcript as well as the contextual information. These chat phases align with the study's goal to facilitate PSTs noticing skills: first to practice "perceiving" salient classroom events, then to "interpret" those events through professional knowledge with the help of NiCo, and finally to "decide" on instructional responses. The PSTs were aware of these aims and they followed these guidelines for chatting with NiCo:

- Chat with NiCo about any classroom situations that have caught your eye or that you notice when you look at the video again.
- NiCo will guide you through the conversation. When you have finished discussing one situation, start a new one.
- You can refer back to the contextual information and watch the video again and again throughout the work phase.

During the week, the PSTs also had the opportunity to use NiCo to chat about a classroom situation they experienced during their practice day at the school. In the second session, we had another circle of video and chat using NiCo.[2] After two circles of using NiCo, the PSTs then provided feedback on the chatbot, its usability, and their work with it in a second survey.

3.2 Sample

The study comprises 25 pre-service teachers from a university in northern Germany that participated in the seminars, 81% of whom identified as female, with an average age of 25 years (min = 22, max = 35). The sample consists of almost equal number of PSTs studying to teach at academic-track schools (48%) and at non-academic-track schools (52%). They had an average of 7.9 months of teaching experience (min = 1, max = 24). 21 PSTs participated in Session 1 of the study, 19 in Session 2, with 15 PSTs participating in both sessions.

3.3 Video vignettes

As stimulus material, we used two video vignettes from the Teacher Education and Development Study – Inclusive Mathematics Education (TEDS-IME), because they provide rich opportunities for practicing the three noticing processes aligned with our learning objectives: perceiving classroom events, interpreting, and deciding on targeted pedagogical actions concerning students' individual understanding and condition as well as the whole class. Both videos focus on secondary algebra education. In the 3-min video "Taxi ride", the students in the video work on an exercise where they are asked to describe the cost of a taxi ride using an algebraic expression with two variables. The video shows the teacher's introduction to the task and two dialogues between students at different levels of learning prerequisites. This vignette was chosen to train PSTs' ability to perceive subtle differences in student reasoning during dialogues and to interpret how learners connect contextual problems to formal algebraic representations. The five-minute long video "Scale model – exploration" shows a whole-class discussion at the end of a lesson on the scale model to introduce the principle of equations. In a previous work phase, students were asked to build equations with stones and boxes and to also express these equations as sketches and formal algebraic expressions. In the whole class discussion, four students present their own solutions and comment on each other's solutions, giving insights into their mathematical thinking. By presenting multiple student solutions and peer commentary, this vignette supports the learning objective of noticing varied mathematical understandings and selecting appropriate instructional strategies.

3.4 Assessment instruments

In the two surveys in our study, we used three instruments to assess OTL, participants AI competence, and their user feedback on NiCo, besides items focusing on participants demographics. Item count and example items for each instrument are presented in Table 1. To capture PSTs' previous OTL on algebra education, teacher noticing, and usage of digital media and AI in the classroom, we applied ten five-point Likert items based on the work by Weyers et al. (2023) that enabled the participants to indicate whether they covered the specific topic in their prior studies and at what intensity. For the self-assessment of AI competence, we adopted a subscale of a larger instrument by Buchholtz et al. (in preparation) on teacher beliefs toward AI in teaching and learning mathematics. To comprehensively collect pre-service teachers' feedback on their work with NiCo as the focus of this study, we combined newly developed six-point Likert scale items focusing on motivation, utility, future use, and usability with open-response items on learning outcomes, benefits, and barriers to working with NiCo, suggested changes, and future expectations.

1 The analysis of AI-PST-chats is not part of this study.

2 Note that the two seminars received the videos in reverse order to control for serial effects.

TABLE 1 Overview of instruments with example items.

Instrument	Example item(s)	Format information	*n* items
Opportunities to learn	"Have you covered the following topics in your studies (in seminars/lectures)?" - Basic mental models of the understanding of variables - Teacher noticing - Use of AI application in the classroom	Five-point Likert items from "No" (0) to "Yes, as the focus of an entire seminar or lecture" (4).	10
Self-assessed AI competence	"I feel able to meet the technical challenges of using AI technologies in mathematics education." "I am open to using generative AI tools in my teaching."	Five-point Likert items from "Strongly disagree" (1) to "Strongly agree" (5)	16
Feedback on NiCo (part I)	"I would use NiCo in my daily teaching to develop my skills." "Working with NiCo is technically challenging."	Six-point Likert items from "Strongly disagree" (1) to "Strongly agree" (6)	15
Feedback on NiCo (part II)	"What has worked well for you when using NiCo?" "To what extent did working with NiCo support or hinder practicing teacher noticing?"	Open-response items	7

Interpretations and asks the PST to focus on perception first.

3.5 AI chatbot construction and prompting

We designed NiCo as an AI assistant based on the large language model GPT-4o by OpenAI (2024b), which we provided with a specific role description (persona prompting; Zheng et al., 2024) and instructions on how to act in chats with PSTs, which governed all chats and ranked above PSTs' prompts. For this meta-prompt, which influences the AI's behavior throughout the whole chat, we followed Park and Choo (2024) recommendations for effective prompt engineering, which include stating the AI's role, chat goal, recipient, theme, and chat structure in the prompt. Drawing on our cognitive-psychological framework, the meta-prompt comprised at its core guiding PSTs with attention-directing questions to elicit perception of salient classroom events, supporting them in the interpretation in light of professional knowledge, and help them to identify cues to provoke decision-making about instructional actions. For the phrasing, we further aligned with the CLEAR framework, which provides instructions for generating concise, logical prompts (Lo, 2023; Park and Choo, 2024). We included few-shot examples of questions to elicit teacher noticing and of feedback in the prompt to better convey the intended meaning to the GenAI (Park and Choo, 2024). To better facilitate a dialogue between the AI and the PSTs, we furthermore included chain-of-thought and query-loop prompting methods. This ensured that the AI would go through the dialogue step-by-step and would wait for input from the PSTs (Schorcht et al., 2023). We developed the prompt iteratively, with multiple feedback loops, piloting, and chats with the GPT builder by OpenAI (2024a). During iterative piloting, we validated each prompt block against its corresponding noticing facet, refining the meta-prompt to ensure clear mapping between our theoretical constructs and NiCo's behavior. The final meta-prompt is described in the results section 4.1 as a central achievement of this study and is presented with its respective prompt blocks in the electronic Supplementary material (ESM).

To comply with the General Data Protection Regulation valid at this university, the chats between NiCo and the PSTs were conducted through the platform fobizz[3], which enables the creation and use of GPT-4o chatbots while maintaining data privacy for the users. The platform also allows the direct export of all chats for later analysis.

3.6 Data analysis

In this study, we focus on the PSTs' responses to items related to their OTL, self-assessed AI competence, and, in particular, their feedback to NiCo. For the OTL items, we calculated the classical indicators for each item. For the AI competence items, one item had to be excluded due to poor item-total correlation. The remaining 15 items showed good reliability (Cronbach's α = 0.90) and were thus transformed into a single sum score. The feedback rating scale items were grouped into the four themes of motivation, utility, future use, and usability, and then analyzed individually based on their indicators. The open-response items for feedback on NiCo were analyzed with MAXQDA (VERBI Software, 2024) using the structuring qualitative text analysis following Kuckartz (2014). In this analysis, we first defined four major categories to structure PSTs' feedback: learning outcomes, advantages and support, disadvantages and difficulties, and suggestions for improvement. We then inductively created subcategories for the four main categories based on the material and comprehensively defined all categories in a coding manual (see the Supplementary material for the coding manual). The entire category system is presented in the Results section.

The first author and another trained coder individually coded 35% (6 out of 17) of the data. The interrater reliability in this double coding was substantial (κ = 0.79)[4], indicating a precise, objective coding manual. Due to the manageable number of responses and as an additional precaution to ensure coding with scientific rigor, the first author and the trained coder then double-coded all

3 https://fobizz.com

4 For a conservative calculation, we estimated interrater reliability not based on occurrences in a response but on code overlap on segments of at least 90% (VERBI Software, 2024) and used the recommendations of Brennan and Prediger (1981) to calculate the kappa value.

10.3389/feduc.2025.1605921

remaining responses. All disagreements were thoroughly discussed and resolved by consensus.

4 Results

We present the findings for each research question in a separate section. We first take a look at the design of NiCo (Research Question 1), then consider PSTs' prior experiences and skills (Research Question 2), before focusing on PSTs' feedback (Research Question 3).

4.1 The AI chatbot NiCo

To build our chatbot and determine NiCo's behavior during the chats with the PSTs, as well as the structure and content of the chat, we created a comprehensive role description and instructions for the chat, which were given to the AI as a meta-prompt based on recommendations for efficient prompt engineering (Park and Choo, 2024; Zheng et al., 2024). For transparency and comprehensibility, the entire meta-prompt is presented and annotated in the Supplementary material. It defines the role of the AI, the recipients, goals, topics, and structure of the chats, and provides examples to better describe the expected output. In addition, the meta-prompt specifies two main functions of the chats: eliciting PSTs' teacher noticing skills and providing feedback to PST's responses. We organized the meta-prompt to mirror our noticing framework by including distinct noticing examples and questions: (a) perception questions that direct GPT-4o to support PSTs in highlighting salient classroom events, (b) interpretation questions that scaffold sense-making based on professional knowledge, and (c) decision-making questions that elicit instructional responses (see the Supplementary material for the specific description of the meta-prompt). Specifically, the meta-prompt defines NiCo's role as a supportive facilitator for PSTs with the aim of helping them to develop their teacher noticing skills by engaging PSTs in chats about classroom situations. NiCo is instructed to guide the discussion by asking structured questions about the PSTs' perceptions, interpretations, and decisions regarding the algebra teaching situations observed and to provide feedback that is specific, empathic, and activating. Further, the prompt requires NiCo to incorporate preloaded knowledge where appropriate, reference specific classroom situations, and maintain an academic yet supportive tone to ensure a thoughtful, step-by-step conversation.

To provide the AI with information about teacher noticing and algebra education, we compiled a document on each topic and made it accessible for the AI.[5] This constitutes a platform-based variant of Retrieval-Augmented Generation (RAG; Lewis et al., 2020), in which external knowledge is semantically embedded and dynamically retrieved by the model to inform its responses. While we did not implement a custom retrieval mechanism, the internal architecture of the platform allows for context-sensitive access to uploaded documents during inference. This approach differs from traditional prompt-only conditioning and enabled more flexible, content-aware dialogue generation. However, we decided not to give the AI specific information about the videos that the PSTs watched in the seminars in order to explore how well the chatbot understood the situations and how the PSTs coped with describing them. As examples help the AI to better understand the intended outcome (Park and Choo, 2024), we created a set of questions and prompts to elicit teacher noticing based on previous studies in the field of teacher noticing (Jacobs et al., 2010; van Es et al., 2017) and included it in the meta-prompt. For the prompt part on providing feedback to PSTs, we draw on the study of Jacobsen and Weber (2023) and their high-quality prompt for providing feedback and generated a complementary example of feedback for a PST's remark.

To better understand how the meta-prompt and the chats with NiCo work, we present an excerpt of a chat from PST 1[6] with NiCo in Table 2. The chat deals with the beginning of the video vignette “Taxi ride”, where the teacher introduces the task. The chat transcript demonstrates how NiCo first asks for general characteristics of the lesson, before eliciting perceptions of a specific situation. In turn 6, NiCo draws the PST's attention to the fact, that they have already made judgments and provided PSTs' prior experiences and skills.

4.2 PSTs' Prior experiences and skills

To consider PSTs' prior experience in the context of research question two, we analyzed their OTL in mathematics pedagogy and, particularly, algebra education as the focus of the videos viewed in the seminar and OTL toward teacher noticing, digital media and AI (see Table 3) as the influence of OTL on students' knowledge has been shown in many studies. PSTs had addressed students mathematical thinking, basic mental models, functions, and algebra education in general as the focus of one seminar session or as part of several sessions on average, presumably providing them with a sound (at least overview) knowledge of these topics. However, with regard to specific topics in algebra education, such as ways of introducing algebraic expressions or equation equivalence, the PSTs indicated that, on average, this had been part of only one seminar session in their previous studies. This indicates that PSTs lack specific knowledge in this area and therefore may explain possible difficulties in perceptual noticing important events in the videos shown. In addition, nearly all PSTs had only few prior OTL toward teacher noticing, which indicates a lack of opportunities to develop their interpretive and decisional noticing skills. On average, digital applications and media had been the focus of several sessions, while AI applications in particular were the focus of only one session or were addressed as part of several sessions. Overall, these results suggest that PSTs had some prior experience with AI, but not integrated into mathematics educational topics. This was also in line with the PSTs' self-assessed AI competence, which the

5 Unfortunately, due to the processing restrictions of fobizz, we were only able to upload two documents with descriptions on the approach of teacher noticing (Bastian et al., 2022; Bastian et al., 2024) and basic mental models in algebra education (Korntreff and Prediger, 2021) and could not provide more extensive information to the AI.

6 To identify responses from the same person while maintaining anonymity, we assigned random numbers to the PSTs.

TABLE 2 Example chat[6].

No	Speaker	Text
1	PST 1	I would like to talk to you about a teaching situation.
2	NiCo	Gladly! At what grade level did the situation take place, what was the topic of the lesson and what kind of situation was it (e.g., a student question, a plenary discussion, a teacher's explanation, etc.)?
3	PST 1	The situation took place in a 7th grade class at a Gymnasium [academic-track school]. The topic of the lesson was "Describing functional relationships with terms and several variables." During this lesson, the students were to set up complex terms in real-world contexts using the example of the costs of a taxi ride. The situation involves the teacher introducing the lesson and explaining the assignment.
4	NiCo	That sounds like an interesting and stimulating lesson! Let's start by looking at how the situation was perceived. Perception: Please describe in detail what happened in the situation. What did you perceive in the situation? What did the teacher say and do, and how did the students react?
5	PST 1	The teacher reminds the class of a task called "table task" in which the students have already learned the variable concept as a variable. She uses the smartboard as visual support. Then she moves directly on to the assignment for today's lesson. The students are to describe the costs of a taxi ride with terms. The teacher already lists which fixed costs are included in a ride and which costs change per kilometer or per minute. She does not involve the students in the introductory situation and delivers a monologue. I only noticed one boy who seemed to be listening actively. All the other students are not looking at the teacher but gazing around the classroom dreamily, looking annoyed and unmotivated. Cognitive activation has not taken place.
6	NiCo	I noticed that you perceived the situation very accurately and described in detail what the teacher did and how the students reacted. That's an important first step. However, you immediately incorporated some interpretations, such as assuming that the students were annoyed and unmotivated. It would be helpful to separate these from the pure perception. Could you describe again what exactly you saw, without jumping to conclusions? This will help you to give more well-founded interpretations later on. I think it's great that you are so attentive and deal with the situation so intensively.

[6]This example chat and all other answer examples from the PSTs were originally in German and have been translated by the first author while preserving the meaning.

PSTs described as being, on average, in the middle range of the scale (see Table 3).

4.3 PSTs' feedback on NiCo

To examine PSTs' evaluations of NiCo, its usability, and their chats with it, and thus answer research question three, we first look at the rating scale feedback before taking a closer look at the advantages, disadvantages, and opportunities for improvement with the open-response feedback.

TABLE 3 PSTs' opportunities to learn and their self-assessed AI competence.

Variable	Mean	Minimum	Maximum
Opportunities to learn			
Algebra education	2.2	0	4
Students' mathematical thinking	1.6	0	3
Basic mental models (in general)	2.3	1	4
Basic mental model on variables	1.3	0	3
Algebraic expressions and their equivalence[a]	1.2	0	3
Equations and their equivalence[a]	1.2	0	3
Functions[a]	1.8	1	4
Teacher noticing	1.4	0	4
Use of digital media in the classroom	2.8	0	4
Use of AI applications in the classroom	1.7	0	4
Self-assessed AI competence	3.3	2.1	4.3

Opportunities to learn were assesses on a five-point Likert scale ranging from "No" (0) to "Yes, as the focus of an entire seminar or lecture" (4). Beliefs on self-assessed AI competence was assessed with 15 items on a five-point Likert scale ranging from "Strongly disagree" (1) to "Strongly agree" (5) and then scaled using a sum score. [a]from a mathematics pedagogical perspective.

4.3.1 Evaluation of PSTs concerning motivation to use, utility, future use and usability

For the rating scale items, we focused on four areas: motivation, utility, future use, and usability. The results for all items are shown in Table 4. PSTs rated chats with NiCo as motivating and enjoyable, indicating that the chatbot was able to motivate them to actively use it. This is consistent with the PSTs' feedback on future use, where the PSTs tended to indicate that they would use NiCo regularly in their daily teaching. However, this was only true for the specific example of NiCo, as PSTs on average do not plan to use more GenAI tools in general for their competence development, or at least were not inspired to do so by their work with NiCo. This could imply that NiCo's usefulness was perceived primarily as supporting teacher noticing, but not as a useful application of GenAI in general.

In terms of utility, the chats with NiCo were evaluated to support different facets of teacher noticing and to be useful for PSTs' professional development. PSTs described the dialogues with NiCo as being able to facilitate perceiving new details of teaching situations, analyzing them comprehensively (interpretation), and generating options for action in the observed teaching events (decision-making). This needs to be emphasized as it was one of the main aims of this exploratory study.

However, the PSTs also stressed that they did not learn anything new from the chats with NiCo. This was somewhat surprising as it seems to contradict the previous feedback. Based on comments from the open-response feedback, PSTs probably perceived this item as being related to new declarative mathematics (pedagogical) content knowledge. This may indicate that NiCo in its current form is not equipped to provide

TABLE 4 Feedback from PSTs on chats with NiCo.

Variable	Mean	Minimum	Maximum
Motivation			
Working with NiCo is motivating.	4.6	3	6
I enjoy the chats with the NiCo.	4.7	3	6
Utility			
I find NiCo's answers useful.	4.6	4	6
I find NiCo useful for my professional development.[a]	4.7	2	6
The chats with NiCo help me to analyze teaching situations well.	4.8	4	6
Through the chats with NiCo, I become aware of new details in teaching situations.	4.3	2	6
With the help of NiCo, I am able to develop various options for action for the observed lesson.	4.4	2	6
I learn something new from the chats with NiCo.	1.2	1	2
Future use			
I would use NiCo in my daily teaching to develop my skills.	4.3	2	6
I can see myself using NiCo on a regular basis.	3.9	2	6
Through the chats with NiCo, I now want to make greater use of (generative) AI tools for my competence development.	2.0	1	4
Usability			
Working with NiCo is technically easy.[a]	5.4	5	6
Working with NiCo is effortless.[a]	2.8	1	5
NiCo's responses were clear and easy to understand.[a]	3.1	1	5
I find working with NiCo straightforward and simple.[a]	4.3	2	6

PSTs' feedback on NiCo was assessed using a six-point Likert scale ranging from "Strongly disagree" (1) to "Strongly agree" (6). [a]The original item was worded negatively in the actual instrument. To improve readability, this item is inverted here to be consistent with the overall scale. The item text has been changed to the positive wording.

PSTs with new knowledge in these areas, which may be a direct result of the restrictions on uploading background material for NiCo's instruction. Although NiCo was not designed to convey knowledge but to promote teacher noticing skills, reflection, and theory-practice-linkages, this highlights an area for improvement.

In terms of usability, the PSTs indicated that working with NiCo was easy technically easy, but required some effort and that NiCo's responses were not always easy to understand. This will be considered more closely in the analysis of the open-response feedback.

4.3.2 Evaluation of PST concerning learning possibilities, advantages and support, disadvantages and difficulties and improvement

To take a closer look at the PSTs' feedback and opportunities for improvement of the chatbot, we asked the PSTs seven open-response questions and structured their feedback into the categories of learning outcomes, advantages and support, disadvantages and difficulties, and suggestions for improvement. The resulting category system is displayed in Table 5. We will now detail each category and provide some sample comments from the PSTs.

From the analysis of PST's open-ended responses for the category learning outcomes, it appears that PSTs recognized the importance of accurate, detailed observations (six PSTs) and the need to take time to reflect in order to make informed pedagogical decisions (two PSTs). For example, PST 14 stated: "NiCo has shown me that it makes sense to observe situations in class more intensively and in more detail", while another (PST 11) realized "that you can use AI to become more aware of your teaching style." Further, it became evident that working with NiCo gave the PSTs a better understanding of the importance of separating perception from interpretation when noticing classroom situations (three PSTs): "NiCo made me realize once again that when observing a lesson, one should make a very strict distinction between pure observation and subjective interpretation" (PST 1). Three PSTs mentioned that through the use of NiCo they learned to developed various courses of action for a specific situation. On a meta-level, PST 13 noted the usefulness of AI in education as a learning, while PST 3 reported a strengthening of their belief in not only observing but also to improving teaching.

Regarding the benefits of working with NiCo and the support provided by the chatbot, the PSTs focused on the process support in noticing teaching situations, sometimes from a general perspective, sometimes from a subject-specific perspective, as most answers could be assigned to categories in this area. In particular, the PSTs mentioned the chatbot's specific questions (seven PSTs) and feedback (six PSTs), as well as its support in describing situations (seven PSTs), developing alternative courses of actions (eight PSTs) and separating perception from interpretation (eight PSTs). For example, PST 6 stated "NiCo encourages you to provide even more details, to separate perception and interpretation, and then to develop alternative courses of action" and PST 1 mentioned "Support through: very specific questions, constructive feedback on one's own observations" as beneficial. Some PSTs also highlighted the reasons for the strict separation of perception and interpretation, e.g., PST 9 noted:

 10.3389/feduc.2025.1605921

TABLE 5 Category system and number of occurrences.

Category	*n*	Category	*n*
Learning outcomes			
Teacher noticing and reflection		***Meta level***	
Relevance of perception	6	Utility of AI	1
Separation of perception and interpretation	3	Beliefs	1
Development of various courses of action	3		
Relevance of (self-) reflection	2		
Advantages and support			
(Subject-specific) process support		***Usability and technology***	
(Constructive) feedback	6	Easy operation	6
(Specific) questions	7	NiCo's understanding of the situation	3
Support for (detailed) reflection	5	Speed of response	3
Support in (precise) perception/description	7	Comprehensibility of the answers	2
Support in developing alternative courses of action	8	***Affective-motivational support***	
Separation of perception and interpretation	8	Reassurance	1
Guidance and structure	4	Praise	2
Suggestions and explanations	2		
Disadvantages and difficulties			
(Subject-specific) process support			
Input for NiCo		***Usability and technology***	
Difficulty of separating perception and interpretation	3	Technical difficulties	3
Difficulty of the exact description	9	Response inconsistency	1
Output of NiCo		Writing effort	1
Missing or wrong focus	3	***Affective-motivational support***	
Lack of insight	1	Feedback	3
Questionable suitability/quality of the alternative actions	2		
Difficulties with the stimulus material	3		
Suggestions for improvement			
(Subject-specific) process support		***Usability and technology***	
Specific, content-related advice and alternative courses of action	6	Input options	4
Restructuring the dialogue	3	Easy access options	4
Examples	2	Access restrictions	1

Categories in italics are for structuring purposes only. They have not been assigned to an answer by themselves.

> "In general, the way NiCo wants to separate objective perception from interpretations and points them out worked very well. I can well imagine that (as with me) behaviors can be recognized that one would not have noticed on one's own."

Five PSTs mentioned NiCo's support in reflecting on the teaching situations: "By NiCos questions and delving deeper and deeper into the interpretation of classroom events, I have discovered things that I did not notice at first glance" (PST 14). The structured guidance and clear structuring of the dialogue through NiCo's responses (four PSTs) as well as its general suggestions and explanations (two PSTs) were also found to be helpful.

Considering the chatbots usability, PSTs highlighted the ease of use (six PSTs) and fast response time (three PSTs). Three PSTs also emphasized NiCo's good understanding of the situations described, e.g., PST 2 mentioned: "Understanding and addressing individual issues or aspects worked well. NiCo recognizes the context of the conversation and addresses it appropriately." The comprehensibility of NiCo's responses was perceived as supportive by 2 PSTs. In addition, at least some PSTs seemed to appreciate Nico's praise and reassurance provided in its feedback, as evidenced by three responses that fell into the affective-motivational support categories.

When considering the disadvantages and difficulties of working with NiCo, the PSTs again mostly evaluated the process

10.3389/feduc.2025.1605921

support from a mathematics pedagogical or general pedagogical perspective. The difficulty of describing the perceived situation in detail and answering the many questions NiCo asks was mentioned most often (nine PSTs). For example, PST 12 stated that "the frequent request [by NiCo] for a very detailed description of the situation is therefore a major obstacle," all the more so since one would not be able to replay and observe the situation in detail in daily teaching. PST 12 also commented:

> "Partially unclear work instructions: "Describe your perception". However, in the answer to my description, it would be criticized that I have already interpreted. Perhaps I have a wrong understanding of the term, but "perception" includes in my opinion also the interpretation."

This indicates the difficulty of separating perception and interpretation for some PSTs (three PSTS) but may also imply that NiCo could better justify its inquiry for the separation of perception and interpretation. In terms of NiCo's output, three PSTs complained about the lack of fit of NiCo's responses in the light of their intended focus, as evidenced by PSTs' comments: "Sometimes he [NiCo] gets stuck on a situation that I don't want to describe or shed light on in more detail" (PST 8). This could mean that the meta-prompt given to NiCo was not flexible enough to handle all of the PSTs' requests, but it could also indicate that the PSTs had difficulties in prompting NiCo, as their intention may not have been clear enough in their requests. Further, PST 2 noted the lack of new insight, which is consistent with the rating-scale items, and PST 9 and 10 discussed poor quality of the provided alternatives actions. The "laboratory situation" (PST 12), i.e., the artificial nature of the video vignettes, was also mentioned as an obstacle (three PSTs), as it made it difficult for some PSTs to cognitively engage with the stimulus material.

The PSTs criticized some minor technical difficulties (three PSTs), such as freezing and reloading of the website, the amount of writing required when communicating with NiCo and describing situations (PST 9), and some response inconsistencies when re-entering the same prompt (PST 5), which affected NiCo's usability. The latter may also be indicative of a lack of understanding of GenAI, as the probabilistic nature of its responses by design leads to differing answers when re-entering the same prompt. Finally, three PSTs noted that NiCo's feedback, which they perceived as "overly embellished" (PST 1) and repetitive, interfered with the AI's functionality. As some PSTs also mentioned these affective-motivational features as being supportive, this highlights heterogeneous preferences for feedback. Adaptable feedback that matches the preferences of the individual PSTs could be considered for further development of NiCo.

In terms of suggestions for improvement, PSTs made suggestions for the process support and the usability of NiCo. Most frequently, the PSTs asked for specific mathematics pedagogical advice and corresponding courses of action (6 PST). For example, PST 2 suggested:

> "That you might learn something new and get to know specific new methods. NiCo often says that you should incorporate exercises and questions. But what might that look like? It would be nice to get recommendations for such exercises and their formats."

This again underlines the need of the PSTs for subject-specific guidance and indicates that NiCo is not yet able to fully provide this. Three PSTs discussed changes to the structure of the dialogue with NiCo. PST 6 suggested describing the broad outline of the lesson first, rather than just the specific situation, so that NiCo could consider this when asking questions and perhaps make better fitting suggestions. PST 8 recommended that NiCo suggest topics to talk about in the described situation to avoid focus misunderstandings between the PST and NiCo. Moreover, two PSTs wanted suggestions on how to work with NiCo and example prompts. On the usability side, the PSTs mentioned improved input options–namely dictation/speech recognition and perhaps even video upload–to facilitate the dialogues with NiCo and reduce the writing effort. They also asked for easier access options, especially for smartphone use, to enable the use of NiCo in daily teaching life. PST 14 stated:

> "An app or other way to use NiCo on the go, if you don't have a laptop or similar with you, would be useful. Especially on the way home on the bus or train, you can review the day and it is precisely at such moments that NiCo could be helpful."

Finally, PST 13 advocated restricting access so that parents would not "mutate into amateur teachers or hobby pedagogical researchers". This indicates a potential conflict of use if NiCo is implemented on a larger scale.

5 Discussion and limitations of the study

5.1 Summary and discussion

In this study, we developed and tested an AI chatbot to facilitate PSTs noticing skills (perception, interpretation and decision-making) by acting as a dialogue partner and feedback authority on perceived classroom situations. The chatbot, NiCo, was developed based on the GenAI model GPT-4o, prompted with a meta-prompt using the latest suggestions for effective prompt engineering (Park and Choo, 2024) and applied chain-of-thought and query-loop methods (Schorcht et al., 2023) (see Supplementary material). We prompted NiCo to guide PSTs through the noticing of teaching situations in a structured way following our conceptualization of teacher noticing as perception, interpretation, and decision-making and supported the model via RAG prompting with external knowledge from previous studies on teacher noticing (Jacobs et al., 2010; Kaiser et al., 2015; van Es et al., 2017), which could be accessed dynamically to inform context-sensitive responses. Moreover, the prompt included detailed instruction for NiCo to provide specific, empathic, and activating feedback to the PSTs (Jacobsen and Weber, 2023) in the sense of persona prompting. By embedding our three-facet noticing framework directly into the chatbot's architecture, this work extends current AI-in-education research by operationalizing theoretical constructs within an interactive GenAI tool, a novelty that shifts AI applications from generic tutoring toward theory-driven pedagogical scaffolds.

As shown in the results (RQ1), the used meta-prompt facilitated a dialogue between the PSTs and NiCo, even though the GenAI was not given the specific teaching situation and only worked with the PSTs' descriptions. This allowed NiCo to be used in a variety of situations, including events from the PSTs' own experience, and provides initial evidence of the feasibility of the design used. However, the applicability to a variety of teaching situations may have reduced the specificity of NiCo's responses, as some PSTs noted in their feedback that the AI's questions and suggestions were not appropriate for the focused teaching situation and that they wished for more detailed, content-specific help.

In terms of prior experience and skills (RQ2), PSTs reported some prior experience with digital media and AI, and assessed their AI skills as moderate, indicating a sufficient skill base to work with GenAI. Their generally positive assessment of their skills and GenAI is consistent with previous research on PSTs' beliefs about AI and their willingness to apply AI tools (Thararattanasuwan and Prachagool, 2024; Zhang et al., 2023). However, some feedback revealed misunderstandings about how GenAI works, and a few PSTs requested scaffolding, such as prompt examples, for their work with NiCo. Thus, PSTs with difficulties or less experience in working with GenAI need to be considered when adapting the chatbot for future use. Moreover, although PSTs stated sufficient experience in mathematics pedagogy in general, they seemed to have less experience in specific aspects of algebra education, which were evident in the video-vignettes used, and the construct of teacher noticing. This may imply that the PSTs lacked some knowledge and skills for noticing the teaching situations in a productive way, which is in line with previous research (Lee et al., 2025; Zhang et al., 2024). Further, it could again explain why the PSTs asked for more content-specific help; finally, some PSTs criticized the separation of the three facets of teacher noticing.

Considering feedback (RQ3), the PSTs perceived their work with NiCo and the dialogues overall as motivating and useful to develop and improve teacher noticing skills in perception, interpretation and decision-making. PSTs reported that NiCo supported them in the detailed perception and comprehensive interpretation of teaching situations as well as the multiple generation of alternative courses of action. This provides evidence for the feasibility of an AI chatbot for facilitating dialogues about noticed teaching situations from the perspective of PSTs. This is also supported by the PSTs' willingness to use NiCo in the future. Overall, by centering user perceptions of AI-mediated noticing, our study contributes novel insights into how generative chatbots can foster reflective practice in ways that traditional professional development formats cannot, underscoring their potential for scalable, just-in-time professional learning.

However, PSTs indicated that they did not learn anything new from working with NiCo, which probably referred to mathematics pedagogical content knowledge and general pedagogical methods. Although the knowledge and method transfer was not the main purpose of NiCo, as its meta-prompt focused on facilitating noticing processes, linking PSTs' own knowledge to the teaching situation, and providing feedback, this highlights an area for improvement for future research. However, the currently available general GenAI models still show somewhat weak performance in content-specific areas of teacher education (Buchholtz and Huget, 2025). The platform used in this study did not allow to include much background knowledge for the GenAI. Future research could include a more robust knowledge base and adapt the basic GenAI model for the chatbot, for example using more specified retrieval-augmented generation methods (Zhao et al., 2024), to allow for more accurate and meaningful answers. This would also address the PSTs' request for more specific mathematics pedagogical advice and alternative courses of action. While our current implementation shares conceptual similarities with the logic of retrieval-augmented generation (RAG), it remains technically limited to a platform-based architecture, where relevant documents were statically uploaded. These documents were embedded and made available for semantic access during inference, but the system lacked mechanisms to actively select and retrieve specific pieces of information in response to the learners' prompts. In contrast, a fully developed RAG architecture would allow for real-time retrieval from a larger, structured knowledge environment, such as curriculum documents, annotated classroom transcripts, or didactic frameworks. Such a system could significantly enhance domain specificity, as large language models tend to reach their limits when it comes to generating pedagogically precise or content-sensitive instructional advice. The more specific the educational context, the more crucial it becomes to connect the GenAI's reasoning processes with trusted, structured educational resources.

Some PSTs found the feedback given by NiCo to be overly positive, repetitive, and, thus, intrusive, while others described the AI's feedback as helpful. This difference in feedback reception adds to the findings of Jacobsen and Weber (2023), who used a comparable prompt to elicit feedback. In the future, feedback preferences should be incorporated into the prompts while maintaining the quality of the prompt and the output.

Overall, the usability of NiCo was rated positively, with the PSTs stating that NiCo was easy to use, fast, and understandable in its questions and answers. However, PSTs criticized the amount of typing required to work with the chatbot and thus suggested easier input methods, such as speech recognition or direct audio input. This could also allow for a more natural experience, as PSTs could have a real conversation with the AI. Further, PSTs suggested creating easier ways to access NiCo. Hence, in order for teachers to be able to use a chatbot like NiCo to support their daily teaching experience, aspects of usability and accessibility need to be included in the creation of future chatbots.

5.2 Limitations of the study

Because this study represents an initial study with a small sample size, the results are not easily generalizable to other contexts and the number of category occurrences is not representative. The construction of the chatbot and the PSTs' feedback only allow an insight into the feasibility of the chatbot approach, the generation of hypotheses about its functioning, and the chatbots' strengths as well as areas of improvement. In addition, the constructed chatbot was limited since the platform used only allowed for a small amount of background material. Future research should include a more comprehensive knowledge base.

In addition, in this study, we only examined the feedback from the PSTs, thus focusing on a user-centered perspective. The usefulness of chatting with NiCo was only self-assessed, so no external assessment took place. Analysis of the actual chat transcripts in future research will provide further insight into the usefulness, functionality, and ability of a NiCo to improve PSTs noticing skills.

Furthermore, a fundamental limitation of our study may arise from the probabilistic nature of GPT-4o's responses. Because the model generates text by sampling from learned token distributions rather than via genuine semantic understanding, its feedback may lack the depth or novelty required to introduce new pedagogical insights. This probabilistic generation can lead PSTs to perceive the interaction as reflective yet ultimately reiterative, which may explain why some PSTs reported "learning nothing new." Future work should further explore hybrid architectures that combine deterministic, knowledge-grounded modules with generative language models to balance open-ended reflection with reliable, content-rich guidance, thereby addressing the tension between reflective scaffolding and the inherent unpredictability of GenAI outputs.

6 Conclusion

Given the need for scalable and efficient methods to develop noticing skills of (pre-service) teachers and new possibilities of AI application in teacher education, this study explored the use of a GenAI chatbot to provide individualized support and feedback to PSTs in noticing classroom situations. The findings show the potential perceived by PSTs to facilitate dialogues about noticing specific classroom situations, the overall good usability of the concept, and some areas for improvement. By embedding a theory-driven noticing framework directly into the chatbot's prompt design, we demonstrate a novel approach to operationalizing teacher noticing within AI-mediated interactions, advancing both research and practice.

Overall, this study illustrates an innovative way to foster teacher noticing skills, that may be introduced on a larger scale in the future and, thus, would allow more (pre-service) teachers to obtain support in the development of their noticing skills at an easily accessible level. Practically, NiCo could be integrated into existing teacher education and professional development programs as an on-demand reflective partner, allowing PSTs and in-service teachers to engage in targeted noticing exercises alongside traditional seminars, workshops or internships. This blended model would harness NiCo's scalable availability to reinforce workshop content, support peer-led video clubs, or supplement mentoring sessions.

For future research, we recommend to evaluate the impact of sustained NiCo use on observable changes in teacher noticing skills for example by classroom observation or through the use of AI-PST chats to complement the findings of the present study.

Furthermore, subject-specific adaptations of the chatbot need to be explored, for instance by embedding discipline-tailored knowledge bases to deepen the content relevance. Additionally, hybrid chatbot architectures need to be investigated that combine generative dialogue with deterministic, knowledge-grounded modules to balance open-ended reflection and domain specificity.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Written informed consent was obtained from the participants for the publication of any potentially identifiable images or data included in this article.

Author contributions

AB: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. NB: Conceptualization, Resources, Writing – review & editing. GK: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that Generative AI was used for the creation of this Manuscript. AI DeepL was used for language editing.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/feduc.2025.1605921/full#supplementary-material

References

Al-Shammari, A., and Al-Enezi, S. (2024). Role of artificial intelligence in enhancing learning outcomes of pre-service social studies teachers. *J. Soc. Stud. Educ. Res.* 15, 163–196.

Amador, J. M., Bragelman, J., and Superfine, A. C. (2021). Prospective teachers' noticing: A literature review of methodological approaches to support and analyze noticing. *Teach. Teach. Educ.* 99, 1–16. doi: 10.1016/j.tate.2020.103256

Bastian, A., Kaiser, G., Meyer, D., Schwarz, B., and König, J. (2022). Teacher noticing and its growth toward expertise: An expert–novice comparison with pre-service and in-service secondary mathematics teachers. *Educ. Stud. Maths.* 110, 205–232. doi: 10.1007/s10649-021-10128-y

Bastian, A., König, J., Ross, N., Klee-Schramm, I., Sommer, D., Strauß, S., et al. (2025). Teacher competence in inclusive mathematics education: Examining the effects of an innovative professional development program on teacher noticing. *Res. Sq.* [Preprint]. doi: 10.21203/rs.3.rs-5752892/v1

Bastian, A., König, J., Weyers, J., Siller, H.-S., and Kaiser, G. (2024). Effects of teaching internships on preservice teachers' noticing in secondary mathematics education. *Front. Educ.* 9:1–15. doi: 10.3389/feduc.2024.1360315

Berliner, D. C. (2001). Learning about and learning from expert teachers. *Int. J. Educ. Res.* 35, 463–482. doi: 10.1016/s0883-0355(02)00004-6

Blömeke, S., Gustafsson, J. E., and Shavelson, R. J. (2015). Beyond dichotomies: Competence viewed as a continuum. *Z. Psychol.* 223, 3–13. doi: 10.1027/2151-2604/a000194

Blömeke, S., Jentsch, A., Ross, N., Kaiser, G., and König, J. (2022). Opening up the black box: Teacher competence, instructional quality, and students' learning progress. *Learn. Instruct.* 79, 1–11. doi: 10.1016/j.learninstruc.2022.101600

Brennan, R. L., and Prediger, D. J. (1981). Coefficient Kappa: Some uses, misuses, and alternatives. *Educ. Psychol. Meas.* 41, 687–699. doi: 10.1177/001316448104100307

Buchholtz, N., and Huget, J. (2025). *Möglichkeiten für die Anwendung von ChatGPT für angehende Mathematiklehrkräfte - Von guten und schlechten Prompts. Mitteilungen der Mathematischen Gesellschaft in Hamburg.* Hamburg: University of Hamburg.

Buchholtz, N., Schorcht, S., Bata, K., Baumanns, L., Decker, M., Dilling, F., et al. (in preparation). Development of a survey instrument to assess mathematics pre-service teachers' beliefs towards AI in the teaching and learning of mathematics. *Comput. Educ. Open.*

Buchholtz, N., Schorcht, S., Baumanns, L., Huget, J., Noster, N., Rott, B., et al. (2024). Damit rechnet niemand! - Sechs Leitgedanken zu Implikationen und Forschungsbedarfen zu KI-Technologien im Mathematikunterricht. *Mitteilungen GDM* 117, 15–24.

Dindyal, J., Schack, E. O., Choy, B. H., and Sherin, M. G. (2021). Exploring the terrains of mathematics teacher noticing. *ZDM Maths Educ.* 53, 1–16. doi: 10.1007/s11858-021-01249-y

Fernández, C., Llinares, S., and Rojas, Y. (2020). Prospective mathematics teachers' development of noticing in an online teacher education program. *ZDM* 52, 959–972. doi: 10.1007/s11858-020-01149-7

Galiç, S., Urhan, S., Dost, Ş, and Lavicza, Z. (2025). Examining mathematics teachers noticing the rationality: Scenario-based training with AI Chatbot. *Sci. Educ.* 34, 2759–2790. doi: 10.1007/s11191-025-00618-3

Gheyssens, E., Consuegra, E., Engels, N., and Struyven, K. (2021). Creating inclusive classrooms in primary and secondary schools: From noticing to differentiated practices. *Teach. Teach. Educ.* 100:103210. doi: 10.1016/j.tate.2020.103210

Goodwin, C. (1994). Professional vision. *Am. Anthropol.* 96, 606–633. doi: 10.1525/aa.1994.96.3.02a00100

Hill, H. C. (2009). Fixing teacher professional development. *Phi Delta Kappan* 90, 470–476. doi: 10.1177/003172170909000705

International Organization for Standardization (2019). *Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems. ISO standard No. 9241-210:2019.* Geneva: International Organization for Standardization.

Jacobs, V. R., Lamb, L. L. C., and Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. *J. Res. Maths. Educ.* 41, 169–202.

Jacobsen, L., and Weber, K. (2023). The promises and pitfalls of ChatGPT as a feedback provider in higher education: An exploratory study of prompt engineering and the quality of AI-driven feedback. *AI* 6:35. doi: 10.31219/osf.io/cr257

Jong, C., Schack, E. O., Fisher, M. H., Thomas, J., and Dueber, D. (2021). What role does professional noticing play? Examining connections with affect and mathematical knowledge for teaching among preservice teachers. *ZDM Maths. Educ.* 53, 151–164. doi: 10.1007/s11858-020-01210-5

Kaiser, G., Blömeke, S., König, J., Busse, A., Döhrmann, M., and Hoth, J. (2017). Professional competencies of (prospective) mathematics teachers—cognitive versus situated approaches. *Educ. Stud. Maths.* 94, 1–22. doi: 10.1007/s10649-016-9713-8

Kaiser, G., Busse, A., Hoth, J., König, J., and Blömeke, S. (2015). About the complexities of video-based assessments: Theoretical and methodological approaches to overcoming shortcomings of research on teachers' competence. *Int. J. Sci. Maths. Educ.* 13, 369–387. doi: 10.1007/s10763-015-9616-7

Kasepalu, R., Prieto, L. P., Ley, T., and Chejara, P. (2022). Teacher artificial intelligence-supported pedagogical actions in collaborative learning coregulation: A wizard-of-Oz study. *Front. Educ.* 7:736194. doi: 10.3389/feduc.2022.736194

Kleinknecht, M., and Gröschner, A. (2016). Fostering pre-service teachers' noticing with structured video feedback: Results of an online- and video-based intervention study. *Teach. Teach. Educ.* 59, 45–56. doi: 10.1016/j.tate.2016.05.020

König, J., Santagata, R., Scheiner, T., Adleff, A.-K., Yang, X., and Kaiser, G. (2022). Teacher noticing: A systematic literature review of conceptualizations, research designs, and findings on learning to notice. *Educ. Res. Rev.* 36:453. doi: 10.1016/j.edurev.2022.100453

Korntreff, S., and Prediger, S. (2021). Verstehensangebote von YouTube-erklärvideos – konzeptualisierung und analyse am Beispiel algebraischer konzepte. *J. Math. Didaktik* 84, 1–30. doi: 10.1007/s13138-021-00190-7

Kuckartz, U. (2014). *Qualitative text analysis: A guide to methods, practice & using software.* Thousand Oaks, CA: SAGE, doi: 10.4135/9781446288719

Labadze, L., Grigolia, M., and Machaidze, L. (2023). Role of AI chatbots in education: Systematic literature review. *Int. J. Educ. Technol. High. Educ.* 20:56. doi: 10.1186/s41239-023-00426-1

Lachner, A., Jarodzka, H., and Nückles, M. (2016). What makes an expert teacher? Investigating teachers' professional vision and discourse abilities. *Instruct. Sci.* 44, 197–203. doi: 10.1007/s11251-016-9376-y

Lee, D., and Yeo, S. (2022). Developing an AI-based chatbot for practicing responsive teaching in mathematics. *Comput. Educ.* 191:104646. doi: 10.1016/j.compedu.2022.104646

Lee, D., Son, T., and Yeo, S. (2025). Impacts of interacting with an AI chatbot on preservice teachers' responsive teaching skills in math education. *J. Comput. Assist. Learn.* 41:e13091. doi: 10.1111/jcal.13091

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., et al. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *arXiv* [Preprint]. doi: 10.48550/arXiv.2005.11401

Lo, L. S. (2023). The CLEAR path: A framework for enhancing information literacy through prompt engineering. *J. Acad. Librariansh.* 49:102720. doi: 10.1016/j.acalib.2023.102720

Mason, J. (2002). *Researching your own practice: The discipline of noticing.* Milton Park: Routledge.

Metsäpelto, R.-L., Poikkeus, A.-M., Heikkilä, M., Husu, J., Laine, A., Lappalainen, K., et al. (2021). A multidimensional adapted process model of teaching. *Educ. Assess. Eval. Accountabil.* 85, 1–30. doi: 10.1007/s11092-021-09373-9

Mischo, C., Wolstein, K., and Peters, S. (2023). Professional vision of early childhood teachers: Relations to knowledge, work experience and teacher child-interaction. *Early Years* 43, 828–844. doi: 10.1080/09575146.2022.2028129

Mosher, M., Dieker, L., and Hines, R. (2024). The past, present, and future use of artificial intelligence in teacher education. *J. Special Educ. Prep.* 4, 6–17. doi: 10.33043/8aa9855b

Neumann, K., Kuhn, J., and Drachsler, H. (2024). Generative künstliche intelligenz in unterricht und unterrichtsforschung – chancen und herausforderungen. *Unterrichtswissenschaft* 52, 227–237. doi: 10.1007/s42010-024-00212-6

OECD (2019). *Providing quality early childhood education and care: Results from the starting strong survey 2018*. Paris: OECD Publishing, doi: 10.1787/301005d1-en

Olea, C., Tucker, H., Phelan, J., Pattison, C., Zhang, S., Lieb, M., et al. (2024). "Evaluating personified expert effectiveness," in *Proceedings of the 110th international conference on artificial intelligence and soft computing*, (Berlin), 63–81. doi: 10.5121/csit.2024.141106

OpenAI (2024b). *Hello GPT-4o*. Available online at: https://openai.com/index/hello-gpt-4o/ (accessed January 29, 2025).

OpenAI (2024a). *GPT builder – what is the GPT builder for in ChatGPT and why did we make it?*. Available online at: https://help.openai.com/en/articles/8770868-gpt-builder (accessed January 29, 2025).

Park, J., and Choo, S. (2024). Generative AI prompt engineering for educators: Practical strategies. *J. Special Educ. Technol.* 7, 1–7. doi: 10.1177/01626434241298954

Pepin, B., Buchholtz, N., and Salinas-Hernández, U. (2025). A scoping survey of ChatGPT in mathematics education. *Digit. Exp. Maths. Educ.* 11, 9–41. doi: 10.1007/s40751-025-00172-1

Santagata, R., König, J., Scheiner, T., Nguyen, H., Adleff, A.-K., Yang, X., et al. (2021). Mathematics teacher learning to notice: A systematic review of studies of video-supported teacher education. *ZDM Maths. Educ.* 53, 119–134. doi: 10.1007/s11858-020-01216-z

Saúde, S., Barros, J. P., and Almeida, I. (2024). Impacts of generative artificial intelligence in higher education: Research trends and students' perceptions. *Soc. Sci.* 13:410. doi: 10.3390/socsci13080410

Schorcht, S., Baumanns, L., Buchholtz, N., Huget, J., Peters, F., and Pohl, M. (2023). Ask smart to get smart: Mathematische Ausgaben generativer KI-Sprachmodelle verbessern durch gezieltes prompt engineering. *Mitteilungen GDM* 115, 12–23.

Schorcht, S., Buchholtz, N., and Baumanns, L. (2024). Prompt the problem – investigating the mathematics educational quality of AI-supported problem solving by comparing prompt techniques. *Front. Educ.* 9:1386075. doi: 10.3389/feduc.2024.1386075

Sherin, M. G., Jacobs, V. R., and Philipp, R. A. (2011). "Situating the study of teacher noticing," in *Mathematics teacher noticing*, eds M. G. Sherin, V. R. Jacobs, and R. A. Philipp (Milton Park: Routledge), 3–13.

Sherin, M. G., Richards, J., and Larison, S. (2025). "Professional vision and teacher noticing: Looking back and looking forward," in *Teacher professional vision. Theoretical and methodological advances*, Vol. 1, eds A. Gegenfurtner and R. Stahnke (Milton Park: Routledge), 29–42.

Son, T., Yeo, S., and Lee, D. (2024). Exploring elementary preservice teachers' responsive teaching in mathematics through an artificial intelligence-based Chatbot. *Teach. Teach. Educ.* 146:104640. doi: 10.1016/j.tate.2024.104640

Stürmer, K., Seidel, T., and Schäfer, S. (2013). Changes in professional vision in the context of practice. *Gruppendynamik Organ.* 44, 339–355. doi: 10.1007/s11612-013-0216-0

Tarantini, E. (2023). "Reflective teacher education in the digital age," in *Open and inclusive educational practice in the digital world*, eds D. Ifenthaler, D. G. Sampson, and P. Isaías (Cham: Springer International Publishing), 213–231. doi: 10.1007/978-3-031-18512-0_13

Thararattanasuwan, K., and Prachagool, V. (2024). Exploring perspectives of teacher students toward generative AI technologies. *Int. Educ. Stud.* 17, 22–28. doi: 10.5539/ies.v17n5p22

Thomas, J., Dueber, D., Fisher, M., Jong, C., and Schack, E. O. (2020). Professional noticing into practice: An examination of inservice teachers' conceptions and enactment. *Invest. Maths. Learn.* 12, 110–123. doi: 10.1080/19477503.2019.1681834

van Es, E. A. (2011). "A framework for learning to notice student thinking," in *Mathematics teacher noticing*, eds M. G. Sherin, V. R. Jacobs, and R. A. Philipp (Milton Park: Routledge), 134–151.

van Es, E. A., and Sherin, M. G. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *J. Technol. Teach. Educ.* 10, 571–596.

van Es, E. A., Cashen, M., Barnhart, T., and Auger, A. (2017). Learning to notice mathematics instruction: Using video to develop preservice teachers' vision of ambitious pedagogy. *Cogn. Instruct.* 35, 165–187. doi: 10.1080/07370008.2017.1317125

VERBI Software. (2024). *MAXQDA analytics Pro 2024*. Available online at: https://maxqda.de (accessed June 1, 2024).

Weber, K. E., Gold, B., Prilop, C. N., and Kleinknecht, M. (2018). Promoting pre-service teachers' professional vision of classroom management during practical school training: Effects of a structured online- and video-based self-reflection and feedback intervention. *Teach. Teach. Educ.* 76, 39–49. doi: 10.1016/j.tate.2018.08.008

Wei, Y., Zhang, Q., Guo, J., and Chen, M. (2023). Learning to teach through noticing: A bibliometric review of teacher noticing research in mathematics education during 2006–2021. *Hum. Soc. Sci. Commun.* 10, 1–5. doi: 10.1057/s41599-023-01718-7

Weyers, J., König, J., Scheiner, T., Santagata, R., and Kaiser, G. (2023). Teacher noticing in mathematics education: A review of recent developments. *ZDM Maths. Educ.* 5, 1–16. doi: 10.1007/s11858-023-01527-x

Zawacki-Richter, O., Marín, V. I., Bond, M., and Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *Int. J. Educ. Technol. High. Educ.* 16:39. doi: 10.1186/s41239-019-0171-0

Zhang, C., Schießl, J., Plößl, L., Hofmann, F., and Gläser-Zikuda, M. (2023). Acceptance of artificial intelligence among pre-service teachers: A multigroup analysis. *Int. J. Educ. Technol. High. Educ.* 20:49. doi: 10.1186/s41239-023-00420-7

Zhang, N., Ke, F., Dai, C.-P., Southerland, S. A., and Yuan, X. (2024). Seeking to support preservice teachers' responsive teaching: Leveraging artificial intelligence-supported virtual simulation. *Br. J. Educ. Technol.* 26, 1–22. doi: 10.1111/bjet.13522

Zhao, P., Zhang, H., Qinhan, Y., Wang, Z., Geng, Y., Fu, F., et al. (2024). Retrieval-augmented generation for AI-generated content: A survey. *arXiv* [Preprint]. doi: 10.48550/arXiv.2402.19473

Zheng, M., Pei, J., Logeswaran, L., Lee, M., and Jurgens, D. (2024). *When "a helpful assistant" is not really helpful: Personas in system prompts do not improve performances of large language models: Findings of the association for computational linguistics: EMNLP 2024 (15126–15154)*. Kerrville, TX: Association for Computational Linguistics.