PRE-SERVICE TEACHERS' DEVELOPING BELIEFS ABOUT AI IN EDUCATION: A SEMESTER-LONG STUDY

Benjamin Rott, André Bresges

University of Cologne

This longitudinal qualitative study investigates how pre-service teachers' experiences with and beliefs about artificial intelligence (AI) tools develop during a semester-long university seminar. Fifteen pre-service teachers participated in a seminar on mathematics and physics education with weekly sessions. They explored and reflected upon educational applications of AI chatbots and reported on their experiences. Data were collected at three time points through open questions and analysed using qualitative content analysis. Results show clear development from limited prior experience toward confident and reflective use, accompanied by a shift from vague perceptions to more nuanced views emphasizing opportunities while acknowledging risks. Implications are discussed.

INTRODUCTION & BACKGROUND

The rise of artificial intelligence (AI) promises a revolution of daily life including education. Since the release of ChatGPT in November 2022, AI chatbots, built on large language model technology (Buchholtz et al., 2024), have become very popular. Main uses of AI chatbots are generating human-like text and carrying out a variety of tasks, including answering questions, writing essays, summarizing and translating texts, or generating pictures. Criticism of AI chatbots often refers to their tendency to generate plausible-sounding but incorrect or nonsensical answers (so-called hallucinations) and to reproduce biases present in their training data (e.g., Jancarik & Dusek, 2024). Still, the quality of their output is increasing so fast, that by today they are even used for scientific discoveries (cf. Jee, 2023). From a student's perspective, tasks and most non-routine problems from school mathematics can be easily solved by popular AI chatbots. And from a teacher's perspective, they can help in lesson planning, designing tasks (including differentiated versions for advanced learners and learners at risk), and even grading students' tests (for an overview, see Almarashdi et al., 2024).

In a recent "systematic review of ChatGPT in transforming mathematics teaching and learning," Almarashdi et al. (2024) identified 31 articles reporting on "empirical evidence on the use of ChatGPT and the extent to which ChatGPT has been playing a facilitative and transformative role in mathematics teaching and learning" (Almarashdi et al., 2024, p. 4). The reviewed studies covered topics such as student support and personal learning, teacher assistance and instructional planning, and mathematical proficiency and technical limitations. However, given their already existing capabilities, further studies exploring the benefits and limitations of AI chatbots in

2026. In H. Krzywacki, M. Hähkiöniemi, E. Haataja, M.S. Hannula & P. Hästö (Eds.). *Proceedings of the 49th Conference of the International Group for the Psychology of Mathematics Education: Research Reports*, pp. 1243-1250. PME.

school contexts are unnecessary, in our view. To make matters worse, because of the rapid development of AI chatbots, such studies are often outdated before publication. Instead, we need to research how users interact with AI chatbots, the risks and opportunities they perceive, how they intend to use utilize such technologies, and how their beliefs evolve.

Nevertheless, studies addressing users' beliefs (in the sense of denotative beliefs on a continuum between subjective and objective knowledge, see Rott, 2021, p. 3 ff.) and the development of these beliefs are scarce. Although some of the studies reviewed by Almarashdi et al. (2024) address users' perspectives on using ChatGPT's to teach mathematics or similar topics, none explicitly cover this topic. Zhuang and Zhang (2025, p. 3) come to a similar conclusion: "Most existing research relies on short-term investigations and survey-based data to examine teachers' perceptions of AI, with little attention to how these perceptions evolve through sustained engagement with AI tools over time."

In fact, one of the few studies addressing the beliefs of pre-service teachers' (PSTs) about AI (chatbots) and the development of these beliefs was conducted by Zhuang and Zhang (2025). In a mathematics methods course, the researchers investigated PSTs perceptions of using generative AI to support teacher questioning strategies. Over the course of a semester, the researchers conducted two written reflections, two surveys, and a final interview with six PSTs. Zhuang and Zhang reported a development towards more positive views and perceived usefulness of AI with increased familiarity and improved prompt-writing proficiency. Additionally, less experienced PSTs reported benefiting from simulated student responses generated by tools such as Student GPT.

As Zhuang and Zhang (2025) conclude, a single study with a small and non-diverse sample can only be the beginning of research on these topics. Therefore, the *rationale* of this study is to *explore PSTs' beliefs about using AI in educational settings and how these beliefs evolve when they actively use AI chatbots in these settings*.

THE STUDY

In the summer term of 2024, the authors held a joint seminar in mathematics and physics education at a university in the centre of Germany on the topic of "Research and Development in the Classroom". 15 PSTs (either at the end of their Bachelor studies or in their Master studies, 7 female & 8 male) participated in weekly 1.5-hour sessions for 14 weeks. Using techniques from *design thinking* approaches (e.g., Lawson, 2006), the PSTs explored possible uses of ChatGPT, Perplexity.ai, and other AI chatbots in supporting teachers' work. These activities included lesson planning, developing introductory stories to interest and motivate students, analysing students' written work, and designing a mega-prompt for an AI chatbot to act as a tutor fostering students in conducting physical experiments. The PSTs also visited a local school to conduct a field test of their ideas in three groups of 3, 5, and 7 PSTs.

At the beginning [T1], in the middle [T2], as well as at the end [T3] of the semester, the students were asked to reflect on their thoughts about AI by answering four open-text questions:

> [Q1] What experience have you already gained with artificial intelligence (interactions with chatbots like ChatGPT, programming machine learning algorithms, etc.)?
>
> [Q2] Do you see AI more as an opportunity or a threat? Please explain your answer.
>
> [Q3] Can you imagine using AI support in your teaching (as a teacher)? What specific applications do you have in mind?
>
> [Q4] Finally, we ask you to create a concept map (this can also be included here as a photo/screenshot). The goal is for you to illustrate which terms come to mind regarding the general topic of "AI support in the classroom" and what connections you see between them.

Using a personalized pseudonym, the PSTs wrote their answers into a CryptPad (an online document that works similar to a shared google.docs document without collecting data of the users); thus, texts from previous answers were visible to the PSTs when answering for the second or third time, allowing them to refer to and reflect upon their previous texts when answering for the second or third time, respectively. Answers were collected by the researchers after each written reflection so that previous answers could not be altered retroactively. Unfortunately, creating a concept map was not easy using the CryptPad, leading to fewer answers to the fourth question on all three times of survey. Thus, the concept-map prompt will not be analysed in this report.

All 15 PSTs completed the questionnaire on the first time of survey; 12 PSTs answered the questionnaire on the second time of survey; and 8 PSTs answered on all three points of survey. A review of the PSTs' answers by the first author suggests meaningful and interpretable responses. This evaluation is supported by the length of the answers: Not counting omitted answers, on average, the PSTs used 40.3 words for each answer (min: 3, max: 307).

To analyse the PSTs written answers, the first author read all students' answers and coded them according to the Qualitative Content Analysis framework (Mayring, 2014), using a very basic category system to identify meaningful responses to questions 1-3. For example, answers to question 2 were either coded as "opportunity" or "threat" and then paraphrased. To illustrate this, take a look at K11S07's answer to question 2 at T1 (translated by the first author.): "Opportunity, as people can be supported and relieved in their work. AI could be used as a supplement here. Processing of large amounts of data possible (both an opportunity and a risk, but with appropriate regulations, the advantages outweigh the disadvantages)." In the coding process, this was shortened to "opportunity: support (unspecified)", "opportunity: data processing", and "risk: data processing".

To validate the human codes, the first author used Google's NotebookLM to analyse the data. The first step was ensuring that the "virtual research assistant" had processed

the data correctly by answering basic questions such as “How many PSTs answered the third question at the end of the semester?” or “What opportunities did K11S07 list at the beginning of the semester?” and comparing the answers to the data coded by the first author. Perfect answers to the basic questions were achieved after preparing the raw data for the data import for the third time, ensuring that NotebookLM could process all data correctly. A second step was to pose advanced prompts such as “Summarize the opportunities listed by the PSTs at the beginning of the semester in a table. For each entry, name the source in an additional column.” Doing this, the results identified by the human rater could be validated.

RESULTS

PSTs’ experiences with AI applications

The first question addresses the experiences of the PSTs regarding AI use. Even though AI chatbots had been publicly available for almost two years, at the beginning of the semester [T1], the majority of the PSTs had little direct experience with such technology. Typical answers include “I have had little direct experience with artificial intelligence so far.” (R12U11) and vague statements such as “I already have experience with ChatGPt and CoPilot. [sic]” (M05F02) The few specific uses of AI chatbots mentioned at this point include generating ideas (brainstorming), shortening texts, and translating texts (which, according to the respective PST, had not been carried out satisfactorily).

There is a clear progression regarding the PSTs’ experiences which shows at the middle and especially at the end of the semester. They refer to activities from the seminar and often add reflective thoughts. For example, A11H12 wrote at the end of the semester [T3]: “I have now created my own ChatGPTs and (mega) prompts. I discovered that even with the same prompt, ChatGPT produced completely different outputs. […] I have now also created my own images (in my free time) with the help of AI. I tried to create worksheets with ChatGPT. These did not look very good yet.”

NotebookLM came to a similar conclusion, summarizing the development of the PSTs’ experiences as follows: “While the students at the beginning of the semester experienced AI as *users* and in the middle of the semester as *applicators*, at the end of the semester they developed into *developers* and *configurators* of AI-supported tools.”

Perceived risks and opportunities regarding AI use

At the beginning of the semester, answering the second question regarding potential benefits or threats of AI use, the PSTs mentioned both risks and opportunities with roughly equal frequency. However, answers were often vague; for example, E09M02 mentioned “opportunity: make work easier” and “risk: relying too heavily on AI” [T1].

Over the course of the semester, the answers got more specific (which is in line with the PSTs’ greater experience, see above) and tend to emphasize opportunities over risks. In the middle of the semester [T2], E09M02 writes “I still believe that AI offers

an opportunity to provide support and also to relieve the burden on teachers, for example." Codes were given as follows: "opportunity: save work and time", "opportunity: support students during lessons when the teacher has no time". At the end of the semester [T3], E09M02 reflects: "I increasingly see AI support as an opportunity, but we must be aware of the risks. AI is a tool and should only be viewed as such." She then stated "opportunity: develop ideas, get inspiration", "opportunity: support (for teachers)". Five of the eight PSTs that responded to all three points of the survey showed such a development. The other three PSTs (M05F02, R11Ö03, and K04M06) only briefly responded to the second question of the third survey, which makes it difficult to detect a development.

AI use for the PSTs' teaching

Regarding the third question about the (potential) use of AI in their teaching, again, a clear progression is visible in the PSTs' answers. At the beginning of the semester, most of the students were confident that they would use AI in their teaching and preparation of teaching. However, many had no clear idea of how to do so (e.g., "Yes, I can." by M05F02 [T1] or "I can imagine it, but not yet exactly where. I will certainly use the algorithms I've learned about with the children at some point." by K04M06 [T1]) or had only vague ideas how to use AI specifically (e.g., "Yes, because as a teacher you always have a lot to do and if you are offered help or inspiration, you gladly accept it." by M06Z04 [T1]).

Later in the semester, the PSTs answer this question with more specific uses of AI regarding teachers (lesson planning, differentiating for weak and strong students, creating tasks and worksheets) and students (explanation of difficult expressions and technical terms, answering questions in the absence of teachers, serving as an assistant with guidance specified by a teacher with a "mega prompt"). This progression in the PSTs' perceived use cases can be illustrated by E02J11's answers (translated):

> [T1] Question 3: Use for lesson planning --> faster preparation
>
> Include in lesson planning --> let students work with it
>
> [T2] Question 3: Again, I can only answer YES. I already enjoy using it in my studies and can imagine continuing to do so in the future. It simply saves me a lot of time and allows me to focus on other important aspects of my work.
>
> [T3] --> I can still answer that question with a yes. The seminar taught me how to develop a prompt and how to use it. I find this application very valuable for teachers, as it allows them to tailor their lessons to individual needs and make them more efficient. This frees them up to concentrate on other important aspects of teaching.

DISCUSSION

One goal of this study was to better understand PSTs' knowledge of and experience with AI tools prior to explicitly covering them in a seminar. Most of the PSTs reported having some experience with AI chatbots prior to the seminar. Only A11H12, K04M06, and R12U11 responded that they do "not have that much experience with

AI, yet" [T1]. Five PSTs (E02J11, K11S07, S07B07, K04M06, and B09P05) wrote about specific knowledge of machine learning algorithms, neural networks, and programming with languages such as Python. K04M96's case is interesting, because he reported experience with automata theory (Turtle Coder, Dynamic Labyrinths) from a previous lecture, but no experience with AI chatbots. The other PSTs described "basic experience" with AI chatbots, such as occasionally generating texts or images and performing translations. However, they had no systematic or in-depth knowledge of AI, and only had vague ideas about how AI could help them in their future profession as teachers. These findings are consistent with those of Zhuang and Zhang (2025).

The second goal was to analyse how the PSTs' experiences with and beliefs about AI tools, generally and specifically for teaching, develop over the course of a seminar that encourages, supports, and reflects on the use of AI tools. The PSTs' responses to the reflective questions clearly demonstrate their improved understanding of how to leverage AI for professional purposes, while also recognizing the limitations and potential drawbacks of these technologies, such as unpredictable outputs, criminal applications (e.g., deepfakes), and changes in the job market. The PSTs attribute this change to their significant gain of experience using AI and reflecting on its use within their groups and in the seminar as a whole.

Interestingly, most answers point to generic teaching-related uses of AI instead of *mathematics- or physics-specific uses*. Even though 11 of the 15 PSTs are studying to become mathematics teachers, specific uses of AI regarding mathematics lessons are rare among the answers to any of the four questions at any point in time. Only three PSTs (A11H12, H05H06, and S07B07) mention mathematics specifically and only at the beginning and at the middle of the semester, not at the end of the semester, when answers are most specific and based on experiences. They suggest that ChatGPT can help with statistical tasks and calculate values such as the mean value or that AI can generally be used as a calculator. However, A11H12 adds that so far, calculation tasks were not solved satisfactorily by the AI chatbot [T3, Q1].

For the PSTs, it seemed to be easier to point out potential uses of AI in scientific lessons, specifically physics lessons. Of the 15 students in the sample, 8 are studying to become physics teachers (please note, in Germany, PSTs study to become teachers for two subjects; in this case, 4 of the 8 PSTs for physics also have mathematics as one of their subjects; thus, the 11 mathematics and 8 physics PSTs add up to 15 students in total). Even though less PSTs study physics than mathematics, there were twice as many physics-related answers than mathematics-related. The PSTs suggested to use AI chatbots to help with experimenting, that is explaining experimental setups, creating and checking safety instructions, assisting as a "second teacher" and giving step-by-step guidance, troubleshooting, and even data analysis and data visualization, as well as providing background information regarding experiments.

These results, of course, have *limitations*. The sample size is small, and although the study is longitudinal, the timeframe is limited to one semester with no follow-up

reflection one or two semesters later. Thus, it would be beneficial to replicate the study with a larger, more diverse group (e.g., PSTs at the beginning of their university studies) and to add a fourth measurement time in a later semester.

CONCLUSION

The results are consistent with unsystematic observations of the authors and many colleagues in the teaching profession (not covered in this study): Learners—that is, school or university students—tend to use AI chatbots for simple internet searches and to complete requirements set by their institutions, such as homework or weekly worksheets for their tutorials. However, most learners do this unreflectively, relying on the AI chatbot's initial responses. They often copy and submit the AI-generated responses without further prompting or checking them. Understanding the limitations of AI-generated responses and the necessity for mitigation strategies (Danyaro et al., 2025; Jancarik & Dusek, 2024; Sawyer, 2024), such as prompt engineering or fine-tuning (e.g., Schorcht et al., 2024), highlights the risks and missed-out opportunities associated with this observed behaviour.

Therefore, it seems necessary to offer students, including PSTs, opportunities to learn about, to experience, and to reflect upon the use of AI chatbots. PSTs need to learn to not trust AI chatbots without reflecting upon their answers and how to improve AI output. This is the conclusion of the authors as well as the PSTs, for example, see the following quotes:

> "However, I also see the danger of "mindless" use. I have observed, both in others and in myself, that the simplest questions are typed into the AI and people no longer think for themselves. In some cases, AI is used not as an aid, but as a solution" (H05H06, part of Q2 at T3).

> "Thanks to the seminar, I now know how to ask Chat GPT more specific questions about a piece of information and how to get a better answer." (E02J11, part of Q1 at T2)

> "Nevertheless, one must be aware of the dangers. And don't always trust AI answers 100%, but see them as an aid and still think about them yourself and check one or two answers." (E02J11, part of Q2 at T3)

Another reason PSTs should be trained in AI use is to learn about potential applications of which they were previously unaware. Examples mentioned by the participants of this study include "troubleshooting when programming with Python" (R12U11, Q1 at T2), "explaining complex concepts in simple terms" and "personalized feedback" (B09P05, Q2 at T2), as well as "mega prompts or even custom GPTs" (S07B07, Q1 at T3). By the end of the semester, some students even reflect on the importance of data privacy when working with AI. They suggest solutions such as using only specific (non-commercial) chatbots, school tablets instead of personal ones, and communicating with a chatbot through a client or interface instead of directly.

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), 1–14. https://doi.org/10.29333/ejmste/15739

Buchholtz, N., Schorcht, S., Baumanns, L., Huget, J., Noster, N., Rott, B., Siller, H.-S., & Sommerhoff, D. (2024). Damit rechnet niemand! Sechs Leitgedanken zu Implikationen und Forschungsbedarfen zu KI-Technologien im Mathematikunterricht [No one expects that! Six key ideas on the implications and research needs of AI technologies in mathematics education]. *Mitteilungen der GDM, 117*, 15–24.

Danyaro, K. U., Abdullahi, S., Abdallah, A. S., & Chiroma, H. (2025). Hallucinations in Large Language Models for Education: Challenges and Mitigation. *International Journal of Teaching, Learning and Education, 4*(6), 13–19. https://dx.doi.org/10.22161/ijtle.4.6.2

Jancarik, A., & Dusek, O. (2024). The problem of AI hallucination and how to solve it. In F. Moriera (Ed.), *Proceedings of the 23rd European Conference on e-Learning – ECEL 2024* (pp. 122–128). ACI, UK.

Jee, H. (2023). Emergence of artificial intelligence chatbots in scientific research. *Journal of exercise rehabilitation, 19*(3), 139–140. https://doi.org/10.12965/jer.2346234.117

Lawson, B. (2006). *How Designers Think*. Routledge.

Mayring, P. (2014). *Qualitative content analysis: Theoretical foundation, basic procedures and software solution*. Klagenfurt. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-395173

Rott, B. (2021). *Epistemological beliefs and critical thinking in mathematics*. Springer. https://doi.org/10.1007/978-3-658-33539-7

Sawyer, A. G. (2024). Artificial Intelligence Chatbot as a Mathematics Curriculum Developer: Discovering Preservice Teachers' Overconfidence in ChatGPT. *International Journal on Responsibility 2024, 7*(1), 1–26. https://doi.org/10.62365/2576-0955.1106

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*, 1–15. https://doi.org/10.3389/feduc.2024.1386075

Zhuang, Y., & Zhang, S. (2025). Pre-service mathematics teachers' perceptions of using GenAI for practicing teacher questioning: A semester-long study. *EURASIA Journal of Mathematics, Science and Technology Education, 21*(9). 1–16. https://doi.org/10.29333/ejmste/16764