Journal of Mathematics Teacher Education in Texas
Volume 15, Issue 3 (2026)

Using a Generative AI Digital Teaching Simulation to Examine Pre-service Teachers' Skills in Eliciting a Student's Mathematical Understanding

Shreyashi Halder
ETS

Jamie N. Mikeska
ETS

Calli Shekell
Pennsylvania Western University

Heather Jorgenson
George Washington University

Background

Eliciting students' thinking is a core teaching practice in mathematics, as it enables teachers to determine what a student understands and make informed instructional decisions (NCTM, 2014; TeachingWorks, 2019). When eliciting students' ideas, teachers pose questions that allow students to share their understanding of mathematical concepts, describe their thought processes, and explain their reasoning (Orr & Bieda, 2020). Teachers make in-the-moment instructional decisions about which ideas are worth exploring and support students in articulating their thoughts clearly based on their understanding (Shaughnessy & Boerst, 2018). However, research has shown that pre-service teachers (PSTs) often struggle to elicit students' thinking effectively in mathematics (Krupa et al., 2017; Monson et al., 2020), tending instead to use questions or comments aimed at clarifying students' understanding (Hallman-Thrasher, 2017). This teaching approach promotes teacher-to-student knowledge transmission rather than fostering student-centric classroom discourse (Molinari & Mameli, 2013; Wang et al., 2021).

Research studies found that PSTs can improve this core instructional practice through practice-based learning (Tyminski et al., 2021; Wang et al., 2021) that involves "targeted practice with repeated opportunities to rehearse and refine this skill" (Hallman-Thrasher, 2017, p. 545). Digital teaching simulations serve as approximations of practice that offer PSTs a flexible and repeated practice space to experiment with mathematics teaching strategies (Mikeska et al., 2023; Wang et al., 2021). Recently, generative artificial intelligence (GenAI) has been integrated into digital simulations to automate the systems that support personalized learning (Bywater et al., 2025; Mikeska, Beigman Klebanov et al., 2025) through repeated practice (Law, 2024). Incorporating GenAI can reduce the limitations of traditional digital teaching simulations, which are challenging to access for repeated practice at any time (Bondie et al., 2021), and increase the system's scalability (Bywater et al., 2025; Mikeska, Beigman Klebanov et al., 2025). In our study, we explored the potential of a GenAI teaching simulation to provide: (a) PSTs with the opportunity to practice eliciting student thinking and (b) mathematics teacher educators (TEs) with valuable insights about what their PSTs learn from eliciting a GenAI student's problem-solving process and conceptual understanding. This study addresses two research questions. Based on PSTs' interactions eliciting student thinking in the GenAI teaching simulation, we ask: (1) How accurately do PSTs describe the GenAI student's process for solving an addition word problem? (2) To what extent do PSTs identify the GenAI student's mathematical understanding and knowledge gaps?

Method

Study Context

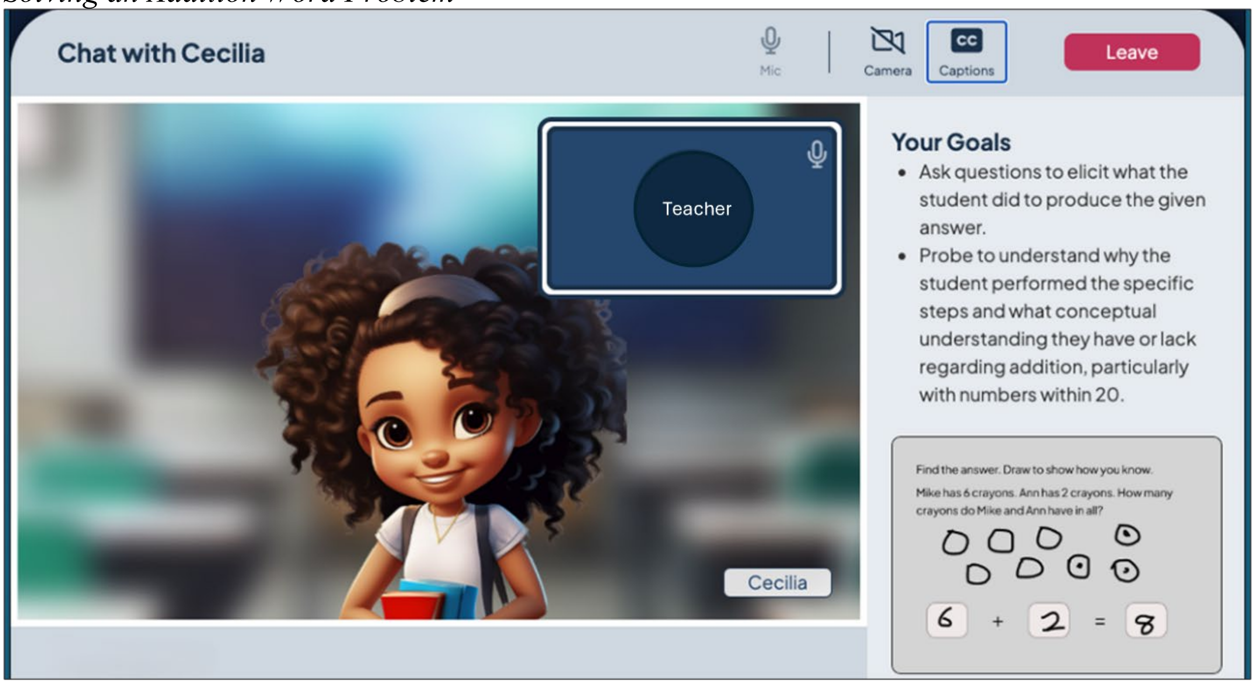
This case study took place during Spring 2025 in an elementary mathematics methods course taught by a TE at a private college in the Northeast region of the United States. Eleven of 12 PSTs initially consented to participate in the study; however, one participant could not engage in the simulation due to technical challenges. All participants were English speakers aged between 18 and 24 years and held a high school diploma or equivalent degree. Five participants had no prior teaching experience, two had substitute teaching experience, two had worked as mentors, and one was an after-school program coordinator.

Data Collection

As part of their coursework, the PSTs participated in two instructional cycles using the Strategies for Adding GenAI simulation. In the simulation, PSTs interacted with a first-grade GenAI student named "Cecilia" and elicited her process for solving an addition problem and her conceptual understanding about this mathematics topic (see Figure 1). Each instructional cycle included four main steps.

Figure 1

Online Interface of the GenAI Teaching Simulation, Displaying Cecilia's Written Work for Solving an Addition Word Problem



Note. Adapted from "Generative AI Teaching Simulations as Formative Assessment Tools Within Preservice Teacher Preparation" by J. N. Mikeska, A. Bhatia, S. Halder, B. Beigman Klebanov, T. Maxwell, B. Longwill, K. Behl, and C. Shekell, 2025, in *Proceedings of the Artificial Intelligence in Measurement and Education Conference (AIME-Con): Full Papers*, p. 214. Copyright 2025 by the National Council on Measurement in Education.

First, the PSTs prepared for the simulation by reviewing the simulation's goal and information about the scenario. The preparation phase included reviewing Cecilia's written work and considering questions PSTs could use to elicit her problem-solving process and conceptual understanding of adding numbers within 20. According to the pre-designed GenAI student profile (Mikeska, Bhatia et al., 2025), Cecilia understands: (a) the meaning of the numbers, (b) how to count on from one number, and (c) the use of models (in this case, the use of drawings) to represent or solve the addition problem. The student does not understand: (a) how to add numbers fluently via mental math, (b) that addition is commutative, (c) that counting on from the larger addend (regardless of the position) is more efficient, and (d) how to use other strategies (besides drawing and counting on) to solve addition problems. This level of detail in the student profile was not included in the instructions for the PSTs since PSTs are intended to elicit the student's understanding and misunderstandings during their conversation with Cecilia. Second, they engaged in a verbal conversation with Cecilia within the online platform. Third, they completed a post-simulation survey. PSTs also completed an online survey after the first simulation to provide information about their personal and professional backgrounds.

Finally, after all experiencing the same simulation, they engaged in a reflection session during their 50-minute class time, which was facilitated by their mathematics methods course TE. In the reflection activity after the first simulation, PSTs reviewed Hallman-Thrasher et al.'s (2020) categorization of questions by purpose. They then analyzed their own transcripts and categorized the prompts they used in their conversation with Cecilia. They also reflected on potential modifications to their questioning in the future. Afterwards, each PST compared their transcript with a partner to further reflect on how to elicit Cecilia's thinking and adjust their prompts in the future. In the reflection activity after the second simulation, which occurred about one week after the first simulation, the PSTs categorized their questions again, this time both categorizing their questions and focusing on the information each question garnered about Cecilia's mathematical understanding. Then, they participated in a whole-class discussion about both questioning and Celicia's understanding, revisiting ideas discussed earlier in the course around strategies for adding. In the discussion, PSTs assessed Cecilia's conceptual understanding and areas of confusion based on their simulation experiences. They also created a list of what they struggled to elicit from Cecilia and completed a debrief assignment to reflect on their elicitation techniques.

Data Analysis

In this study, we conducted qualitative content analysis (Gläser-Zikuda et al., 2020) of three open-ended questions from the post-simulation survey to determine what the PSTs ascertained from their elicitation regarding Cecilia's problem-solving process and her conceptual understanding about strategies for adding within 20. To address RQ1, we analyzed PSTs' responses to a question asking them to describe the step-by-step procedure the GenAI student used to solve the addition word problem. Based on the design of the GenAI simulation and the pre-designed GenAI student's profile (Mikeska, Bhatia et al., 2025), there were three key aspects of Cecilia's problem-solving process: (a) drawing a pictorial representation (circles) to show the two groups of crayons, (b) counting on from the first addend (six) to find the sum, and (c) using dots on the last three circles to count on. As shown in Table 1, for each aspect of the student's profile, two researchers developed a codebook using a deductive coding approach (Chandra & Shang, 2019). The codebook outlined four different responses per aspect that could have been represented in the PSTs' written responses, from the most to the least accurate. The researchers engaged in group coding for all open-ended responses and applied four codes – one code per aspect – to each PST response, resolving any disagreements to finalize the codes. The researchers calculated descriptive frequencies and percentages to determine which aspects the PSTs accurately identified about Cecilia's problem-solving process.

Table 1
Key Aspects of the GenAI Student's Problem-solving Process and Example PST Responses

Aspect 1: Pictorial Representation of Two Groups	
Coding Category	**Example PST Survey Response**
DRAW – Two Groups: Mentioned drawing two groups (6 circles for Mike's crayons, 2 circles for Ann's crayons)	*She drew 6 circles for Mike, then drew 2 circles for Ann. After that, she counted on from 6.* (PST_009_Sim2)
DRAW – Picture (no groups): Mentioned drawing a picture but did not specify two separate groups were drawn	*Drew 8 circles to represent crayons then counted on to get the answer.* (PST_003_Sim1)
DRAW – Inaccurate: Mentioned drawing a picture but description does not align with the word problem the GenAI student solved	*She drew 6 circles for Mike and 3 circles for Ann.* (PST_004_Sim1)
DRAW – Not Mentioned: No mention of drawing a picture	*The student used counting on as her strategy. I asked if she could solve it in another way and she said no.* (PST_008_Sim1)
Aspect 2: Counting on to Find the Sum	
Coding Category	**Example PST Survey Response**
SUM – Counted On: Mentioned counting on from 6 to find the sum	*She drew 6 circles then 2 for Ann and counted on saying 6, 7, 8.* (PST_008_Sim2)
SUM – Counting/Adding: Mentioned counting or adding to find the sum but does not specify that they counted on from 6	*Drew 6 circles then 2 then added together.* (PST_003_Sim2)
SUM – Inaccurate: Mentioned the strategy inaccurately (e.g., counts all)	*The student drew 6 circles for Mike and 2 circles for Ann. She counted all the circles to add them.* (PST_004_Sim2)
SUM – Not Mentioned: No mention of finding the sum	*The student told me she drew six circles for Mike and then drew two more circles for Ann.* (PST_007_Sim1)
Aspect 3: Use of Dots to Count On	
Coding Category	**Example PST Survey Response**
DOTS – Accurate and Complete: Mentioned how the dots in the picture were used to count on to find the sum	*She started drawing the dots when she started counting on from Mike's crayons to Ann's crayons. She told me that she knew that one dot represented the last of Mike's and the two others represented both of Ann's.* (PST_001_Sim2)
DOTS – Accurate and Unclear: Mentioned the dots in the picture but does not explain how they were used to count on to find the sum	*There are 3 circles with a dot inside each of them. I was confused about this and tried to question the student about why they did it, and they had the same response that did not answer the question.* (PST_011_Sim1)
DOTS – Inaccurate: Mentioned the dots in the picture but states that they were not used correctly by the student	No examples
DOTS – Not Mentioned: No mention of the dots in the picture	*She drew 6 circles for Mike, then drew 2 circles for Ann. After that, she counted on from 6.* (PST_009_Sim2)

Note: PST_000_Sim is the format of PSTs' identification codes. This code includes the PST number and whether the example comes from their first (Sim1) or second (Sim2) simulation.

Two survey questions asked the PSTs to describe what the GenAI student understands and does not understand about strategies for adding, constituting the data to answer RQ2. The coding process for RQ2 followed the same process as RQ1; two researchers coded the PSTs' written responses to determine whether the PSTs could identify these aspects of the GenAI student's conceptual understanding and gaps in her understanding. The researchers then calculated the frequencies and percentages for each aspect across the PSTs' responses.

Study Findings

RQ1: PSTs' Understanding of the GenAI Student's Problem-Solving Process

As shown in Table 2, most PSTs (80%) were able to elicit and describe the GenAI student's strategy for the first step in her process, which involved drawing two groups of circles to represent the crayons that Mike and Ann had in the addition problem. Over half of the PSTs mentioned that Cecilia used the counting on strategy to determine how many crayons in all. Only about a third of PSTs were able to elicit and describe how Cecilia used the three dots to illustrate the counting on strategy. Findings show that it was rare for PSTs to inaccurately describe one of the key aspects of Cecilia's problem-solving process. In addition, some PSTs did not fully notice, elicit, and describe key aspects, particularly in the later steps of Cecilia's problem-solving process.

Table 2

Distribution of PSTs' Descriptions of the GenAI Student's Problem-Solving Process

Key Aspects of the GenAI Student's Problem-Solving Process	Timepoint One (n=10 responses)		Timepoint Two (n=10 responses)		Total (n=20 responses)	
	n	%	n	%	N	%
DRAW - Two Groups	6	60%	10	100%	16	80%
DRAW - Picture (no groups)	2	20%	0	0%	2	10%
DRAW – Inaccurate	1	10%	0	0%	1	5%
DRAW - Not Mentioned	1	10%	0	0%	1	5%
SUM - Counted On	7	70%	6	60%	13	65%
SUM - Counting/Adding	2	20%	2	20%	4	20%
SUM – Inaccurate	0	0%	2	20%	2	10%
SUM - Not Mentioned	1	10%	0	0%	1	5%
DOTS - Accurate and Complete	3	30%	4	40%	7	35%
DOTS - Accurate and Unclear	2	20%	0	0%	2	10%
DOTS – Inaccurate	0	0%	0	0%	0	0%
DOTS - Not Mentioned	5	50%	6	60%	11	55%

RQ2: PSTs' Articulation of the GenAI Student's Conceptual Understanding

As shown in Table 3, results indicate that PSTs struggled to elicit and identify key aspects of Cecila's understanding, especially any knowledge gaps, in this topic area. Less than half of PSTs' responses (45%) noted that Cecilia understood how to count on from one number. Similarly, only a few PST responses acknowledged that Cecilia understood the meaning of the numbers in the addition word problem or how to use models to solve an addition word problem. A few PSTs (35%) identified one knowledge gap: Cecilia did not know how to use other strategies (besides drawing and counting on) to solve the addition problem. Although the TE conducted a reflection and debriefing after the first time point, many PSTs were unable to identify the GenAI student's understandings and misunderstandings related to the problem during the simulation at both time points. The findings suggest that PSTs may need more targeted support in learning how to elicit students' thinking more effectively and in determining how to interpret first-grade students' mathematical ideas, especially regarding gaps in students' understanding.

Table 3
Distribution of PSTs' Articulation of the GenAI Student's Conceptual Understanding

Aspects of the GenAI Student's Conceptual Understanding		Timepoint One (n=10 responses)		Timepoint Two (n=10 responses)		Total (n=20 responses)	
		Yes n (%)	No n (%)	Yes n (%)	No n (%)	Yes n (%)	No n (%)
Indicates that the student understands	The meaning of the numbers in the addition problem	1 (10%)	9 (90%)	3 (30%)	7 (70%)	4 (20%)	16 (80%)
	How to count on from one number	5 (50%)	5 (50%)	4 (40%)	6 (60%)	9 (45%)	11 (55%)
	The use of models (e.g., drawings) to represent or solve the problem	2 (20%)	8 (80%)	1 (10%)	9 (90%)	3 (15%)	17 (85%)
Indicates that the student does not understand	How to fluently add the numbers (e.g., mental math)	0 (0%)	10 (100%)	0 (0%)	10 (100%)	0 (0%)	20 (100%)
	That addition is commutative	0 (0%)	10 (100%)	0 (0%)	10 (100%)	0 (0%)	20 (100%)
	That counting on from the larger addend (regardless of the position) is more efficient	0 (0%)	10 (100%)	0 (0%)	10 (100%)	0 (0%)	20 (100%)
	How to use other strategies to solve the problem	3 (30%)	7 (70%)	4 (40%)	6 (60%)	7 (35%)	13 (65%)

Discussion and Conclusion

In this study, we analyzed the open-ended survey responses of PSTs enrolled in an elementary mathematics methods course who engaged in a GenAI digital teaching simulation aimed at eliciting a student's thinking. The results indicated that the PSTs were able to identify some of the GenAI student's problem-solving process, such as drawing two groups or using the counting on strategy when solving an addition word problem but struggled to consistently identify all key aspects. Similarly, while some PSTs were able to identify a few aspects of the GenAI student's understanding, most of them struggled to use their elicitation to identify the GenAI student's knowledge gaps. Findings similar to the study by Orr and Bieda (2020) suggest that PSTs may need support in learning how to probe student thinking in more depth to understand nuances in how students use and apply problem-solving processes, moving beyond more surface-level elicitation. Doing so may allow PSTs to support their students as their concept of number develops.

Collectively, these results suggest several implications for using GenAI teaching simulations within mathematics teacher education courses and for supporting PST learning of how to elicit student thinking. Using GenAI teaching simulations can provide TEs with insight into how PSTs use their elicitation strategies to understand student thinking. This insight can allow TEs to identify growth areas for where to focus their future instruction with PSTs (Hallman-Thrasher, 2017). Guidance from TEs can support PSTs in identifying gaps in their elicitation strategies and in improving their ability to engage in complex instructional practices. In addition, similar guidance can help PSTs apply what they are learning with respect to the content in the context of student thinking.

In the current study, we primarily focused on how PSTs can benefit from using GenAI digital teaching simulations to practice specific skills, like eliciting students' thinking. However, this study did not assess PST learning outcomes, a gap we leave to future studies. In addition, the small sample size limited our ability to comprehensively examine the factors influencing how PSTs may have benefited from this simulation. Furthermore, the time interval between the two simulations was insufficient to measure any significant changes that might occur over time, which also provides an opportunity for future research.

In conclusion, PSTs need to understand and accurately identify the strategies students use to solve mathematical problems. Hence, PSTs require practice in learning how to elicit students' thinking effectively. Beyond eliciting students' strategies, PSTs need to learn how to assess the extent of students' understanding of a topic using purposeful probing questions. For that, PSTs may benefit from repeated practice, which GenAI teaching simulations can offer, to ensure their success in using this core instructional practice when they enter real classrooms.

Acknowledgements

We gratefully acknowledge the support of ETS's Research Institute and K–16 Strategic Business Unit, whose funding and collaboration made the development and refinement of the generative AI teaching simulation possible. ETS's K-16 Strategic Business Unit incorporated a version of this simulation into an online learning module designed to support elementary paraeducators in learning how to engage students with subject matter, a high-leverage teaching practice critical for educators' success in the classroom. We also extend our sincere appreciation to the preservice teachers who participated in this study.

References

Bondie, R., Mancenido, Z., & Dede, C. (2021). Interaction principles for digital puppeteering to promote teacher learning. *Journal of Research on Technology in Education*, *53*(1), 107–123. https://doi.org/10.1080/15391523.2020.1823284

Bywater, J. P., Lilly, S., & Chiu, J. L. (2025, July). Labeling disagreements: Illuminating the classification of mathematics teacher questions. In A. I. Cristea, E. Walker, Y. Lu, O. C. Santos, & S. Isotani (Eds.), *Proceedings of the International Conference on Artificial Intelligence in Education* (pp. 49–59). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-98417-4

Chandra, Y., & Shang, L. (2019). Inductive coding. In *Qualitative research using R: A systematic approach* (pp. 91–106). Springer, Singapore. https://doi.org/10.1007/978-981-13-3170-1_8

Gläser-Zikuda, M., Hagenauer, G., & Stephan, M. (2020). The potential of qualitative content analysis for empirical educational research. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, *21*(1), Article 17. https://doi.org/10.17169/fqs-21.1.3443

Hallman-Thrasher, A. (2017). Prospective elementary teachers' responses to unanticipated incorrect solutions to problem-solving tasks. *Journal of Mathematics Teacher Education, 20*(6), 519–555. https://doi.org/10.1007/s10857-015-9330-y

Hallman-Thrasher, A., Rhodes, G., & Schultz, K. (2020). Supporting mathematics teacher educators' practices for facilitating prospective teachers' mathematical explanations in content courses. *The Mathematics Enthusiast, 17*(2), 883–906. https://doi.org/10.54870/1551-3440.1506

Krupa, E. E., Huey, M., Lesseig, K., Casey, S., & Monson, D. (2017). Investigating secondary prospective teachers' noticing of student thinking. In E. O. Schack, J. Wilhelm, & M. H. Fisher (Eds.), *Teacher noticing: Bridging and broadening perspectives, contexts, and frameworks research in mathematics education* (pp. 49–72). Springer. https://doi.org/10.1007/978-3-319-46753-5_4

Law, L. (2024). Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. *Computers and Education Open, 6*. https://doi.org/10.1016/j.caeo.2024.100174

Mikeska, J. N., Beigman Klebanov, B., Bhatia, A., Halder, S., & Suhan, M. (2025, July). Evaluating the use of generative artificial intelligence to support learning opportunities for teachers to practice engaging in key instructional skills. In A. I. Cristea, E. Walker, Y. Lu, O. C. Santos, & S. Isotani (Eds.), *Proceedings of the International Conference on Artificial Intelligence in Education* (pp. 378–391). Cham: Springer Nature Switzerland. https://link.springer.com/content/pdf/10.1007/978-3-031-98417-4_27

Mikeska, J. N., Bhatia, A., Halder, S., Beigman Klebanov, B., Maxwell, T., Longwill, B., Behl, K., & Shekell, C. (2025, October). Generative AI teaching simulations as formative assessment tools within preservice teacher preparation. In *Proceedings of the Artificial Intelligence in Measurement and Education Conference (AIME-Con): Full Papers* (pp. 212–220). https://aclanthology.org/2025.aimecon-main.23.pdf

Mikeska, J. N., Howell, H., & Kinsey, D. (2023). Do simulated teaching experiences impact elementary preservice teachers' ability to facilitate argumentation-focused discussions in mathematics and science? *Journal of Teacher Education*, *74*(5), 422–436. https://doi.org/10.1177/00224871221142842

Molinari, L., & Mameli, C. (2013). Process quality of classroom discourse: Pupil participation and learning opportunities. *International Journal of Educational Research*, *62*, 249–258. https://doi.org/10.1016/j.ijer.2013.05.003

Monson, D., Krupa, E., Lesseig, K., & Casey, S. (2020). Developing secondary prospective teachers' ability to respond to student work. *Journal of Mathematics Teacher Education, 23*(2), 209–232. https://doi.org/10.1007/s10857-018-9420-8

National Council for Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all*. Reston, VA.

Orr, S., & Bieda, K. (2020). Preparing to elicit student thinking: Supporting PST questioning in an university teaching experience. In A. I. Sacristán, J. C. Cortés-Zavala, & P. M. Ruiz-Arias, (Eds.). *Mathematics Education Across Cultures: Proceedings of the 42nd Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education, Mexico* (pp. 1672–1676). https://doi.org/10.51272/pmena.42.2020-264

Shaughnessy, M., & Boerst, T. A. (2018). Uncovering the skills that preservice teachers bring to teacher education: The practice of eliciting a student's thinking. *Journal of Teacher Education, 69*(1), 40–55. https://doi.org/10.1177/0022487117702574

TeachingWorks Resource Library. (2019, August). Eliciting and interpreting. https://library.teachingworks.org/curriculum-resources/teaching-practices/eliciting-and-interpreting/

Tyminski, A. M., Simpson, A. J., Land, T. J., Drake, C., & Dede, E. (2021). Prospective elementary mathematics teachers' noticing of childrens' mathematics: A focus on extending moves. *Journal of Mathematics Teacher Education, 24*(6), 533–561. https://doi.org/10.1007/s10857-020-09472-2

Wang, X., Thompson, M., Yang, K., Roy, D., Koedinger, K. R., Rose, C. P., & Reich, J. (2021). Practice-based teacher questioning strategy training with ELK: A role-playing simulation for eliciting learner knowledge. *Proceedings of the ACM on Human-Computer Interaction,* 5(CSCW1), 1–27. https://doi.org/10.1145/3449125