2. Using Digital Clinical Simulations and Authoring Tools to Support Teachers in Eliciting Learners' Mathematics Knowledge

NANCY ANDERSON, RACHEL SLAMA, ROYA MOUSSAPOUR, GREG BENOIT, AND JUSTIN REICH

Abstract: Mathematics instruction that builds on student thinking has been linked to rich math learning environments and gains in student math achievement, yet it is challenging for teachers to find professional learning opportunities to develop these skills. Open-source digital clinical simulations with authoring features hold promise for supporting math teachers' professional learning at scale, particularly around skills that can be developed with low-stakes, repeated practice before classroom implementation. We engage a cohort of 19 middle-years math teachers in a midsized urban district in a program to regularly play and design simulations focused on eliciting learner math knowledge. Analysis of interview and simulation data revealed that teacher performance in the simulation related to their self-reported description of math discursive practice. At baseline, most participants did not report attending to the work of deepening students' reasoning as part of math instruction. Similarly, in the simulation, participants typically did not ask follow-up questions and instead engaged with their students in ways aligned with initiate-respond-evaluate conversation patterns. This simulation may be a promising diagnostic tool to help teachers implement classroom discussions to promote effective discursive practices. The paper concludes by discussing design implications for digital clinical simulations that support teachers' use of productive math talk.

Introduction

Mathematics instruction that builds on student thinking has been broadly linked to rich math instructional environments (e.g., Clarke, 2008) and gains in student math achievement (e.g., Bobis et al., 2005; Jacobs, Franke, Carpenter, Levi, & Battey, 2007). To elicit student thinking, teachers ask questions or assign tasks in which students share their thinking about specific academic content, allowing the teacher to evaluate student understanding, guide instructional moves, and surface ideas that may benefit the whole class (Teaching Works, 2018). Teachers require ample practice to develop the requisite skills and professional judgment to support deep math learning (Ball & Forzani, 2009; United States Department of Education, 2016), but they often lack opportunities to practice critical competencies in low-stakes settings (Grossman et al., 2009; Levine, 2009).

We describe an open-source, chat-based digital clinical simulation called *Eliciting Learner Knowledge* (*ELK*) designed by our team of learning scientists and engineers to allow teachers to practice eliciting learner math knowledge in a low-stakes setting. The purpose of this chat-based simulation is to help teachers develop effective questioning strategies for understanding student conceptions and misconceptions that they can later integrate into the complex assemblage of their classroom practice. One player role-plays the teacher based on a predetermined teacher profile. The other player role-plays the student. The players engage in a seven-minute chat conversation about a specific student understanding. The objective of the game is for the teacher to identify what the student knows and does not know about a topic. A short true/false quiz at the end of the game where the "teacher" must respond as if she or he is the student tests the teacher's knowledge of the student's understanding. *ELK* also has an authoring tool component where users can quickly generate their own grade-level specific scenarios. We depict a composite chat transcript that results from the "student"-"teacher" interaction (see Figure 1).

3:30:00: Teacher: Hello Joel. Can you tell me how you solved 8 divided by ¼?
3:30:20: Student: Yes, I know that there are four fourths in one whole. So times 8 is 32.
3:30:53: Teacher: So I heard you say that if there are four fourths in one whole and since we have 8 wholes, that's 32. Is that what you said?
3:30:58: Student: Yes.
3:31:21: Teacher: Okay, tell me a bit more. Why did you think about the number of fourths in one whole?
3:31:39: Student: Because there are four fourths in one whole.
3:31:57: Teacher: Yes, you are right. But what is it about the expression 8 divided by ¼ that told you to count fourths?
3:32:14: Student: Well, it's division, so that means how many times ¼ goes into 8.
3:33:29: Teacher: I see. But let me check that I understand. You are thinking about division as how many times we can take out or repeatedly subtract groups. So to find the number of times we can take ¼ from 8, you first found how many times to take it out of one? Is that what you're thinking?
3:33:34: Student: Yes.

Figure 1. Eliciting Learner Knowledge composite chat transcript.

Through a yearlong fellowship program with 19 math teachers teaching Grades 3–9 in a midsized urban school district, we study the effects of playing and co-designing scenarios within the simulation on learners' perception of the role of math discussion in the classroom and change in teacher practice. We discuss implications for the design and scale of digital clinical simulations.

Research Questions

In this paper, we explore the following research questions:

1. How do teachers describe their discursive math practices in interviews?
2. How does teacher performance in a simulation relate to the descriptions of their practice?
3. To what extent do teachers deploy new discursive math practices in simulations?

Background and Context

Eliciting Learner Knowledge and Student Math Learning

To effectively elicit a math learner's knowledge, a teacher must draw out a student's thinking through carefully chosen questions and tasks that consider and check alternative interpretations of the student's ideas and methods. Effective teachers build on students' preconceptions and misconceptions about a topic (e.g., Fuson, Kalchman, & Bransford, 2005) that can provide a bridge to new material and engage the individual learner, surfacing mathematical ideas that support mathematical thinking for the whole class (e.g., Van Zoest et al., 2017). However, to drive gains in student math achievement, teachers must move beyond eliciting learner knowledge and attending to students' strategies to interpreting their understandings and deciding how to respond (e.g., Jacobs, Lamb, & Philipp, 2010). Teaching skills related to attending to and interpreting student thinking are difficult to learn (Shaughnessy & Boerst, 2018) and have proved challenging to assess among teachers. These skills occur interactively in the classroom (Jacobs, Lamb, & Philipp, 2010); student responses to teachers' questions are difficult to predict (e.g., Shaughnessy & Boerst, 2018) and student responses may be context-dependent (Sleep & Boerst, 2012). Teaching is complex, and teachers need opportunities to engage in targeted, low-stakes, improvisational rehearsal of these practices (Grossman et al., 2009; Levine, 2009).

Using Digital Simulations to Practice Eliciting Learner Math Knowledge

ELK leverages games, technology, and playfulness to help teachers grow their teaching practice in math. We leverage widely accessible, mobile, web-based technology to help teachers to practice everyday interactions between teachers and students that can shape K–12 students' academic trajectories (Reich, Kim, Robinson, Roy, & Thompson, 2018, 2019). Most opportunities for teachers to practice have focused on replicating the full complexity of teaching as closely as possible through mock teaching or digital simulations, whereas *ELK* tries to create more targeted, low-stakes opportunities to practice specific dimensions of teaching that can be systematically improved and reintegrated into the teacher's practice. Participants role-play as students with conceptual misunderstandings and teachers try to elicit student thinking. This activity aims to support the teacher to engage naturally with productive practice (Thompson et al., 2017).

Measuring Teacher-Student Math Discourse

Classroom talk typically adheres to a triadic structure wherein the teacher initiates an exchange, the student responds, and the teacher responds to what the student said. In traditional instruction, the triadic structure is typically of the form "initiate-response-evaluate" (IRE) in which the teacher offers feedback such as an evaluation or correction in the "third slot." By contrast, in discourse-intensive instruction, the triadic structure is in the form "initiate-response-follow-up" (or IRF; Cazden, 2001; Nassaji & Wells, 2000; Wells, 1996) in which the teacher responds with follow-up questions that seek additional information about what the student said.

Chapin, O'Connor, and Anderson (2013) describe types of questions or "talk moves" that help teachers and other students examine the intended meaning, mathematical merit, or relevance of a student contribution. *Revoicing* consists of the teacher's repeating the student's prior turn while also offering the student the chance to verify or clarify the teacher's interpretation of his or her intended meaning (O'Connor & Michaels, 1993, 1996). *General press for reasoning* refers to any prompt that encourages the student to keep talking about his or her ideas. *Specific press for reasoning* questions aim to develop students' thinking by building awareness for the fundamental ideas at the heart of their explanation, connecting individual aspects of their explanation in more robust ways, and situating their explanation within the relevant mathematical concepts under inspection (Truxaw & DeFranco, 2008). Prior research suggests that it typically takes more than two specific and consecutive follow-up questions to increase the accuracy or robustness of a student's original explanation (Franke et al., 2009).

Instructional Design and Simulation Co-Design Process

As part of a yearlong fellowship, we met with fellows one evening per month for a four-hour in-person block of time between October 2019 and January 2020. During each four-hour meeting, fellows participated in the following types of activities: (a) review of instructional strategies, (b) playing *ELK*, (c) debriefing and reflecting on their performance in the simulation, and (d) designing their own *ELK* scenarios based on a series of low-fidelity prototype templates. Fellows also received biweekly coaching sessions at their school sites focused on the instructional strategies presented at the meetings.

Methods

We employed a design-based implementation research approach (DBIR; e.g., Collins, Joseph, & Bielaczyc, 2004), which allowed for rapid-cycle testing of our digital clinical simulations. Our mixed-methods approach included analysis of interview data and simulation chat transcripts.

Research Site and Participants

We recruited a cohort of 19 math teachers to participate in our yearlong fellowship through a widely circulated open call for participation. All fellows teach in a midsize, urban school district in a middle-years math classroom (defined as Grades 3 through 9). They teach at 16 different schools and in mainstream, inclusion, and sub-separate classroom environments. Nearly all of the fellows' schools (94%) have a student population with a majority of students who are Black or Latino. Most fellows (81%) teach at schools where more than half of the student body (51%) experience poverty. Fellows have between 2 and 30 years of teaching experience. Of the 19 fellows, 47.4% self-report as White, 10.5% as Asian, 21.0% as Black or African American, 5.3% as Hispanic/Latino, 5.3% as Other, and 10.5% did not report their race.

Data Collection and Analysis

We report on two data sources—interviews and simulation chat transcripts—to capture fellows' perception of their discursive math practice (Research Question 1), the link to their performance in the simulation (Research Question 2), and the degree to which they deploy new discursive practice in the simulation (Research Question 3). Below, we describe each of the data sources.

Fellow interviews. In September and October 2019, we conducted a set of semistructured interviews with fellows (*N* = 16) that asked participants a range of demographic and contextual information about professional learning at their school and how they facilitate discourse in the math classroom. Using an established framework defining four goals of the math teacher's role in facilitating classroom discourse (Chapin et al., 2013), we double coded a subset of questions according to these *etic* codes: (a) *share*: helping students share their reasoning; (b) *deepen*: helping students deepen their reasoning; (c) *attend*: helping students attend to the reasoning of others; and (d) *engage*: helping students engage with the reasoning of others. The percent agreement between the two raters was 60%. All disagreements were resolved by discussion between the raters.

Simulation chat transcripts. We collected and coded fellows' transcripts from the *ELK* platform in October, November, December, and January. Two researchers coded each of the transcripts for the type of exchange (*independent*, *dependent*, *other*, *not applicable*), as well as the teacher turn type (*revoicing*, *general press for reasoning*, *specific press for reasoning*, and *other*). Coders also marked the number and length of sequences—or chains—of extended student-teacher exchanges because of the prior link established between this form of classroom talk and student learning (Franke et al, 2009). We calculated interrater reliability using Cohen's kappa for each exchange type (independent and dependent). Interrater reliability for Rater 1 and 2 was moderate for independent—Kappa = 71% ($p < 0.001$), 95% confidence interval, or CI (0.63, 0.78)—and dependent exchanges—Kappa = 70% ($p < 0.001$), 95% CI (0.62, 0.79) (McHugh, 2012). All disagreements were resolved by discussion between the two raters.

Results

In Research Question 1, we examined how teachers described their discursive practice before participating in the professional learning. We identified examples of attending to the goal of helping students *share* their reasoning in all participant interviews. Sixty-nine percent of participants reported helping students *attend* to the reasoning of others and 63% of participants reported helping students *engage* with the reasoning of others. Only 19% of participants reported helping students *deepen* their own reasoning. Examples of helping students share their reasoning included talk-based routines or activities; these included number talks, estimation puzzles, and "do now" tasks. When describing the student talk that occurs during these activities, participants focused on student explanation or justification. Strategies for helping students attend to the reasoning of others included supporting students in talking to each other in order to reach consensus on a solution to a problem, students examining and questioning each other about their problem-solution strategies or computational methods, and standing before the class and explaining a strategy before soliciting peer feedback. Engage-facilitation strategies focused on students' using each other's ideas to develop a full explanation as well as examining mistakes and correcting faulty solution strategies. Only three participants described using teacher questioning as a strategy for supporting students who struggle to explicate their thinking or to deepen student reasoning.

In Research Question 2, we sought to understand how teacher performance in a simulation relates to descriptions of their own practice. Based on the *ELK* chat transcripts, we summarized teachers' discursive practice by calculating the mean length and standard deviation for the total number of conversational chains per month. The average chain length varied from 1.25 to 1.45 exchanges across the four data-collection points (see Table 1). These data indicate that discursive chains were typically only one exchange in length. When playing the role of the teacher in the chat, participants most often initiated an exchange and followed the student response with a new question that was unrelated to the mathematics ideas within the student's prior response. These findings suggest that teachers' performance in a simulated environment are aligned with how they describe their classroom discussion practices. Only three participants mentioned teacher questioning strategies as an aspect of this work. These data suggest that participants provide students with opportunities to engage in discussion on a regular basis but do not attend to the teacher's role in asking follow-up questions that aim to deepen students' reasoning.

Month	Length of chain	
	Mean	SD
October	1.38	0.81
November	1.33	0.68
December	1.45	0.97
January	1.25	0.56

Table 1. Mean conversational chain length.

In Research Question 3, we sought to understand the extent to which teachers deploy new discursive practices in simulations. We coded chains by length to see if there was change through time in the frequency with which participants engaged in interactions with "students" in the simulation. The frequency of chains made up of one exchange remained consistent throughout the intervention (see Table 2). We also looked at the frequency of follow-up questions by type that participants asked during each month of the intervention. Our results revealed no clear change in frequency or prevalence of any of the three types of follow-up questions (revoicing, general press for reasoning, or specific press for reasoning).

Chain Length	October	November	December	January	Total, October - January	Total, November - January
1	77.4	73.1	75.5	79.1	76.3	75.9
2	13.2	22.4	11.3	17.9	16.7	17.7
3	3.8	3.0	9.4	1.5	4.2	4.3
4	5.7	1.5	1.9	1.5	2.5	1.6
5+	0.0	0.0	1.9	0.0	0.4	0.5

Table 2. Percent of chains of length 1, 2, 3, 4, or 5+.

Discussion

Consistent with data on teachers' self-described practice, our simulation data suggest that participants' chats aligned with direct-instruction conversational formats. On average, chains remained shorter than two exchanges across each *ELK* administration. The teacher player in a chat typically asked the student player a question and awaited a response. After the student response, the teacher then asked another question that did not seek to follow up on the student's ideas in his or her prior contribution. Instead, the teacher asked a new question that was unrelated to the student's prior contribution and sought to redirect the student toward alternate solution-strategy pathways.

Future simulation design and research considerations include supporting teachers in unpacking the relevant mathematics, identifying common student misconceptions related to that content, and analyzing student work. In order for the teacher to "poke at" certain key ideas or identify the misconception, the teacher player must know the relevant mathematics deeply. Materials that help teachers practice and role-play in a variety of ways might also support effective discursive practice in the simulation. This may include modeling teacher-student scenarios, highlighting model *ELK* questioning practice, and analyzing how other teachers respond in *ELK*. This study also highlights the promise of *ELK* as a diagnostic tool. Participants' self-reported discursive practice was closely aligned with their performance in the simulation, in which students had very few opportunities to expand on their original explanation, illuminate their correct concepts and misconception, or connect new understanding to prior knowledge.

Conclusion

Digital clinical simulations with authoring features hold promise for supporting math teachers' professional learning, particularly around skills that benefit from low-stakes, repeated practice before classroom implementation such as facilitating a discussion. *ELK* shows particular promise as a diagnostic tool that reflected teachers' understanding of their role in promoting math discourse. Future iterations of the simulation will explore the integration of supports directly into the tool as an intervention to promote effective discursive practice.

References

Ball, D. L., & Forzani, F. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education*, 60(5), 497–511.

Bobis, J., Clarke, B., Clarke, D., Thomas, G., Wright, R., Young-Loveridge, J., & Gould, P. (2005). Supporting teachers in the development of young children's mathematical thinking: Three large-scale cases. *Mathematics Education Research Journal*, 16(3), 27–57.

Cazden, C. (2001). *Classroom discourse* (2nd ed). Portsmouth, NH: Heinemann.

Chapin, S., O'Connor, M., & Anderson, N. (2013). *Classroom discussions: Using math talk to help students learn* (3rd ed.). Sausalito, CA: Math Solutions.

Clarke, B. (2008). A framework of growth points as a powerful teacher development tool. In D. Tirosh & T. Wood (Eds.), *International handbook of mathematics teacher education: Vol. 2. Tools and processes in mathematics teacher education* (pp. 235–256). Rotterdam, The Netherlands: Sense.

Collins, C., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *Journal of the Learning Sciences*, 13(1), 15–42.

Franke, M. L., Webb, N., Chan, A. G., Ing, M., Freund, D., & Battey, D. (2009). Teacher questioning to elicit students' mathematical thinking in elementary school classrooms. *Journal of Teacher Education*, 60, 380–392.

Fuson, K. C., Kalchman, M., & Bransford, J. D. (2005). Mathematical understanding: An introduction. In M. S. Donovan & J. Bransford (Eds.), *How students learn mathematics in the classroom* (pp. 217–256). Washington, DC: National Research Council.

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record*, *111*(9), 2055–2100.

Jacobs, V. R., Franke, M. L., Carpenter, T. P., Levi, L., & Battey, D. (2007). Professional development focused on children's algebraic reasoning in elementary school. *Journal for Research in Mathematics Education*, 38, 258–288.

Jacobs, V. R., Lamb, L. L. C., & Phillip, R. A. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education*, 41, 169–202.

Levine, A. (2009). *Educating school teachers*. Princeton, NJ: The Education Schools Project. Retrieved from http://edschools.org/pdf/Educating_Teachers_Report.pdf

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), 276–282.

Nassaji, H., & Wells, G. (2000). What's the use of 'triadic dialogue?' An investigation of teacher-student interaction. *Applied Linguistics*, 21, 376–406.

O'Connor, M. C., & Michaels, S. (1996). Shifting participant frameworks: Orchestrating thinking practices in group discussion. In D. Hicks (Ed.), *Child discourse and social learning* (pp. 63–102). Cambridge, UK, and New York, NY: Cambridge University Press.

O'Connor, M. C., & Michaels, S. (1993). Aligning academic task and participation status through revoicing: Analysis of a classroom discourse strategy. *Anthropology & Education Quarterly*, 24(4), 318–335.

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.

Reich, J., Kim, Y. J., Robinson, K., Roy, D., & Thompson, M. (2018). Teacher practice spaces: Examples and design considerations. In *Proceedings of the 2018 International Conference of Learning Sciences*, London, UK. doi:10.31235/osf.io/ewn2g

Reich, J., Kim, Y. J., Robinson, K., Roy, D., & Thompson, M. (2019). Exploring authenticity and playfulness in teacher practice spaces. In J. H. Kalir (Ed.), *Proceedings of the 2018 Connected Learning Summit* (pp. 214–223). Pittsburgh, PA: ETC Press.

Sleep, L., & Boerst, T. A. (2012). Preparing beginning teachers to elicit and interpret students' mathematical thinking. *Teaching and Teacher Education*, 28(7), 1038–1048.

TeachingWorks. (2018). *High-leverage practices*. Retrieved from http://www.teachingworks.org/work-of-teaching/high-leverage-practices

Thompson, M., Roy, D., Wong, T., Tan, P., Huttner-Loan, E., & Reich, J. (2017, March). *Online practice for preservice teachers to elicit students' preconceptions*. Presented at the 2017 Society for Information Technology and Teacher Education Conference, Austin, TX.

Truxaw, M. P., & DeFranco, T. C. (2008). Mapping mathematics classroom discourse and its implications for models of teaching. *Journal for Research in Mathematics Education*, 39, 489–525.

United States Department of Education, Office of Innovation and Improvement. (2016). *STEM 2026: A vision for innovation in STEM education*. Washington, DC: Author. Retrieved from https://innovation.ed.gov/files/2016/09/AIR-STEM2026_Report_2016.pdf

Van Zoest, L. R., Stockero, S. L, Leatham, K. R., Peterson, B. E., Atanga, N. A., & Ochieng, M. A. (2017). Attributes of instances of student mathematical thinking that are worth building on in whole-class discussion. *Mathematical Thinking and Learning*, 19(1), 33–54.

Wells, G. (1996). Using the tool-kit of discourse in the activity of learning and teaching. *Mind, Culture, and Activity*, 3, 74–101.

Acknowledgments

This work was made possible by the generous support of the Bill & Melinda Gates Foundation as part of the Middle Years Math initiative. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Gates Foundation or any other organizations with which the authors are affiliated.