Simulations as a Tool for Practicing Questioning

Corey Webel, Kimberly Conner and Wenmin Zhao

Abstract In this chapter we discuss some of the affordances and constraints of using online teaching simulations to support reflection on specific pedagogical actions. We share data from a research project in which we implemented multiple iterations of a set of simulated teaching experiences in an elementary mathematics methods course. In each experience, preservice teachers contrasted the consequences of different pedagogical choices in response to a particular example of student thinking. We share how their evaluations of their choices shifted within experiences at certain points, and their criteria for "good" questions began to evolve. We end with implications for how simulations can promote critical reflection on teaching practice.

Keywords Representations of practice · Teaching simulations · Questioning
Preservice teacher education · Elementary mathematics

Introduction

Representations of practice are being increasingly used to engage preservice teachers (PSTs) in problems of instruction (Amador et al. 2017; Bartell et al. 2013; de Araujo et al. 2015; Herbst et al. 2011; Sun and van Es 2015). These can include videos, animations, comic strips, vignettes, photos, and real or manufactured representations of student work. These representations have various affordances and limitations, but in general, they help PSTs and their instructors decompose instructional practice into manageable pieces that can be described, interpreted, analyzed, and practiced. Part of the purpose of this monograph is to explicate various ways that representations of practice can be used in teacher education to promote learning through, for example, stimuli for reflection, criteria-based analysis, or structured observation.

C. Webel (✉) · K. Conner · W. Zhao
University of Missouri, Columbia, MO, USA
e-mail: webelcm@missouri.edu

O. Buchbinder and S. Kuntze (eds.), *Mathematics Teachers Engaging with Representations of Practice*, ICME-13 Monographs,
https://doi.org/10.1007/978-3-319-70594-1_6

In this chapter, we focus on a particular kind of representation of teaching which we describe as a *teaching simulation*, created with tools provided by the Lesson*Sketch* online platform (www.lessonsketch.org). A simulation, as we use the term, is a representation of practice that provides an opportunity for pedagogical action, in addition to opportunities for activities like noticing student thinking and reflecting on teaching. We designed the simulations to gain insights about the decision-making processes of novice teachers as they make pedagogical choices. They allow us to set up opportunities for learning directly from teaching (Hiebert et al. 2007), as PSTs try different pedagogical actions and then reflect on their consequences.

In this paper, we reflect on some specific design considerations that have resulted from the first two years of implementation. We specifically targeted the questioning practices of the PSTs in our program, documenting their pedagogical choices as well as their rationales and reflections. We analyzed what PSTs noticed about the consequences of the questions they selected in specific pedagogical situations and documented changes in their choices and explanations within and across multiple experiences with the simulations.

Background

Questioning Practices

International comparisons in mathematics teaching have shown that low-level questions, which require students to recall specific facts or carry out certain procedures, are especially prevalent in the United States (Givvin et al. 2005; Kawanaka and Stigler 1999; Stigler et al. 1996). Similarly, sequences of closed questions, intended to direct students through a series of procedural steps until they obtain the correct answer, have been referred to as *funneling* (Herbel-Eisenmann and Breyfogle 2005; Wood 1998). These types of questions position students as recipients of information rather than contributors to their own knowledge development (Boaler 2003; Webb et al. 2006), and are unlikely to spur correct and complete explanations on the part of students (Franke et al. 2009).

Recent research has described questioning practices that, in contrast to funneling or recall questions, are responsive to student thinking, drawing out and building on the specifics of students' ideas rather than imposing the teacher's idea (Jacobs and Empson 2016; Kazemi and Stipek 2001; Sherin 2002). Building on this research, The National Council of Teachers of Mathematics' *Principles to Actions* (2014) advocated teacher questions that "build on, but do not take over or funnel, student thinking," and those that "make mathematical thinking visible" (p. 41). In this study, we are looking specifically at follow-up questions that teachers might pose immediately after eliciting an initial explanation about a student's solution. Based on the literature described above, we defined types of questions as "low leverage"

Table 1 Classifications of question types

Type	Description	Example
Low leverage (directive)	Suggests a specific alternate strategy, does not refer to student work	Do you know how to do find common denominators?
Low leverage (invalidate)	Specific to student's work, but invalidates the student's strategy	Is it really two whole brownies?
Low leverage (funnel)	Responds to student work, but funnels to a correct answer. Often includes a binary choice (either/or, yes/no, etc.)	Are those pieces [in your diagram] sixths, or tenths?
High leverage (elicit)	Elicits student's thinking	Can you tell me more about the sixths in your diagram?
High leverage (build)	Help students build on their own thinking	Based on your diagram, who would you say gets the most amount of brownie?

or "high leverage" as shown in Table 1. For example, the low leverage (funnel) example refers to a student's work, but presents a binary choice that funnels the student toward a correct answer. In contrast, the high leverage questions reference specific aspects of student work ("your diagram"), but either elicit more information from the student about the work (i.e., they make mathematical thinking visible), or push students to consider the meaning of their work without conveying that it is correct or incorrect (i.e., they provide opportunities for students to build on their own thinking).

Learning to Ask Better Questions

Teacher educators have used various approaches to help PSTs and practicing teachers improve their questioning (Milewski and Strickland 2016; Moyer and Milewicz 2002; Nicol 1999; Spangler and Hallman-Thrasher 2014; Wagner 1973). For example, Moyer and Milewicz (2002) introduced a questioning framework to support PSTs in recognizing questions with different features. PSTs who worked with the framework began to ask more follow-up questions, but inconsistently (e.g., in some cases, they did so only when students had incorrect answers). Spangler and Hallman-Thrasher (2014) used imaginary task dialogues to support PSTs' ability to anticipate and respond to student thinking. When PSTs enacted the tasks with real students, the researchers found that while PSTs were able to develop and use a repertoire of "standard" responses, such as "How did you get that?" and "Can you tell me what you were thinking?", they struggled to respond to students in ways that were task-specific. Nicol (1999) found that PSTs struggled to reconcile different purposes for questions, such as learning more about student thinking but also helping them arrive at a correct solution. All of this work shows that supporting PSTs in developing high leverage questioning practices is challenging, and that learning about frameworks or categories of questions does not always translate into the ability to respond to students with high leverage questions. One explanation is that knowledge is situated; that is, "how a person learns a particular set of knowledge and skills, and the situation in which a person learns, become a fundamental part of what is learned" (Borko et al. 2000, p. 195). This view of knowledge as situated implies that that if PSTs are going to draw on the knowledge and skills that they gain in their education courses, the context of their learning experience needs to feel like teaching. Such experiences could include approximations that represent some authentic aspects of practice but also provide low-risk opportunities for novices to try, fail, and learn from their practice (Grossman et al. 2009).

Representations of Practice

To approximate practice, one must first represent it. There are many ways to represent teaching practice, including vignettes, depictions of student work, photos, animations, comic-strips, and videos. These have various affordances, but in general they aid in the decomposition of practice and support reflection on specific pedagogical situations (e.g., Kuntze et al. 2015). Also important are the ways that learners are asked to engage with representations (Beilstein et al. 2017). Videos in particular have been shown to help PSTs analyze and attend to details about the work of teaching (Star and Strickland 2008; Sun and van Es 2015). For example, Sun and Van Es (2015) found that PSTs who took a video-based course attended to and took up student ideas better than the students who took the previous course that did not utilize videos. They concluded that "learning to systematically analyze teaching with video can help PSTs learn to enact practices that afford opportunities to access and examine student thinking" (p. 210).

We define simulations as representations of practice that provide the possibility of pedagogical action. When a PST engages in a simulation, they can engage in activities similar to those associated with other representations (noticing, interpreting, describing, and reflecting), but in addition, they can make choices that actually *affect* the representation. They can see the results of those choices, and can make judgments about those choices on the basis of their effects. While we have designed our simulations within Lesson*Sketch*, other types of simulated experiences have been employed in mathematics teacher education, such as the use of trained actors or peers playing the role of students (Baldinger et al. 2016; Lampert et al. 2013; Shaughnessy et al. 2015). These similarly put novices in the position of making choices that have consequences within the simulation, though such "rehearsals" require decisions to be made quickly and may not afford as much time for reflecting on specific decisions.

In this project, we used the Lesson*Sketch* platform to design online storyboard teaching scenarios, which include some aspects of the teaching context such as a classroom, students, student work, dialogue (represented by text bubbles), etc. Lesson*Sketch* provides tools to aid in the reflection process; moments where the situation is paused and the user can be asked to make a choice, provide a comment, or ask a question. Finally, Lesson*Sketch* includes a "media chooser" tool, in which the user can be asked to select one out of a number of representations (in our case, these represented possible teacher actions). Each choice represents a unique path, and the designer can establish in advance how the situation will unfold in response to particular choices made by the user. The use of this feature is what distinguishes our Lesson*Sketch* experiences as simulations (see Kosko 2016 for a similar use of Lesson*Sketch*).

In contrast to interviews with real students (e.g., Moyer and Milewicz 2002; Nicol 1999), or with peers playing the role of students (Baldinger et al. 2016), the Lesson*Sketch* tool allows the designer a high degree of control over what the user can see, do, and notice within the representation (Herbst et al. 2011). Because Lesson*Sketch* experiences are standardized, the quality of the experience is not dependent upon the expertise of facilitators, actors, or peers playing the part of students. This is both a strength of situating the simulation with Lesson*Sketch* (we can compare how different participants respond in the same instructional situation) as well as a limitation (it cannot respond as flexibly to individual differences, and only includes a limited number of choices). In addition, because our simulations are online, they can be accessed easily by many participants and can generate substantial data in a short amount of time. Tweaks to the design can be made with little effort and new iterations can be subsequently tested with new populations.

One of the goals of the simulations was to not only see what could be revealed about PSTs' questioning practices, but to see if the simulations might impact the way they reflected on questions, including their purposes for questions and whether they believed their selected questions were "good." In this paper, we discuss our findings related to the research question, "How might cartoon-based teaching simulations be used to challenge novice teachers' mathematics questioning practices?"

Methods

Participants

In the first year of implementation, we engaged PSTs (n = 53) in three simulations during their first of two elementary methods courses at a four-year university in the Midwestern region of the United States. After analyzing and reporting on this data (Webel and Conner 2015, 2017), we revised these simulations and administered them with a new population (n = 86) the following year. In both administrations, PSTs were generally in their junior (third) year of university study, approximately 20–21 years old. The first course in the sequence targeted fraction concepts for the first eight weeks of the semester, focused specifically on helping PSTs appreciate the role of the unit in constructing and naming fractions (Chval et al. 2013), while the second course focused on measurement and geometry. Course assignments included explorations of mathematics with an emphasis on justification and reasoning, as well as analyzing and interpreting student work. As part of the program, each PST was assigned a field placement in an elementary classroom in which they spent at least 60 hours over the course of each semester during their junior year.

Data Collected from the Simulations

Each simulation involved a mathematical task, a classroom scenario, and a representation of student work. PSTs completed each experience as a homework assignment; the three experiences were spaced out, with about three weeks between them. A map of one experience titled Brandon is provided in Fig. 1.

First, PSTs solved a mathematical task (Step 1) and were asked to describe the mathematical ideas addressed in the task (Step 2). They watched a classroom episode that culminated in the teacher asking Brandon to explain his work. Then the PST was presented with several prompts, including requests to interpret the thinking represented by Brandon's work (Step 3), compose a question for Brandon (Step 4), and then select a question from a pre-established list (Step 5) and provide a rationale for why they believed the selected question would be the best to ask Brandon (Step 6). The choices included both high and low leverage questions. In the example shown in Fig. 1, the high leverage question (Step 5.3) aimed at eliciting Brandon's thinking by focusing on the critical misconception in his solution without directing him down a particular path. One of the low leverage question directed Brandon to a specific (procedural) strategy (Step 5.1), and the other funneled Brandon to a yes or no answer and conveyed that his solution strategy was incorrect (Step 5.2).

After selecting one of these questions, the PSTs viewed a predesigned response from Brandon (Step 7) and then were asked to evaluate their question once more (Step 8). For the high leverage question in Fig. 1, Brandon's response showed explicitly that he was now thinking of the previously established fourths and sixths as tenths, and in doing so had also changed the referent whole from one cup to two cups. This response has potential to help PSTs see these misconceptions more clearly than in Brandon's original response, and also opens up possibilities for Brandon to recognize, on his own, the inconsistencies in his solution (for example, he drew all of the sixths in the bottom cup to be the same size, but does not recognize that not all of the "tenths" are the same size). In response to the low leverage question, "Are fourths the same as sixths?" Brandon's response gave little information about his thinking; he merely responded with the expected answer of "no." Rather than providing an opportunity for Brandon to recognize and confront his misconception, the teacher's question allowed the misconception to remain unexamined while simultaneously conveying that his solution was incorrect.

After viewing Brandon's response to their selected question, PSTs were asked to imagine they could "go back in time" to see what would have happened had they asked the other question (Step 9). They viewed the Brandon's new response, and

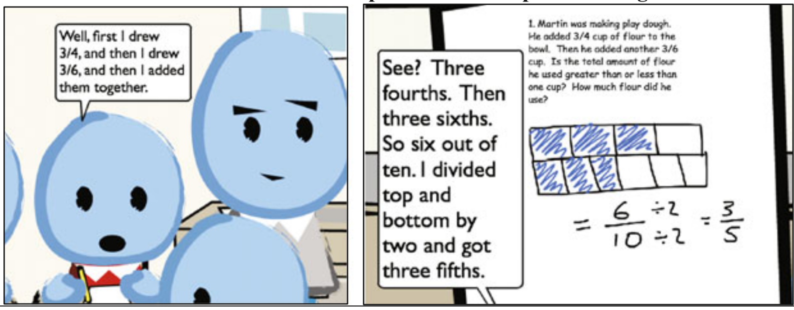
1. Solve a mathematical task

Martin was making play dough. He added $\frac{3}{4}$ cup of flour to the bowl. Then he added another $\frac{3}{6}$ cup. Is the total amount of flour he used greater than or less than one cup? How much flour did he use?

2. Unpack learning goals

What are the important mathematical ideas emphasized in this task? What misconceptions might this task reveal?

3. View classroom episode and interpret thinking



4. Pose a question to the student (and explain)

5. Select a question to ask the student

5.1 "Do you know what you need to do to the denominators before you can add fractions?" (low-directive)	**5.2** "In the problem it says that there are three fourths and three sixths. Are fourths the same as sixths?" (low-funnel)	**5.3** "Can you tell me more about where the fourths, the sixths and the tenths are in your picture?" (high-elicit)
6. Rationale for choice	**6. Rationale for choice**	**6. Rationale for choice**
7. View student response Brandon: Oh, multiply them by each other! Six times four is 24! So is it 24?	**7. View student response** Brandon: "Uh, no…" [with a confused expression] Teacher: "Remember, you can't add fractions unless they are the same size." Brandon: "Okay."	**7. View student response** Brandon: "These bigger pieces are fourths and these smaller ones are sixths. Then when I looked at the whole thing, there were ten pieces, so that was tenths"
8. Reflection about chosen question	**8. Reflection about chosen question**	**8. Reflection about chosen question**
9. See response for alternative question: "Can you tell me more about where the fourths, the sixths and the tenths are in your picture?" (high-elicit)		**9. See response for alternative question: "In the problem it says that there are three fourths and three sixths. Are fourths the same as sixths?" (low-funnel)**

10. Evaluate two questions

Looking back, how would you compare the second question with your first question? (and explain)

Fig. 1 A flowchart of the Brandon experience

then concluded the experience by determining which of the two questions they believed was "better" and explained why (Step 10). In some of the experiences, this was the final step. In the Brandon experience, we included another set of questions that could have been posed to Brandon, and repeated Steps 5–10 with this new set of questions (hereafter referred to as Brandon B. The set of questions shown above will be referred to as Brandon A).

Although we have data from six experiences in total (three from Year 1 and three from Year 2), only four of the experiences follow the format shown in Fig. 1, and these are the focus of this chapter. The two experiences that are excluded did not have Steps 9 and 10, in which PSTs compared the effects of different questions. Our previous analysis of data from Year 1 (as described in Webel and Conner 2017) suggested that Steps 9 and 10 were important for challenging PSTs' perspectives about effective questions. In this chapter, one of our main aims is to document the differences between how PSTs responded to the prompt in Step 8 versus the prompt in Step 10, and so we only include analysis from experiences that included both steps.

The experiences and question sets analyzed in this paper include the following:

- Matthew 2015: Matthew is depicted as questioning whether "3/4 of two brownies" can be an accurate way to describe a picture of two square brownies, with one whole brownie shaded and half of the second brownie shaded. He responds, "No, because they are cut in half, not in four squares." This is the only experience analyzed from the first iteration of the simulations in 2015.
- Matthew 2016: The same experience offered in 2015, but with a new group of PSTs.
- Brandon A: The experience described in Fig. 1.
- Brandon B: A second set of questions offered at the end of the Brandon experience. This set consisted of two questions: "What is the whole?" and "Where is the cup of flour in your picture?"
- Cedric: A new experience involving the task, "If I have four square yards, how many square feet is that?" Cedric draws a picture of a 4 by 3 rectangle, multiplies 4 by 3, and gives the answer of 12 ft^2.

In each experience, the student (Matthew, Brandon, or Cedric) produced work that revealed a significant mathematical misconception that had been previously discussed with the PSTs in the methods class. In none of the responses to questions did any of the simulated students completely resolve their misconception. This reflected our desire to represent student thinking authentically and challenge the naïve belief that misconceptions can be easily resolved in a short exchange (Spangler and Hallman-Thrasher 2014).

Data and Analysis

In this chapter, we describe what questions PSTs selected in each experience (in terms of high leverage or low leverage) and how they evaluated their questions at two time points (Step 8 and Step 10 in Fig. 1). In Step 8, PSTs selected one of three options:

- It was a good question; it accomplished what I wanted it to accomplish
- It was a good question, but [the student] didn't respond in the way I expected
- It was maybe not the best question; I should have asked something different.

After seeing the student's response to the second question (Step 10), they chose from the following options:

- The second question was better than the first
- My first question was better
- They were the same.

For each of these questions, PSTs were asked to type an explanation for their choice. We used a constant comparative process (Glaser and Strauss 1967) to place these explanations into 11 categories according to emerging themes, and then consolidated these themes into five larger conceptual categories. Once codes were agreed upon, we coded approximately 25% of the data individually between two researchers, reaching an agreement rate of 84% on the five large categories. We then coded the rest of the data, individually, and resolved all discrepancies through discussion. Table 2 shows the final codes, some of the most prevalent initial codes, and examples of explanations given by PSTs in each category.

Most of our findings will report numerical patterns in how PSTs answered the multiple choice questions (Steps 5, 8 and 10) across different experiences, focusing mostly on those who chose a low leverage question at Step 5 and whether they expressed doubt about the effectiveness of that question in either Step 8 or 10. However, we will also supplement these findings with summaries of explanation codes (from Table 2) and examples of explanations that PSTs provided to support their choices.

Results

Impact of the Simulated Experiences on PSTs' Question Preferences

The experiences did appear to have some influence on how PSTs thought about the questions they chose initially, particularly if they chose a low leverage question. Figure 2 shows all of the experiences in which PSTs had opportunities to compare the effects of a high and low leverage question.

For example, in the Matthew 2015 experience, 30% of PSTs initially selected a high leverage question in Step 5, and after the experience, 43% preferred that question over the alternative low leverage question that they had viewed in Step 9. Of the 70% who initially chose a low leverage question, only 13% still preferred that question after seeing the high leverage option (the remaining PSTs did not prefer one question over the other).

Table 2 Most prevalent codes for evaluations of questions after seeing student response

Final code	Initial code	Example explanation for choice
Directing: Question led to teacher take over of strategy/ thinking	The question provided an opportunity for the teacher to explain or tell.	"This is enough information for me as the teacher that I need to pull Brandon aside and have a mini lesson with him."
Addressing misconceptions: PST claims that the question helped the student understand, focused on a misconception, or failed to "fix" a misconception	The student understands now.	"Brandon understands that the denominator needs to be the same in order to add the fractions."
	The question directed the student to the misconception.	"I did give him a clue about what he should do next, he just didn't use it to find his answer."
	The question did not fix the student's misconception.	"This question was useless because Brandon has no idea how to find a common denominator."
Understanding student thinking: PST claims that the question helped the teacher to better understand Brandon's thinking or allowed the student to explain his thinking	The question helped the teacher understand the student's thinking.	"I think it was a really great question to ask Brandon because although he did not discover the correct answer I, the teacher got a much better understanding of his thinking."
	The question did not provide information about the students' thinking.	"I was hoping he would give some more explanation as to why he added them all up."
Building on student thinking: PST claims that the question provided an opportunity for Brandon to come to a new realization *on his own*	The question caused (or will cause) student to extend *his own thinking* to come to a new realization.	"I wanted Brandon to realize his confusion without me having to point it out to him. By asking this question, he reevaluated his answer and decided it may have not been the best solution."
	The question was too leading or gives away the answer.	"I should not have asked this question because the teacher gave away the answer and it did not probe Brandon to think on his own about the problem."
Other	Other (does not give a clear evaluation of the question).	"I expected Brandon to ask why the pieces must be the same size in order to add them."

Across all of the experiences, the chart shows inconsistent results for those who started by choosing a high leverage question (left hand side of Fig. 2)—sometimes, after viewing both types of questions, more PSTs expressed preference for the high

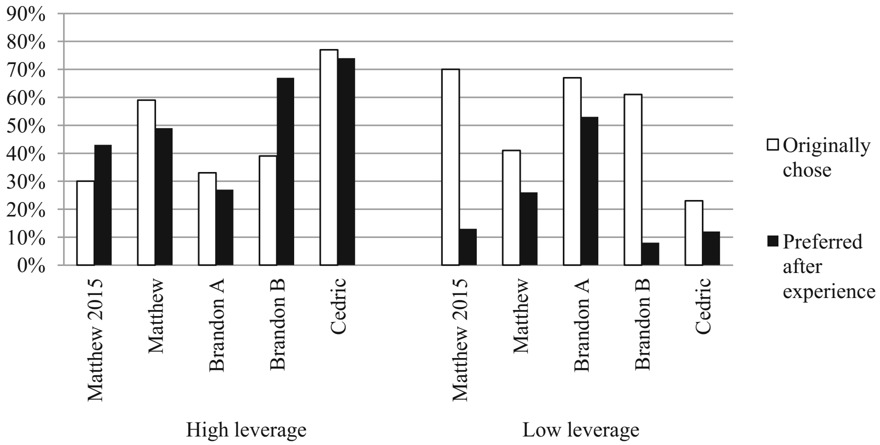


Fig. 2 Percentages of types of questions selected at the beginning of each experience and preferred at the end of each experience

leverage question they had initially selected, and sometimes they did not. In contrast, those who initially selected the low leverage questions (right hand side of Fig. 2) were more consistent in stating that they did not prefer their initial question after viewing both responses. This suggests that in general, the experiences supported PSTs in being more critical of their low leverage questions—but also did not necessarily increase their confidence in their high leverage questions.

The rationales PSTs wrote for their evaluations of these questions gave us indications about why some PSTs were impacted by the responses given by Brandon and some were not. For example, after initially selecting the low leverage question for Brandon A ("In the problem it says that there are three fourths and three sixths. Are fourths the same as sixths?") and seeing Brandon's reaction ("uh, no"), one PST explained why she thought the question was effective:

> I think this was a good question to ask Brandon because he realized that the sixths and fourths are different sized parts. He also realized that you cannot add fractions if the denominators are different numbers (and represent different sized parts). This is leading Brandon in the right direction of adding his fractions again, but the correct way. Because he knows that you cannot add fractions if the denominators are different, Brandon's next step would be to find common denominators.

This PST considered Brandon's response ("uh, no") to constitute evidence of understanding, and conveyed in her reflection that the question helped resolve his misconception. After seeing both questions, this PST still believed that the initial question was better:

> The first question prompted him into knowing that you cannot add fractions if the parts are different sizes and the denominators are different. The second question did not change Brandon's idea about the 6/10ths being incorrect. He was able to identify where the fourths were represented and where the sixths were represented, but he did not notice that the parts were different sizes. He still continued to count all of the parts together.

This was typical of the 39 PSTs who initially chose and then maintained their preference for the low leverage question; all but four of their explanations were coded as “addressing misconceptions.” On the other hand, of the 10 PSTs who changed their minds after going through the Brandon experience, six explanations were coded as “understanding student thinking.” For example, one PST wrote,

> I think the second question was a lot better because we actually get the chance to observe Brandon’s thinking and strategies. He is able to explain his thought process for us. The other question was more of the teacher telling Brandon what is right and what is wrong.

This suggests that when PSTs change their minds about the low leverage question they initially picked, they are doing so because they are attending to positive consequences of the question other than whether it supposedly resolves the student’s misconception; in this case, the PST values getting more information about Brandon’s thinking.

Important Features of the Experiences

The quotes in the previous paragraph suggest that *the opportunity to see the results of different pedagogical actions* was an important part of the experience. That is, the PST’s criteria for effective questioning began to shift when she compared the consequences of a low leverage question with the consequences of a high leverage question. Across all of the experiences, we saw that, indeed, PSTs became more critical of their initial low leverage questions only *after* comparing with a high leverage question (Fig. 3). For example, in the Brandon B experience, 36% of the PSTs who initially selected a low leverage question selected “It was maybe not the best question; I should have asked something different” after seeing Brandon’s response (Step 8). But after comparing with the high leverage question (Step 10), the percentage who selected “The second question was better than the first” was 62%. This shows that more PSTs had begun to doubt the effectiveness of the question they originally chose. In fact, in all of the experiences, more PSTs expressed doubt about their selection after seeing both high and low leverage questions.

This pattern was stronger in some experiences than others. In fact, the Brandon A experience was the least effective in terms of prompting PSTs to be more critical about their initial question choice. The Matthew experience, in contrast, revealed that some PSTs questioned their choice after seeing Matthew’s initial response, but substantially more PSTs doubted their original choice after seeing both questions. Explanations for these patterns are suggested by the PSTs’ evaluations of their questions. For example, after choosing a low leverage question and seeing Matthew’s response, one PST wrote,

> I still think the question was a good question, it shows the teacher that Matthew doesn’t understand what the partitioned pieces represent of the whole. She knows this is where she’ll have to work more with him and maybe the whole class.

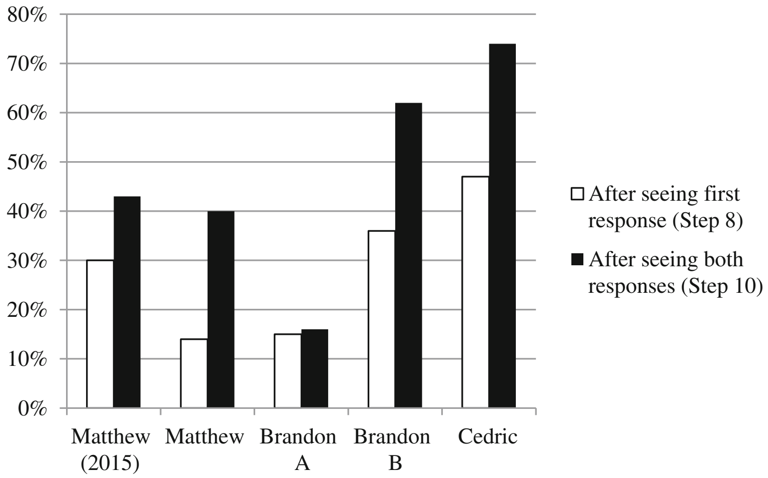


Fig. 3 When PSTs expressed doubt about their low leverage question choice. *Note* The total for each experience is the number of PSTs who initially selected a low leverage question, which was different for each experience

But after seeing Matthew's response to the high leverage question, the same PST wrote,

> The second one might have been better because it has him draw out what he is thinking. Therefore, you can see exactly what he is thinking because sometimes the explanation can get confusing and he might not be saying what he is thinking. But if he draws it out then you know for sure what he is thinking.

This PST, and several of her peers, only questioned her first choice when shown how Matthew responded to the high leverage question. Initially, she based her evaluation of the question on whether Matthew understood the mathematical idea, but after the second question, she based her evaluation on how clearly she could see Matthew's thinking. This suggests that providing a contrast between different pedagogical moves created a learning opportunity for several PSTs, and encouraged them to consider affordances of questions that they might not initially see as valuable (such as "knowing for sure" what a student is thinking).

In summary, the results show that, in general, individual experiences tended to support reductions in the number of PSTs who preferred low leverage questions. PSTs who changed their minds often shifted their criteria for evaluating questions from addressing misconceptions to drawing out or building on student thinking, while those who did not change their minds continued focusing on whether the question resulted in "fixing" a misconception. When we looked more closely at the features of the experiences, we saw that providing a student response to the initially chosen low leverage question (Step 8) resulted in some doubts about this question, but that providing a second response to a contrasting (high leverage) question (Step 10) increased the number of PSTs expressing doubts about that initial question.

Discussion

These findings suggest that the teaching simulations have some potential for challenging PSTs' initial questioning practices and provide some affordances that are not present in other representations. For example, some representations are primarily examples of student work (e.g., Bartell et al. 2013; Jacobs et al. 2010). Preservice teachers are able to analyze this work and even say what they would do in response, but they do not get to see the consequences of their decisions. In our simulations, the user engages in many of the same analyses, but then makes a choice that has an effect. This means that rather than getting feedback about their analysis and decisions from an instructor, feedback is contained *within* the simulation, in the form of the response from the student. This feedback can cause PSTs to reevaluate their interpretation of the students' thinking, their thoughts about the mathematics itself, and/or their pedagogical choice. In this sense, the PST is learning *from* teaching (Hiebert et al. 2007), rather than just learning *about* teaching.

Videos are another popular representations of work used in teacher education (e.g., Beilstein et al. 2017; Sun and van Es 2015; van Es and Sherin 2010). Videos have the benefit of realism—the students and teachers in the videos are real people doing the real work of teaching, in real time. However, this realism comes at a cost. First, the complexity of a video means that there are many things PSTs might pay attention to (what students are wearing, what students in the background are doing, how desks are arranged, etc.), which may or may not be the particular object of learning intended by the teacher educator. A simulated experience, while less realistic, allows the designer to reduce the complexity of an instructional situation to focus attention on specific objects of learning (Herbst et al. 2011). Secondly, as with analyzing student work, when watching a video PSTs do not have the possibility of making any choices. They can watch what happens, but they can only participate vicariously. In this sense, simulations provide the additional affordance of providing the opportunity for PSTs to engage in the scenario and "interact" with the student (albeit in a limited manner), allowing them to do some of the work of teaching rather than just observing and talking about it (Ball and Forzani 2009). Finally, when watching a video, PSTs can only see what actually happened in the recorded episode. In our simulations, PSTs see multiple versions of what *might* have happened, and then consider the affordances of different decisions that could have been made at a particular moment in time. Our data supports the conjecture that this weighing of different outcomes led to increased critique of practice, and indeed, this appears to be one of the main affordances of our simulations.

At the same time, there are certainly challenges involved with designing and using simulations within the Lesson*Sketch* environment. For example, since we created the teacher and student contributions, we cannot be sure that these represent realistic interactions or that PSTs will accept them as possible events that might occur in a real classroom. When designing the experiences, we drew on many of our own experiences working with teachers and students and sought to avoid simplistic or inauthentic interactions, but ultimately, we cannot say that the interactions we designed represent real teaching. On the other hand, the experiences we have designed are, in a sense, stories, and stories need not be true to be educative. What they need to do is feel authentic to the listener; the user of an experience must be able to imagine that real students and teachers could do and say the things depicted.

Another challenge with our simulations is that the interaction between student and teacher are necessarily much shorter than a real interaction. Jacobs and Empson (2016) argue that single talk turns (like posing a question) are sometimes inadequate for capturing the intent of a teaching move "because teachers often need to persist to support or extend children's thinking" (p. 188). Thus, judging PSTs' intentions based on a single question might be viewed as overly simplified. We would agree that real interactions with real students are messier and less structured than in our simulations, and that in such interactions, teachers have many more opportunities to either build on or take over student thinking. The nature of a designed simulation makes it difficult to create scenarios in which users make more than a few consecutive decisions, as possible outcomes increase exponentially with the addition of more decision points.

However, we would also argue that our simulations mitigate this in three ways. First, the PSTs not only select a question, but also explain why they chose the question and then evaluate it afterwards, which gives us more information about where they expect the conversation to go after asking the question. Second, in some experiences, we included indications (e.g., with thought bubbles) about the how the simulated teacher envisions the ensuing conversation, increasing our confidence that PSTs who choose a question are doing so with an understanding of the teachers' intent. Third, the question choices come after the teacher in the scenario has already elicited some initial information about the student's thinking. The question that comes next reveals what the PST plans to do with that thinking. In particular, if the teacher's initial move is to take over student thinking, it is not likely that later moves will start building on thinking. In this case, we assume that the teacher's goal is to direct students towards a particular approach, and indeed our analysis of rationale for PSTs' question critique supports this—PSTs who pick leading questions are much more likely to talk about "fixing" students' misconceptions by explaining or telling. The converse, however, is not assumed. If a teacher begins with a question that draws out or builds on thinking, they may or may not take over student thinking later in the interaction. We have some indications of this in the PSTs' evaluations of their selected questions; for example, in some cases they talked favorably about a high leverage question, explaining that it provided an opportunity for them to explain how to do the problem. These cases give some support to the idea that the PSTs' question choice only gives partial indication of the PSTs' overall intention for an interaction with a student.

A final challenge involves interpreting PSTs' pedagogical decisions across multiple simulation experiences. It has proven difficult to design different pedagogical situations in which we can be confident that the underlying features of questions are similar enough that we can tell whether PSTs are attending to them for consistent reasons. For example, in Brandon B, 33% of PSTs initially selected a high leverage question, but in the Cedric experience, this percentage was 77% (see Fig. 2). Is this because the PSTs' criteria for effective questioning was different, or because there is something about the Cedric experience (the mathematical task, the student work, the question choices) that is influencing their initial choice? There is simply too much variance across the experiences to know. Within an experience, this is less of an issue, because the only variation is the questions—the mathematical task and student work are the same. That is why, instead of comparing directly across simulations in terms of the numbers of PSTs who select each type of question, we have examined change *within* an experience, and documented whether PSTs who start by selecting a low leverage question become critical of that choice.

Conclusion and Future Considerations

Representations of teaching can clearly provide learning experiences for PSTs; simulations are a particular kind of representation that has certain benefits and limitations. Simulations slow down the action, reduce complexity, and allow decisions and reflection on their consequences. Also, because these are completed online, we generate data quickly in a form that is relatively easy to organize and analyze.

One feature that is important for teacher learning, especially in the context of trying out teaching within simulations or rehearsals, is feedback (Baldinger et al. 2016; Lampert et al. 2013); our feedback comes primarily from the simulation itself. In our second year of the project, we decided to include, at the end of each experience, explanations about what was advantageous about the high leverage questions (e.g., they did not do the mathematics for the student, they elicited additional thinking, they provided opportunities for students to build on their own ideas). We wondered whether these would start to be internalized in subsequent experiences. Our data do not allow us to address this question, but this raises the question of whether some additional in-class discussion might further support efforts to help PSTs more critically examine their pedagogical choices. Originally we had hoped that the simulations might become stand-alone modules that could be accessed more widely without requisite in-class activities. Additional testing is needed to see whether these response patterns can be strengthened and become consistent across contexts, and also to see whether they translate into changes in practice in real teaching situations, such as one-on-one tutoring sessions or small group tasks.

Acknowledgements The storyboards presented in Fig. 1 were created with the LessonSketch platform. LessonSketch is designed and developed by Pat Herbst, Dan Chazan, and Vu-Minh Chieu with the GRIP lab, School of Education, University of Michigan. The development of this environment has been supported with funds from National Science Foundation grants ESI-0353285, DRL-0918425, DRL-1316241, and DRL-1420102. The graphics used in the creation of these storyboards are © 2015 The Regents of the University of Michigan, all rights reserved. Used with permission.

References

Amador, J. M., Estapa, A., de Araujo, Z., Kosko, K. W., & Weston, T. (2017). Eliciting and analyzing preservice teachers' mathematical noticing. *Mathematics Teacher Educator, 5*(2), 159–178.

Baldinger, E. E., Selling, S. K., & Virmani, R. (2016). Supporting novice teachers in leading discussions that reach a mathematical point: Defining and clarifying mathematical ideas. *Mathematics Teacher Educator, 5*(1).

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

Bartell, T., Webel, C., Bowen, B., & Dyson, N. (2013). Prospective teacher learning: Recognizing evidence of conceptual understanding. *Journal of Mathematics Teacher Education, 16*(1), 57–79.

Beilstein, S. O., Perry, M., & Bates, M. S. (2017). Prompting meaningful analysis from pre-service teachers using elementary mathematics video vignettes. *Teaching and Teacher Education, 63,* 285–295.

Boaler, J. (2003). Studying and capturing the complexity of practice—The case of the "dance of agency". In N. Pateman, B. Dougherty, & J. Zilliox (Eds.), *Proceedings of the 27th Conference of the International Group for the Psychology of Mathematics Education held jointly with the 25th Conference of PME-NA* (Vol. 1, pp. 3–16). Honolulu, Hawaii.

Borko, H., Peressini, D., Romagnano, L., Knuth, E., Willis-Yorker, C., Wooley, C., et al. (2000). Teacher education does matter: A situative view of learning to teach secondary mathematics. *Educational Psychologist, 35*(3), 193–206.

Chval, K., Lannin, J., & Jones, D. (2013). *Putting essential understanding of fractions into practice in grades 3-5*. Reston, VA: National Council of Teachers of Mathematics.

de Araujo, Z., Amador, J., Estapa, A., Kosko, K., Weston, T., & Aming-Attai, R. (2015). Animating preservice teachers' noticing. *Mathematics Teacher Education & Development, 17*(2), 25–44.

Franke, M. L., Webb, N. M., 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*(4), 380–392.

Givvin, K. B., Hiebert, J., Jacobs, J. K., Hollingsworth, H., & Gallimore, R. (2005). Are there national patterns of teaching? Evidence from the TIMSS 1999 video study. *Comparative Education Review, 49*(3), 311–343.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory*. Chicago: Aldine.

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

Herbel-Eisenmann, B. A., & Breyfogle, M. L. (2005). Questioning our patterns of questioning. *Mathematics Teaching in the Middle School, 10*(9), 484–489.

Herbst, P., Chazan, D., Chen, C.-L., Chieu, V.-M., & Weiss, M. (2011). Using comics-based representations of teaching, and technology, to bring practice to teacher education courses. *ZDM Mathematics Education, 43,* 91–103.

Hiebert, J., Morris, A. K., Berk, D., & Jansen, A. (2007). Preparing teachers to learn from teaching. *Journal of Teacher Education, 58*(1), 47–61.

Jacobs, V. R., & Empson, S. B. (2016). Responding to children's mathematical thinking in the moment: An emerging framework of teaching moves. *ZDM Mathematics Education, 48*(1–2), 185–197. doi:10.1007/s11858-015-0717-0.

Jacobs, V. R., Lamb, L. L. C., & Philipp, R. A. (2010). Noticing of children's mathematical thinking. *Journal for Research in Mathematics Education, 41*(2), 169–202.

Kawanaka, T., & Stigler, J. W. (1999). Teachers' use of questions in eighth-grade mathematics classrooms in Germany, Japan, and the United States. *Mathematical Thinking and Learning, 1*(4), 255–278.

Kazemi, E., & Stipek, D. (2001). Promoting conceptual thinking in four upper-elementary mathematics classrooms. *The Elementary School Journal, 102*(1), 59–80.
Kosko, K. W. (2016). Primary teachers' choice of probing questions: Effects of MKT and supporting student autonomy. *IEJME-Mathematics Education, 11*(4), 991-1012.
Kuntze, S., Dreher, A., & Friesen, M. (2015). Teachers' resources in analysing mathematical content and classroom situations: The case of using multiple representations. In *CERME 9-Ninth Congress of the European Society for Research in Mathematics Education* (pp. 3213–3219).
Lampert, M., Franke, M. L., Kazemi, E., Ghousseini, H., Turrou, A. C., Beasley, H., et al. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education, 64*(3), 226–243.
Milewski, A., & Strickland, S. (2016). (Toward) developing a common language for describing instructional practices of responding: A teacher-generated framework. *Mathematics Teacher Educator, 4*(2), 126–144.
Moyer, P. S., & Milewicz, E. (2002). Learning to question: Categories of questioning used by preservice teachers during diagnostic mathematics interviews. *Journal of Mathematics Teacher Education, 5*(4), 293–315.
National Council of Teachers of Mathematics. (2014). *Principles to actions: Ensuring mathematical success for all*. Reston, VA: Author.
Nicol, C. (1999). Learning to teach mathematics: Questioning, listening, and responding. *Educational Studies in Mathematics, 37,* 45.
Shaughnessy, M., Boerst, T., & Ball, D. L. (2015). Simulating teaching: New possibilities for assessing teaching practices. In T. G. Bartell, K. N. Bieda, R. T. Putnam, K. Bradfield, & H. Dominguez (Eds.), *Proceedings of the 37th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 924–927). East Lansing, MI: Michigan State University.
Sherin, M. G. (2002). A balancing act: Developing a discourse community in a mathematics classroom. *Journal of Mathematics Teacher Education, 5,* 205–233.
Spangler, D. A., & Hallman-Thrasher, A. (2014). Using task dialogues to enhance preservice teachers' abilities to orchestrate discourse. *Mathematics Teacher Educator, 3*(1), 58–75.
Star, J. R., & Strickland, S. K. (2008). Learning to observe: Using video to improve preservice mathematics teachers' ability to notice. *Journal of Mathematics Teacher Education, 11*(2), 107–125. doi:10.1007/s10857-007-9063-7.
Stigler, J. W., Fernandez, C., & Yoshida, M. (1996). Traditions of school mathematics in Japanese and American elementary classrooms. In L. P. Steffe, P. Nesher, P. Cobb, G. A. Goldin, & B. Greer (Eds.), *Theories of mathematical learning* (pp. 149–177). Mahwah, NJ: Lawrence Erlbaum.
Sun, J., & van Es, E. A. (2015). An exploratory study of the influence that analyzing teaching has on preservice teachers' classroom practice. *Journal of Teacher Education, 66*(3), 201–214.
van Es, E. A., & Sherin, M. G. (2010). The influence of video clubs on teachers' thinking and practice. *Journal of Mathematics Teacher Education, 13*(2), 155–176.
Wagner, A. C. (1973). Changing teaching behavior: A comparison of microteaching and cognitive discrimination training. *Journal of Educational Psychology, 64*(3), 299–305.
Webb, N. M., Nemer, K. M., & Ing, M. (2006). Small-group reflections: Parallels between teacher discourse and student behavior in peer-directed groups. *Journal of the Learning Sciences, 15*(1), 63–119.
Webel, C., & Conner, K. A. (2015). Designing simulated student experiences to improve teacher questioning. In *Proceedings of the 37th Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 976–979). East Lansing, MI: Michigan State University.
Webel, C., & Conner, K. (2017). Using simulated teaching experiences to perturb preservice teachers' questioning practices. *Mathematics Teacher Educator, 6*(1), 9–24.
Wood, T. (1998). Alternate patterns of communication in mathematics classes: Funneling or focusing. In H. Steinbring, M. G. B. Bussi, & A. Sierpinska (Eds.), *Language and learning in the mathematics classroom* (pp. 167–178). Reston, VA: NCTM.