Learning from avatars: Learning assistants practice physics pedagogy in a classroom simulator

Jacquelyn J. Chini,1,* Carrie L. Straub,2 and Kevin H. Thomas1

1Department of Physics, University of Central Florida, Orlando, Florida 32816-2385, USA
2College of Education and Human Performance, University of Central Florida, Orlando, Florida 32816-2385, USA

(Received 30 November 2014; published 22 February 2016)

Undergraduate students are increasingly being used to support course transformations that incorporate research-based instructional strategies. While such students are typically selected based on strong content knowledge and possible interest in teaching, they often do not have previous pedagogical training. The current training models make use of real students or classmates role playing as students as the test subjects. We present a new environment for facilitating the practice of physics pedagogy skills, a highly immersive mixed-reality classroom simulator, and assess its effectiveness for undergraduate physics learning assistants (LAs). LAs prepared, taught, and reflected on a lesson about motion graphs for five highly interactive computer generated student avatars in the mixed-reality classroom simulator. To assess the effectiveness of the simulator for this population, we analyzed the pedagogical skills LAs intended to practice and exhibited during their lessons and explored LAs’ descriptions of their experiences with the simulator. Our results indicate that the classroom simulator created a safe, effective environment for LAs to practice a variety of skills, such as questioning styles and time. Additionally, our analysis revealed areas for improvement in our preparation of LAs and use of the simulator. We conclude with a summary of research questions this environment could facilitate.

DOI: 10.1103/PhysRevPhysEducRes.12.010117

I. INTRODUCTION

The experiences of science majors in university departments can often dissuade them from seriously considering a career in teaching. Faculty may explicitly encourage their students to explore careers in research or industry [1]. Universities may implicitly send the message that teaching does not require specialized knowledge by requiring little pedagogical training of faculty and building classrooms that facilitate large, teacher-centered courses. Teaching assistants (TAs) frequently play an important role in supporting faculty in transforming courses to be more student centered and to incorporate active learning, but there are questions about whether these TAs have the skills and beliefs to lead student-centered interactions [2]. It is becoming increasingly important to meaningfully train TAs as calls to transform traditional lecture courses to active learning continue; e.g., see Ref. [3]. High-quality teaching requires one to see ideas from multiple perspectives, anticipate common errors in thinking, and make content accessible to diverse learners [4]. The Colorado Learning Assistant (LA) model has emerged as a method for science departments to change the message they send their undergraduate majors about teaching careers and support course transformation by simultaneously supporting student-centered instruction and engaging majors in pedagogical training [5].

Nearly 60 colleges and universities have adopted the learning assistant (LA) model to prepare undergraduate students for the high quality teaching required for the success of these course transformations [6]. This model makes use of experiential learning, where LAs practice their skills on real students; however, one unintended consequence is that LAs’ mistakes may have negative consequences on students’ learning. Peer-learning models take other forms, such as role play, to give LAs opportunities to practice with peers before working with real students. However, students’ behavior and thinking may not be accurately portrayed during a role play with peers and, consequently, role plays can be less effective than other forms of simulation.

In this study, we explore the use of a mixed-reality classroom simulator [7], which combines components of the real world (here, the LA in a physical classroom) with virtual components (here, student avatars in a virtual classroom), to allow undergraduate LAs to practice pedagogical skills. The highly immersive classroom simulator is currently used to help individuals practice skills for teaching, much like a flight simulator for pilots [8]. As described in Sec. II B, this simulator has proven effective for training inservice and preservice K–12 teachers. Here, we address two research questions:

*Corresponding author.
jchini@ucf.edu
(1) What evidence do we see of LAs practicing pedagogical skills, such as the use of divergent and convergent questions [9], attention to students’ prior knowledge [10], and uptake [11]?

(2) Do LAs experience the simulator as a realistic teaching environment?

We begin by discussing prior strategies for teacher training and summarizing past uses of the simulator in Sec. II. In Sec. III, we describe the lesson the LAs taught in the simulator, and in Sec. IV we describe the framework we used to analyze the LAs’ experiences. In Sec. V we present our results, and in Secs. VI and VII we discuss our findings and their implications.

II. LITERATURE REVIEW

High-quality teaching requires knowledge of content and pedagogy as well as their intersection, or pedagogical content knowledge [12,13]. Teachers engage in reflective practice, integrating content knowledge, pedagogy, and classroom experience to develop pedagogical content knowledge for teaching [13]. This reflective practice takes place both informally through actual teaching experiences and formally in teacher preparation institutes. In this section, we review a variety of methods of training several populations of teachers.

A. Training physics teaching assistants

Researchers have identified a lack of research about the training of physics teaching assistants [2,14]. Meaningful TA training is especially crucial as research-based instructional strategies become more widespread and TAs need to use teaching strategies they may not have been exposed to as students [15]. A few studies have explored strategies to improve graduate teaching assistants’ (GTAs) pedagogical content knowledge, using strategies such as weekly preparation meetings [15] and full courses [16] where TAs experience research-based materials and review sample student responses. Other studies have shown that while teacher buy in to reformed curricula is important for its success, other “social and environmental” factors, such as classroom, department, and institutional levels of implementation likely outweigh the effects of specific GTA training activities [2].

The Colorado LA program has emerged as a model for training undergraduate teaching assistants that leverages these social and environmental factors [5]. The Colorado LA program is based on a conceptual framework of experiential learning, which posits that learning is created through a transforming experience [17]. In many cases, faculty must apply for LAs and preference is given to faculty who want to make their courses more student centered. Similar to the common understanding that students must be actively involved to learn physics content, teachers need opportunities to actively engage with techniques they are learning about in pedagogy courses [18]. In this program, the learning experience for the LAs incorporates content, pedagogy, and practice, and allows undergraduate science majors to explore teaching by combining a weekly seminar about pedagogical skills with opportunities to practice those skills. LAs engage in weekly reflection, as they learn about a technique in seminar, practice it with students, and write about their experiences, with real students serving as test subjects. However, one drawback of this method is that LAs do not have multiple attempts at a particular interaction; the interaction occurs and LAs may reflect on it later, but they cannot "pause," "rewind," and try again, which limits the instructors’ control to focus the LAs’ attention on specific skills, as suggested by Etkina [18].

B. Training preservice and inservice teachers

While there is a lack of research about training physics teaching assistants, much work has been done on the training of inservice and preservice teachers. However, even when preservice teachers learn about student-centered teaching during their preparation, they may have difficulty implementing those methods in their teaching [19]. Research has shown that pedagogical training is more beneficial when participants have the opportunity to practice techniques and reflect on their effectiveness [20]. Some have made use of microteaching to peers to approximate teaching experiences with participants who are not real students [18]. In microteaching, participants take on either the role of “teacher” or “student”, giving the teacher the opportunity to plan a lesson, choose the resources to achieve certain goals, gather research evidence about students’ ideas, and revise the lesson based on student feedback. Since the students are actually classmates, the instructor can pause the lesson to provide feedback, and request the teacher try a different approach. However, peers are not trained to behave like real students, so they may not create an authentic teaching experience. Brown described preservice teachers’ dissatisfaction with such experiences: ‘The tradition of field experience is so firmly entrenched that it is often difficult for students to see any value in alternative activities—they are just not “real” enough” [21] (pg. 307). She adds, “Modern technology, however, may offer some viable and satisfying alternatives” (pg. 307).

I. Virtual reality in teacher education

Hixon and So reviewed existing literature on the role of technology in field experiences for preservice teachers and described a number of advantages and disadvantages to simulated field experiences (which they classified as “Abstract Experience with Model of Reality”) [22]. Compared to traditional field experiences, they found virtual field experiences exposed preservice teachers to a wider variety of teaching and learning environments, allowed peers to have common experiences, promoted
reflectivity, and allowed students to prepare cognitively for their upcoming real-life field experiences. Hixon and So also pointed out several limitations of virtual field experiences, including lack of interaction with teachers and students, limited reality and complexity, and technical problems. Similarly, Brown argued that simulations take advantage of peoples’ natural tendency to learn through mimicry and allow possibilities to be explored in a safe environment before entering the professional arena [21].

Vanlehn, Ohlsson, and Nason discussed several additional advantages of training teachers with virtual students [23]. Unlike real students, simulated students can be “reset” to an earlier state, allowing the teacher to retry a lesson segment. The simulated students’ initial knowledge can be modified to meet the teacher educators’ goals for the experience, and the teacher educator can more easily supervise the interaction. Simulated students will not have their feelings hurt by overhearing teachers discussing their understanding or behavior.

2. Mixed-reality simulation provides a new environment for teacher training

New environments are emerging which harness the power of computer simulation, giving teachers new opportunities to learn pedagogical skills by practicing with virtual characters instead of real students [24]. In this study, we used TeachLivE™ (TeachLivE) [7], a mixed-reality, immersive classroom simulator. In this simulator, users enter a real room with props needed for teaching, such as desks and a white board. However, unlike a real classroom, in the simulator students are computer-generated characters (avatars) displayed on a large monitor. Users interact with the student avatars in real time and hold authentic discussion. Simulations are grounded in the theory of situated cognition, which asserts that “what is learned cannot be separated from how it is learned and used” [25] (p. 88). Because developing pedagogical content knowledge is challenging for teachers [13,26], the use of simulation allows experiential learning to take place via realistic, constructed scenarios that give learners an opportunity to rehearse and learn from mistakes in a safe environment.

The simulator is designed to purposefully implement the critical components of strong simulated environments: personalized learning, suspension of disbelief, and cyclical procedures [24]. The simulator derives from role play, an instructional method dating back to the 1800s, but its unique components create a highly immersive environment that enhances traditional role play with peers [21]. First, users of mixed-reality simulation experience a feeling of presence and suspend disbelief when engaged in a simulation, which is unlike role play. In a role-play scenario, individuals practice teaching to peers who often break character or portray behaviors that might not be authentic, resulting in a less realistic scenario. However, in computer simulation, the learning scenario is enacted consistently, resulting in a realistic portrayal. A group of approximately 75 practicing teachers were polled to determine the realism of the mixed-reality computer simulation in this study. Over 80% agreed that the classroom simulator felt like a real classroom and over 90% agreed that the student avatars accurately portrayed student personalities [27]. Second, a computer simulation delivers a consistent instructional scenario, which is designed using common principles of instructional design, with special attention to providing opportunities for learners to demonstrate targeted skills [24]. The instructor can request manipulations of the environment or interaction, for example, suggesting common student ideas to be used in the avatars’ responses and specifying the willingness of avatars to provide responses. In a role play, when peers are asked to assume the role of a student, they likely receive limited training on how to deliver their performance consistently, resulting in a derailment of instructional objectives. In the mixed-reality simulator used in this study, an interactor, a professionally trained human instructor, orchestrates the computer simulation based on the specific instructional objectives. Interactors are trained in basic pedagogy, interactive performance, and technology, as well as the specific standardization points needed to enact the simulation.

3. Prior research on using a mixed-reality simulator for teacher training

Currently over 50 university and school district partners use TeachLivE for teacher professional learning. Researchers have demonstrated positive effects of mixed-reality simulation on pedagogy and content across content areas with varying experience, skills, and levels of teaching [28–31]. A national research study examined the effects of four 10-min sessions in the simulator using a quasiexperimental prepost group design on the pedagogical practices of teachers [27]. Middle school mathematics teachers (N = 157) were randomly assigned either to a group that received mixed-reality simulation or to a comparison condition. Data were collected on the number and type of questions asked during regular classroom instruction as well as during simulation with student avatars. Teachers significantly improved targeted teaching behaviors while experiencing the simulator scenarios, and those improvements transferred into the teachers’ original classroom with real students. Andreasen and Haciomeroglu conducted a mixed methods study with 15 mathematics teacher candidates investigating mixed reality simulation as part of an undergraduate course to improve their ability to teach content [32]. Each group of three participants rotated through a three-stage cycle of teaching in which one member taught the lesson while the other two observed; the simulation included middle school-aged student avatars with misbehavior. After completing the 15 to 20 min simulation, participants watched a video of their interaction and reflected on their performance in writing. Participants reported the need to balance content delivery with behavior...
management. Elford, James, and Haynes-Smith conducted semistructured interviews with 11 teachers and teacher candidates after using mixed-reality simulation to practice a read-aloud activity during one 5-min session followed by after-action review [30]. Three themes emerged after qualitative data analysis: (a) classroom management skills were needed to successfully complete the simulation, (b) they were able to interact with the student avatars as they would real students, and (c) mixed-reality simulation provided a valuable opportunity for teachers to reflect on their instruction and determine how to improve.

A basis exists for equating time in a simulator to time in the actual environment. The National Training and Simulation Association has developed a Transfer Effectiveness Ratio (TER), which estimates actual equipment training time saved as a function of time spent training on a simulation for flight simulators. While comparisons cannot be directly made between classroom and flight simulators, the TER for flight simulation versus actual flight training indicates a 0.33 ratio, which suggests that for every 3 hours spent in a simulator, 1 hour of actual time could be eliminated [33]. Further, Straub, Dieker, Hughes, and Hynes compared 40 min of online, synchronous, professional development to four 10-min sessions of mixed-reality simulation and found teachers’ skills in questioning increased significantly after simulation sessions, yet there was no significant increase after online professional development [27]. Evidence is still emerging for the exact dosage of simulation to achieve the desired effects of transfer into real environments. In the current study, LAs spent approximately 5 min per session and conducted two sessions.

4. Limitations of prior research

The current study differs from prior research with the mixed-reality simulation because it is conducted with LAs with no formal teacher preparation training outside of the LA program offered at the university. While an early research base exists for preparation of teacher candidates [30,32] and practicing teachers [27], a review of literature yielded no studies related to using mixed-reality simulation with LAs. Participants in prior studies had been trained in teacher preparation institutes and received extensive instruction in pedagogy. While LAs trained in the Colorado LA model learn about evidence-based instructional approaches and engage in reflective practice, they have less time dedicated to instructional approaches. It is important to investigate whether a mixed-reality simulation of teaching would be effective for users who have spent less time in formal teacher preparation settings.

III. METHODOLOGY

A. Participants

The present study includes fourteen participants who were enrolled in the university LA program, based on the Colorado LA model. LAs completed the simulator activity (simulation and reflections, described below) as part of their seminar course. No stipend for participation was provided. Fifteen students (100%) participated in the activity and agreed to contribute their data to the research project; of these, fourteen students (93%) completed the written reflections before and after the simulator. The simulator activity was recorded for nine students (60%). Demographic information for students included in the two data sets is summarized in Table I. Most LAs were in their first semester of the LA program. The activity was conducted about two-thirds into the semester.

### Table I. Participant demographics. 

<table>
<thead>
<tr>
<th>Demographic</th>
<th>Reflection data</th>
<th>Video data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Man</td>
<td>9 (64%)</td>
<td>7 (78%)</td>
</tr>
<tr>
<td>Woman</td>
<td>5 (36%)</td>
<td>2 (22%)</td>
</tr>
<tr>
<td>Class standing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sophomore</td>
<td>2 (14%)</td>
<td>2 (22%)</td>
</tr>
<tr>
<td>Junior</td>
<td>3 (21%)</td>
<td>2 (22%)</td>
</tr>
<tr>
<td>Senior</td>
<td>9 (64%)</td>
<td>5 (56%)</td>
</tr>
<tr>
<td>Major</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physics</td>
<td>6 (43%)</td>
<td>3 (33%)</td>
</tr>
<tr>
<td>Engineering</td>
<td>3 (21%)</td>
<td>3 (33%)</td>
</tr>
<tr>
<td>Biology or Biomedical</td>
<td>4 (29%)</td>
<td>2 (22%)</td>
</tr>
<tr>
<td>Health Sciences</td>
<td>2 (14%)</td>
<td>1 (11%)</td>
</tr>
<tr>
<td>LA experience</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st semester</td>
<td>10 (71%)</td>
<td>6 (67%)</td>
</tr>
<tr>
<td>2nd semester</td>
<td>4 (29%)</td>
<td>3 (33%)</td>
</tr>
</tbody>
</table>

B. Setting

TeachLivE is a multicomponent system housed on secure servers at the University of Central Florida. LAs completed the activity on the university campus in a room that housed the TeachLivE classroom simulator. For operation, the lab requires a high definition monitor approximately 60” in size, a webcam, speaker, and an Internet connection. As with many highly immersive simulators used for military defense scenarios, a trained professional orchestrates the behaviors of the student avatars using a combination of prerecorded behaviors and human-in-the-loop interactions, resulting in an interactive, authentic simulation. During their interaction with the simulator, the LAs wore a portable microphone and their movements were tracked by a motion-sensing input device (Kinect™). (See Ref. [24] for further details about using TeachLivE.) LAs participated in groups of three and observed their group mates’ interactions with the simulator to provide feedback. The simulator included five high school-aged student avatars, displayed in Fig. 1, designed to represent a range of common personalities. In the present study, the instructor supplied the interactors with an outline of the activity which included both correct responses to the
student prompts, suggested research-based student ideas, and pedagogical goals for the LAs’ sessions. The instructor also selected the level of misbehavior exhibited by the student-avatars to be low so the LAs could focus on pedagogical skills over classroom management.

C. Mixed-reality simulation activity

LAs practiced leading discussions among the five student avatars, based on a motion graph activity taken from the Open Source Tutorials in Physics Sensemaking, Suite 1 [34]. Figure 2 summarizes the activities LAs participated in before, during, and after the simulation. We used the Action Review Cycle to structure these activities; LAs planned what they hoped to learn (Before Action Review; BAR), experienced the simulation (Action), and reflected on their performance in relation to their goals (After Action Review; AAR) [35]. For the BAR and AAR, LAs wrote responses to reflective prompts [Figs. 2(b) and 2(e)]. Prior to delivering the lesson in the simulation, each LA was randomly assigned to one of three lesson segments about motion graphs; LAs were instructed to get the student avatars to express their ideas and to use questions to move student avatars toward the accepted scientific ideas. To allow for adjustment to using the simulation and to get to know the student avatars, each LA began with a 2 min orientation session [Fig. 2(c)]. Next, each LA was given approximately 5 min to lead a discussion about his or her activity prompt [Fig. 2(d)]. After all three LAs completed their first main activity discussion, the LA group reflected on what had gone well and what they could try differently, with suggestions from course instructors and simulation facilitators. Then, each LA was given another 5 min to repeat his or her discussion. LAs completed their final reflection in writing [Fig. 2(e)]. On average, the LAs used the simulator for 5.3 min (minimum: 5.0; maximum: 6.4) during round 1 and 4.7 min (minimum: 4.0; maximum: 6.5) during round 2. We defined an LA’s turn of speech as starting when the LA started speaking and ending when a student avatar started speaking. On average, each LA had 20 turns of speech (minimum: 11; maximum 28) during round 1 and 17 turns of speech (minimum 11; maximum 23) during round 2.

D. Data analysis

1. Coding of pedagogical skills

Our coding framework (discussed in Section IV) is grounded in the pedagogical strategies that the LAs discussed in their written reflections (BARs and AARs) and verbal group reflections. We used cycles of open, axial and selective coding, and constant comparison to identify pedagogical strategies and to link evidence from the videos with evidence from the written reflections [36,37]. We then applied this framework to LAs’ written and group reflections to identify the skills they intended to practice or assessed themselves as practicing in the simulator and their interactions with the simulator to identify the skills practiced. Two raters independently coded the data in stages, discussing and refining the coding definitions at each step, and then re-coding the data (constant comparison). Interrater reliability was assessed with Cohen’s kappa [37]; Cohen’s kappa is an improvement over “percentage agreement” because it takes into account how frequently the raters would agree by chance [38]. Landis and Koch proposed a scale for interpreting Cohen’s kappa values, with less than 0 as poor, 0–0.20 as slight, 0.21–0.40 as fair, 0.41 and 0.60 as moderate, 0.61–0.80 as substantial, and 0.8–1.00 as almost perfect [39]. We applied the method used by Smith and colleagues to generate a Cohen’s kappa value for our code set [40]. Cohen’s kappa values ranged from 0.48 (moderate) to 0.86 (almost perfect) across the eighteen sessions, with an average of 0.75 (substantial).
Two sessions were in the moderate range, ten were in the substantial range, and six were in the almost perfect range. Disagreements between the two raters were discussed and resolved to determine the final codes. An external rater coded 20% of the data and had a Cohen’s kappa of 0.80 compared to the final code set.

2. Coding of LAs’ Impressions of the Simulator

Codes emerged from students’ responses related to the simulator’s usefulness for practicing pedagogical skills in a college physics environment in terms of its similarities and differences from their real-life teaching experiences [see Fig. 2(e) for complete prompt]. Students were then classified as finding the simulator “realistic” if they had only advantage codes, as “nonrealistic” if they had only disadvantage codes, and as “mixed” if they had codes of both types. One researcher pulled common phrases from students’ responses, which two researchers independently applied to the response set. The initial Cohen’s kappa was 0.72. The two researchers discussed and revised the code descriptions and reanalyzed the data, with a resulting Cohen’s kappa of 0.85. After further discussion, all disagreements were resolved. Finally, an external rater analyzed the data, again resulting in Cohen’s kappa of 0.85 between the final set (presented here) and the external rater. Importantly, the external rater only disagreed on one LA-level classification.
IV. THEORETICAL FRAMEWORK

A. Development of pedagogical skills requires active engagement

Many research-based instructional strategies for teaching physics content rely on theories such as constructivism, emphasizing that students must build new knowledge for themselves [41], and the resources view of knowledge, emphasizing that students bring to any subject productive seeds of knowledge and skills [10]. The literature on teacher preparation has recently called into service these same perspectives to describe the development of pedagogical content knowledge (PCK). For example, Etkina writes, “If one cannot learn physics by just listening and reading but needs to engage in the active process of knowledge construction, the same should apply to PCK; one can only acquire PCK by actively constructing it in the process of teaching (called clinical practice)” [18] (pg. 2). The use of a classroom simulator to practice pedagogical skills is grounded in this perspective, and represents a new modality of clinical practice. Harlow and colleagues have described pedagogical resources as “small, discrete ideas about teaching... activated in specific contexts to understand the teaching and learning of science” [42] (pg. 1099). The Action Review Cycle used with the mixed-reality simulator provides an opportunity for the identification and development of the resources LAs bring to physics teaching. In this study, we adopt this perspective by using the LAs as our main source for developing the framework of pedagogical skills described in Sec. IV B and will further build on this perspective in future studies.

B. Framework for pedagogical skills

We developed a framework for analyzing the teaching strategies used by the LAs in the simulator from the common strategies described by the LAs in their reflections. The strategies described by the LAs were likely influenced by their exposure to topics in the pedagogy seminar. Common strategies that emerged from the LAs’ reflections included eliciting students’ prior knowledge, use of divergent or open versus convergent or closed questions, use of leading questions, use of wait time, use of real life examples, getting all students in the group involved, and (proximal) formative assessment. Because we coded LAs’ pedagogical skill use by turn of speech, we introduced “uptake” to help identify periods of proximal formative assessment. Nystrand et al. define uptake as one person asking another person about something the person previously said [11]. Uptake has the effect of valuing a students’ contribution, likely facilitates the negotiation of understandings, and promotes coherence within the discourse. We expect periods of proximal formative assessment will be marked by cycles of eliciting students’ prior knowledge and uptake. Table II displays examples of how LAs identified and enacted each of these skills and references the literature LAs were exposed to through their seminar related to each topic.

Several codes required special attention for coders to consistently code LAs’ turns of speech. In particular, the questioning codes may overlap. A question was coded as “closed” if it was seeking one particular answer (convergent) or only required the student avatar to respond with “yes” or “no.” On the other hand, a question was coded as “open” if it allowed the student avatar to provide a variety of responses (divergent). A question could be co-coded as open or closed and “leading” if it was used to guide the student avatars’ answers or explanations in a particular direction [44]. In general, a turn of speech could be coded with zero or multiple codes (see Figs. 3 and 4, for example).

C. Framework for strong simulated environments

In order to be used effectively, simulated environments must provide personalized learning, suspension of disbelief, and cyclical procedures [24]. A personalized environment allows participants to choose the skills they would like to practice. Suspension of disbelief occurs when a participant feels physically and cognitively engaged in the simulated environment as the result of a high level of engagement between the participant and the technology [52]. The activity in this study was intentionally built on the cyclical process of the Action Review Cycle at multiple scales. First, written reflections engaged the LAs in planning what they would practice (Before Action Review) and reflecting on how their performance matched their expectations (After Action Review) [35]. Second, the LAs made two attempts at leading a discussion about their activity prompt, with group reflection after the first attempt. Past research has indicated that the mixed-reality simulator used in this study meets these conditions for in-service and preservice teachers [24]. A future study will expand on how the mixed-reality simulator met these requirements for our population, who had less prior exposure to teacher training and teaching experiences.

V. RESULTS

A. Pedagogical strategies

1. Examples of pedagogical skill use

LAs’ pedagogical skill use was coded per turn of speech. Figures 3 and 4 display the skills exhibited by two LAs during their first and second trials in the simulator. The transcripts for these interactions are included in the Supplemental Material [51]. LA-3C was observed to make several positive changes from the first round to the second round. For example, he increased the frequency with which he used open questions and decreased the frequency with which he used closed questions. On the other hand, the changes in LA-3A’s skill use were more mixed. He was
observed to decrease his use of questioning, but to increase his use of wait time. Thus, the LAs were observed to make a variety of changes in their second attempts with the simulator.

2. Observed strategies across LAs

LAs’ conversations with the student avatars, group reflections between simulator teaching experiences, and written reflections (BARs and AARs) were analyzed for the skills described in Sec. IVA. Figure 5 displays the skills observed at least once for each student (whereas Figs. 3 and 4 displayed all the skills used by individual LAs). We did not code for formative assessment during the teaching episodes because we defined it as cycles of other coded skills (eliciting students’ prior knowledge and uptake). However, LA-2A and LA-3B mentioned formative assessment in their BARs, LA-2B, LA-3A, and LA-3B mentioned it during the group reflection, and LA-1C, LA-2A, and all LAs in group 3 mentioned it in their AARs.

### TABLE II. Framework for pedagogical skills

<table>
<thead>
<tr>
<th>Pedagogical skill</th>
<th>LA identification</th>
<th>LA enactment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eliciting students’ prior knowledge [43]</td>
<td>LA-1C: “I would use questioning techniques to understand what their intuitive answers are and probe them to better understand why they think that.”</td>
<td>“Do you know what the definition of velocity is?”</td>
</tr>
<tr>
<td>Use of divergent (open) vs convergent (closed) questions [9]</td>
<td>LA-2B: “The most effective strategy I have at my disposal are the aforementioned questioning strategies. Using open ended questions… I can hear their thoughts on the connection between velocity and displacement.”</td>
<td>Closed: “Can anybody tell me what the slope of that would be?” Open: “Ok, that’s a good observation. So Ed how would you put that in other words?”</td>
</tr>
<tr>
<td>Use of Leading Questions [44]</td>
<td>LA-1A: “I think the main thing is that I would use questions in order to lead them to the answer.”</td>
<td>“So Maria, let’s say from 0 to 4, where do you think the velocity would be, knowing that it’s just the slope?”</td>
</tr>
<tr>
<td>Use of Wait Time [9]</td>
<td>LA-2C: “One aspect of questioning which we have learned about is wait time, which is extremely helpful when questioning students. Giving an ample amount of wait time allows the students to think through the question thoroughly and give better answers.”</td>
<td>Wait time was coded when the LA waited more than three seconds after asking a question before prompting students to answer or rephrasing the question. The percentage was calculated out of the opportunities LAs had to use wait time; the total possible instances was reduced by the number of times a student-avatar responded to a question within three seconds.</td>
</tr>
<tr>
<td>Group Involvement [45,46]</td>
<td>LA-3C: “Hopefully they can have a back and forth, high quality discussion to further their learning.” and “I think the biggest thing would be trying to get CJ involved since that’s the last person who hasn’t really been involved much at all.”</td>
<td>“Let’s get an opinion from someone else. Anyone else? CJ? Maria?”</td>
</tr>
<tr>
<td>Real World Examples</td>
<td>LA-2A: “I can build a bridge by using proximal formative assessment and using real life examples to help explain the concept.”</td>
<td>No strong examples identified.</td>
</tr>
<tr>
<td>(Proximal) Formative Assessment [47–50]</td>
<td>LA-3B: “I will attempt to use open and closed questions for the appropriate situations and perhaps proximal formative assessment to enhance the learning and instruction of this interaction as much as possible.”</td>
<td>Coding described in text.</td>
</tr>
<tr>
<td>Uptakea</td>
<td>This skill was added by the researchers.</td>
<td>“A lot of people get confused at the, uh, the flat part. They think that you’re moving at a constant speed, but since this is a position vs time graph, like Kevin said earlier, we actually don’t know anything about speed, really. We just know how the position’s changing.”</td>
</tr>
</tbody>
</table>

aLAs were not introduced to uptake.
FIG. 3. Pedagogical skill use by LA-3C. This LA increased his use of open questions in the second round of teaching the student-avatars. The transcripts for these interactions are available in the Supplemental Material. Codes are as follows: Eli = Elicit Prior Knowledge; OpQ = Open Question; ClQ = Closed Question; LdQ = Leading Question; RWE = Real World Example; WGp = Whole Group Involvement; Upt = Uptake; WT = Wait Time.

FIG. 4. Pedagogical skill use by LA-3A. This LA decreased his use of open questions, while increasing his use of wait time. The transcripts for these interactions are available in the Supplemental Material [51].
As shown in Fig. 5, the full range of identified pedagogical skills was observed during the simulated teaching experiences (although there were only weak examples of attempts at real world examples). All LAs were observed to make use of open and closed questioning, whole group involvement, and wait time greater than 3 sec in both rounds of their teaching. Most LAs also made use of leading questions during both rounds. Several LAs elicited students’ prior knowledge, but few made use of real world examples or uptake.

All LAs were able to practice at least one of the skills they had personally identified in their BAR, and eight of nine LAs practiced all of the skills they had identified from our framework. LA-2A identified real world examples as a skill he would use, but was not observed to practice this skill. However, he did identify this as a skill he needed more practice on in his AAR. LA-1B also identified in her AAR that she had not made use of real world examples while teaching the student avatars. LA-3B and LA-3C both described eliciting students’ prior knowledge in their AAR, but were not coded for this practice during their teaching. Three LAs (LA-1C, LA-2A, and LA-2B) were observed to practice a new skill in their second round of teaching.

### 3. Frequency of observed pedagogical skills

Table III displays the frequency of use for each pedagogical skill per LA in round 1. Frequency of use of wait time is given as a percentage of turns of speech in which the LA exhibited wait time out of the turns of speech that presented an opportunity for wait time; the total turns of speech for wait time calculations were reduced by the number of times the student avatars answered a question.

Table III. Frequency of pedagogical skill use: Round 1.

<table>
<thead>
<tr>
<th>Round 1 Turns of Speech</th>
<th>Elicit prior knowledge</th>
<th>Open question</th>
<th>Closed question</th>
<th>Leading question</th>
<th>Real world example</th>
<th>Whole group uptake</th>
<th>Wait time &gt;3 sec</th>
</tr>
</thead>
<tbody>
<tr>
<td>LA-1A</td>
<td>22</td>
<td>1</td>
<td>6</td>
<td>9</td>
<td>4</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>LA-1B</td>
<td>28</td>
<td>3</td>
<td>2</td>
<td>12</td>
<td>7</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>LA-1C</td>
<td>21</td>
<td>2</td>
<td>5</td>
<td>11</td>
<td>3</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>LA-2A</td>
<td>22</td>
<td>2</td>
<td>6</td>
<td>9</td>
<td>5</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>LA-2B</td>
<td>20</td>
<td>0</td>
<td>10</td>
<td>7</td>
<td>5</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>LA-2C</td>
<td>25</td>
<td>0</td>
<td>7</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>LA-3A</td>
<td>11</td>
<td>1</td>
<td>6</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>LA-3B</td>
<td>11</td>
<td>0</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>LA-3C</td>
<td>18</td>
<td>0</td>
<td>2</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>
within the 3 sec window so as to represent the number of times the LA could have exhibited wait time greater than 3 sec. In both rounds, the most frequently used skill was closed questioning, followed by wait time and open questioning. As previously mentioned, use of real world examples, eliciting students’ prior knowledge, and uptake were used infrequently.

Table IV shows the shifts in strategy use from round 1 to round 2 per LA. Combining data from Fig. 5 with data from Table IV, we observe that six LAs (LA-1A, LA-1C, LA-2A, LA-3A, LA-3B, and LA-3C) increased use of at least one of the skills identified in their BAR or group reflection from round 1 to round 2. Comparison to Fig. 5 reveals that several of these changes may have been motivated by the group reflection that followed round 1. For example, LA-3A discussed needing to use more wait time, and was observed to do so. The instructor suggested that all LAs in group 3 should make more use of open questions, and LA-3C was observed to increase his use of open questioning and drastically decrease his use of closed questions.

TABLE IV. Shifts in pedagogical skill use from round 1 to round 2.

<table>
<thead>
<tr>
<th>Round 2 Turns of Speech</th>
<th>Elicit prior knowledge</th>
<th>Open question</th>
<th>Close question</th>
<th>Leading question</th>
<th>Real world example</th>
<th>Whole group</th>
<th>Uptake</th>
<th>Wait time &gt;3 s</th>
</tr>
</thead>
<tbody>
<tr>
<td>LA-1A</td>
<td>19</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>LA-1B</td>
<td>23</td>
<td>-2</td>
<td>1</td>
<td>-3</td>
<td>-4</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>LA-1C</td>
<td>22</td>
<td>-2</td>
<td>-1</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>-2</td>
<td>0</td>
</tr>
<tr>
<td>LA-2A</td>
<td>18</td>
<td>-1</td>
<td>4</td>
<td>-2</td>
<td>-2</td>
<td>0</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>LA-2B</td>
<td>14</td>
<td>2</td>
<td>-9</td>
<td>2</td>
<td>-5</td>
<td>0</td>
<td>-2</td>
<td>1</td>
</tr>
<tr>
<td>LA-2C</td>
<td>19</td>
<td>0</td>
<td>-3</td>
<td>-3</td>
<td>-2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LA-3A</td>
<td>14</td>
<td>-1</td>
<td>-3</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>LA-3B</td>
<td>11</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>-2</td>
</tr>
<tr>
<td>LA-3C</td>
<td>12</td>
<td>0</td>
<td>2</td>
<td>-6</td>
<td>0</td>
<td>0</td>
<td>-2</td>
<td>0</td>
</tr>
</tbody>
</table>

TABLE V. Suspension of disbelief codes. Examples of quotes from the LAs’ reflections on the simulator and the number of LAs exhibiting each code.

<table>
<thead>
<tr>
<th>Code description</th>
<th>Exemplar quote</th>
<th>Number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Common similarities between simulator and real life</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avatars had realistic personalities</td>
<td>“It was very similar with regards to the different types of students that we interact with in a real classroom. Every student had their own personality and way of learning just like in real life.”</td>
<td>8</td>
</tr>
<tr>
<td>Simulation felt natural with time</td>
<td>“At first, it felt weird using the simulator. But I got used to it quickly and it felt natural.”</td>
<td>7</td>
</tr>
<tr>
<td>Avatars exhibited similar conceptual difficulties as real students</td>
<td>“The students definitely had similarities to real students in that they shared common principle difficulties.”</td>
<td>3</td>
</tr>
<tr>
<td><strong>Common differences between simulator and real life</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Possible movements were restricted</td>
<td>“The movement was a little bit awkward for me and so I found myself just standing in one spot, whereas in a real classroom I’m normally much more active.”</td>
<td>4</td>
</tr>
<tr>
<td>Reduced tools (e.g., white board, paper)</td>
<td>“The major difference between the simulator and a real classroom was the inability to physically draw a diagram or ask the students to show a diagram that would have helped with testing their understanding for velocity and linear motion.”</td>
<td>4</td>
</tr>
<tr>
<td>Missing nonverbal cues</td>
<td>“As realistic as they were though, I felt disconnected and out of place working with the animated students. I hadn’t realized until afterwards just how much I depend on reading behaviors and emotionally connecting with my students.”</td>
<td>3</td>
</tr>
<tr>
<td>Avatar’s behaviors and attitudes were more extreme than reality</td>
<td>“The students also seemed a little stereotypical, and their attitudes were very pronounced.”</td>
<td>2</td>
</tr>
<tr>
<td>Unrealistic pauses in conversation</td>
<td>“It was strange because there was a time lag and I felt I had to be more intentional in how I spoke, the words I used, and my intonation.”</td>
<td>2</td>
</tr>
<tr>
<td>Unrealistic audio or visual</td>
<td>“I think the only thing that was wrong with the program was that sometimes it would make the caricatures have sharp movements but the rest was almost identical to a real classroom.”</td>
<td>2</td>
</tr>
</tbody>
</table>
B. LAs’ descriptions of their experiences with the simulator

As discussed in Sec. III. D., codes for LAs’ views of the simulation were developed from the LAs’ written AARs. The codes, examples of quotes related to each code, and the number of LAs exhibiting each code are presented in Table V. Most LAs (8/14) found the student-avatars exhibited realistic personalities, while fewer students (2/14) found the attitudes and behaviors to be more pronounced than in real students. A few LAs (3/14) remarked that the student-avatars exhibited similar conceptual difficulties as real students. Half of the LAs (7/14) commented that they quickly felt at ease using the simulation. However, the LAs did point out several ways in which the simulator was different than their real teaching experiences, including restricting their physical interaction (4/14) and nonverbal communication (3/14) with the students, not providing tools they typically use to teach (4/14), and unnatural pauses in communication (2/14). In addition, a couple students (2/14) commented that the overall appearance and sound of the simulator was unrealistic.

Several students mentioned additional similarities or differences between the simulator and their real-life teaching experiences. For example, one student commented that the imposed time limit was similar to the time limits that he operated under during class, while another student liked that the student avatars were able to perform calculations with calculators like real students. On the other hand, one student felt that wearing the microphone made the simulation feel unnatural, while another commented that the student group in the simulation was larger than the groups he typically worked with.

To determine the LAs’ overall view of the simulator, we looked across codes within each LA. LAs were classified as finding the simulator “realistic” if they had only advantage codes, as “nonrealistic” if they had only disadvantage codes, and as “mixed” if they had codes of both types. Most LAs discussed both realistic and nonrealistic attributes of the simulator (10/14, 71.4%). Three students (21.4%) described only positive attributes, while one student (7.1%) only described nonrealistic attributes.

VI. DISCUSSION

1. Ease of practice varied by pedagogical skill

The LAs were observed to practice an array of pedagogical skills. Some skills, such as various forms of questioning and whole group involvement, appeared to be easy to implement and were implemented frequently. We were surprised to find that some skills, such as use of real world examples, were not implemented more frequently. Occasionally, the student avatars actually made a bid for a real world example. For example, in the following passage the student-avatar “CJ” tries to relate velocity to her experiences with driving, but the LA does not carry on the example.

CJ: Um, yeah that’s like, like how hard you’re going or how fast you’re going. Right? Like how, like how much you, like when you step down on the gas of your car, that’s like, like making you, like have more velocity.

LA-2A: Right, so if you speed up, you have a change in position over time, right? Make sense?

There are several reasons that the LAs may not have made use of real world examples. First, the opportunity may not have arisen; however, as demonstrated above, this was not the case for some LAs. Second, the LAs may not have “framed” the interaction with virtual students as an appropriate place to use real world examples [53]. However, there is evidence against this explanation as well. The student avatars made bids for real life examples, which should have caused the LAs to reframe the interaction if they had previously been excluding such examples. Additionally, two LAs discussed planning or failing to use real life examples in the simulator. Thus, it appears that the most likely explanation is that the LAs need more practice in linking the physics concepts to real life examples. It should also be noted that LAs should not be expected to demonstrate every skill within a 5 min interaction.

Another skill the LAs made less use of than expected was formative assessment or proximal formative assessment. Although we did not outright code the LAs’ simulator interactions for formative assessment, we expected formative assessment events would be marked by cycles of eliciting students’ prior knowledge and uptake, and we did not observe many such cycles. Six LAs discussed formative assessment during at least one written or group reflection. Some of these LAs discussed difficulty implementing formative assessment, even proximal formative assessment, with the simulator. For example, LA-1C wrote, “I tried to use formative assessment to assess where students were in their understanding. This was difficult and not always successful because the students were at such varying levels. It was not as useful as it is in a longer period of time in studio classes” in her AAR. It may be the short teaching intervals did not give the LAs enough time to practice some of the skills they desired, although the short time interval is a fairly realistic match to the amount of time most LAs can devote per group in their real life teaching assignments.

2. Learning assistants experienced the simulator as a realistic teaching environment

LAs’ descriptions of the simulator indicated that it provided a realistic teaching experience. Most LAs discussed only either advantages or a mixture of advantages and disadvantages to using the simulation to practice physics pedagogy skills. Several LAs would have liked to have access to a whiteboard, which could be easily facilitated in future sessions. Others requested additional capabilities of the student avatars, such as being able to
share drawings with the LA, which would require changes to the simulator program.

Another way to assess the nature of the simulation is to check whether LAs were able to practice the skills they intended to practice and make improvements from round 1 to round 2. We find that all LAs were able to practice at least one identified skill and most practiced all personally identified skills from our framework. Several LAs were observed to improve their performance from their first to second trials and to identify skills with which they needed more practice. Our study made use of back-to-back trials. Other methods, discussed in Sec. VII below, may facilitate additional improvements between trials.

Additionally, we have evidence that the simulator caused LAs to reflect meaningfully on their practices, an important component of teacher training [20]. Several LAs were able to identify skills they intended to practice, but had not used during the simulation. For example, one LA wrote, "I feel like I could have done better and maybe tried to relate it to something they were familiar with." One LA realized how much she depends on reading her students' behaviors and emotions (see Table V). These self reports indicate that the mixed-reality simulator engaged the LAs in meaningful reflection.

3. Limitations

The purpose of this study was to demonstrate that a mixed-reality classroom simulator is a useful environment for physics LAs to practice pedagogical skills. We offer several precautions in other interpretations of the data presented. First, the set of pedagogical skills analyzed in this paper does not span the full set of skills identified by the LAs. For clarity, we chose to focus on a specific set that was common among the LAs and supported by the literature on effective teaching. Second, since the LAs used the simulation in groups of three, it is possible that they influenced each other; for example, if an LA observed his peers frequently asking open questions, he may push himself to ask open questions. We consider this a feature given our focus on practicing skills, but a different method, such as those described in Sec. VII below, may be needed to address other research questions.

It should also be noted that in the present study we did not code for how well the LA succeeded in using each skill. For example, returning to Table II, we see that there is a disconnect between the example LA descriptions and enactments of certain skills, such as eliciting prior knowledge and asking open questions. This reveals a potentially useful way the mixed reality simulator could be used both to facilitate LAs' reflection about how their actions match their goals and to track changes in how a skill is used over multiple sessions.

VII. IMPLICATIONS AND FUTURE WORK

Initial results indicate that the mixed-reality simulator provided a safe, effective practice environment for LAs, allowing them to make mistakes and try new approaches with student avatars rather than real students. Since the simulator makes use of a trained “interactor” to orchestrate avatar interactions with participants, the teacher educator has more control over the classroom response than in traditional microteaching exercises. The instructor takes part in a design process which gives specific points of standardization for the simulation experience to be used as a guide by the interactor. For example, the interactors can be trained to make use of common student ideas for the practice of formative assessment or certain misbehaviors for the practice of classroom management skills. The simulator could also be useful for a summer “boot camp” for teaching assistants, when real students may not be available.

Findings suggest that mixed-reality simulation can be leveraged to sustain improvement in higher education instruction and development of LAs’ pedagogical skills by identifying areas where teacher educators should focus their efforts. These initial results suggest that our LAs may need more practice with skills such as formative assessment and real world examples.

A mixed-reality simulator also represents a novel environment in which to explore several research questions. For example, LAs or GTAs could use the simulator at the beginning and end of the semester to assess changes in their approaches to teaching. LAs could use the simulator repeatedly over the semester to practice difficult skills, such as proximal formative assessment. Their use of these skills in the simulator could be correlated to their use in real classroom interactions to assess whether skills transfer from the mixed-reality simulator to real classroom practice. Finally, the most effective methods for using the simulator (individuals versus groups; time between reteaching episodes) could be explored for incorporating this exercise into pedagogy courses.

ACKNOWLEDGMENTS

The authors thank Jarrad Pond and Matthew Wilcox for assisting with data analysis and the physics learning assistants for their participation. This work was supported in part by National Science Foundation Grants No. DUE1347515 and No. DUE1246024 and a PhysTEC Grant (supported by the National Science Foundation under Grants No. 0808790, No. 0108787, and No. 0833210.) Since submission of this manuscript, C. S. has taken a position with Mursion.


LEARNING FROM AVATARS: LEARNING …


[47] F. Erickson, chapter 8 Some Thoughts on “Proximal” Formative Assessment of Student Learning, Yearbook of the National Society for the Study of Education 106, 186 (2007).


