TeachLivE National Research Project

Year 2 Findings

University of Central Florida
Carrie Straub, Ph.D.
Lisa Dieker, Ph.D.
Michael Hynes, Ph.D.
Charles Hughes, Ph.D.
Using Virtual Rehearsal in TLE TeachLivE™ Mixed Reality Classroom Simulator to Determine the Effects on the Performance of Science Teachers:
A Follow-up Study (Year 2)

This research project was supported by funding from the Bill & Melinda Gates Foundation.

Abstract

A follow-up study, Phase II, was conducted to mirror procedures from Phase I of a quasi-experimental, pre-post group design examining the effects of repeated virtual rehearsal sessions in a mixed-reality computer simulated environment, the TLE TeachLivE™ (TeachLivE) classroom simulator (see Straub, Dieker, Hynes, & Hughes, 2014). In the first phase, 157 middle school mathematics teachers across the nation received four levels of innovative professional development, including computer simulation, synchronous online instruction, and lesson resources based on the Common Core State Standards (CCSS). The researchers found that four 10-minute sessions in the TeachLivE simulator improved targeted teaching behaviors in the simulator, and that those improvements transferred into the teachers’ classroom settings.

In Phase II the research team used refined methods to examine the effects of four 10-minute sessions of 129 secondary science teachers working with newly developed high school avatars to compare two levels of professional development, lesson resources, and these same resources combined with virtual rehearsal in TeachLivE. Phase II teachers who took part in TeachLivE significantly increased their targeted behaviors compared to colleagues who had not taken part in computer simulation, and again, as in the Phase I study, the results transferred back to their real classroom settings. Results from both Phase I and Phase II research validate that, in the field of teacher education and simulation, professional learning in mixed-reality simulated classrooms can be effective in positively impacting teacher practice.
Using Virtual Rehearsal in TLE TeachLivE™ Mixed Reality Classroom Simulator to Determine the Effects on the Performance of Science Teachers: A Follow-up Study

The current administration has invested billions of dollars in education reform, and changes are underway. The White House’s 2011 Strategy for American Innovation authors stressed the need for better teacher preparation in science, technology engineering, and mathematics (STEM) subjects (National Economic Council, Council of Economic Advisers, and Office of Science and Technology Policy, 2011). Though student scores on the National Assessment of Educational Progress (NAEP) are improving in eighth grade, only 32% of eighth grade students reached or exceeded proficiency in science in 2011 (National Center for Education Statistics, 2011). The Common Core Standards Initiative (2011) aimed to increase content knowledge of students with the expectation that students will take on greater challenges in their academic careers and be better prepared for their careers or college coursework. While raising expectations for students, this shift in standards also has altered the way teachers are expected to approach education, by moving to an increasingly student-centered, inquiry-based approach to learning where students are engaged with “deep learning goals enabled by new pedagogies and accelerated by technology” (Fullan & Langworthy, 2013, p. 4). This need for deeper learning is reinforced in the National Research Council (NRC) Framework and Next Generation Science Standards (NGSS) underscoring the importance of students’ thorough understanding of core scientific ideas and the ability to discuss their ideas with others (Reiser, Berland, & Kenyon, 2012). These shifts in student expectations require a corresponding shift in teacher professional development. Teacher preparation programs need to prepare teachers for these new standards, but in-practice teachers need immediate retooling to prepare students to meet these standards, which are designed to result in better college and career options upon graduation. If teachers must undergo rigorous professional development (PD) to be ready to teach these standards, their training must be efficient and immediate with strong transference of knowledge (Windschitl, Thompson, Braaten, & Stroupe, 2012) to give students a quality understanding of these new standards in practice.

Reform in science instruction, beginning in the 1960s, has increased interest in a new scientific inquiry-focused teaching approach, which can be traced through the Benchmarks for Science Literacy (American Association for the Advancement of Science, 1993), the National Science Education Standards (National Research Council [NRC], 1996), along with Inquiry and the National Science Education Standards (NRC, 2000). The underlying philosophy is that “science practice involves doing something and learning something in such a way that the doing and learning cannot really be separated,” so the word practice becomes the operative word, a skill to be honed, much like riding a bike (Michaels, Shouse, & Schweingruber 2008, p. 34). In inquiry-based science, students are encouraged to take a hands-on, investigative approach to learning the functions of scientific concepts (Duschl, Schweingruber, & Shouse, 2007; Shouse & Schweingruber, 2008). Bybee (2002) indicates that students make cognitive gains and have a significantly improved understanding of the field of science when using inquiry-based methods. Yet, most science classroom discussions follow a pattern of initiation, response, and evaluation (I-R-E) most commonly linked to elementary class structure, but clearly observable throughout primary and secondary schools and even at the college level (Neal, 2008). The pattern is easy to recognize: the instructor asks a question, usually one to which he or she already knows the answer; the student responds, and subsequently receives feedback or constructive criticism.
Instructors use this pattern for various purposes – to remind students of information, spark a discussion, etc. (Neal, 2008).

The Common Core Standards point to a departure from this structure to explore problems and ideas that, in the IRE pattern, might not see the light of day (National Science Teachers Association, 2014; Waring, 2009). Furthermore, the exploration of what might be learned outside of the structure places the students’ education squarely in their hands, piquing their curiosity as they are able to practice discovering new knowledge on their own (Goodwin, 2007; Hawkins, 2007; Walsh 2002; Waring, 2009). The NRC Framework and NGSS focus not so much on memorization of facts on the scientific process but on a thorough understanding of core scientific ideas and the ability to discuss them with others (Reiser, Berland, & Kenyon, 2012). The transition from the typical I-R-E pattern to a more challenging and far-reaching exploration of science will require a conscious effort on the part of teachers directing student discourse, so that students can make sense of the material in a constructive way (Bacolor, Cook-Endres, Lee, & Allen, 2014).

**Retooling Teachers**

To engage students at new levels of thinking related to science, teachers need to demonstrate an array of teaching practices in their classrooms, and teacher preparation and PD should target the practices teachers find most challenging (Windschitl et al., 2012). In the *Measures of Effective Teaching* study, Kane and Staiger (2012) reported that teachers scored lowest for complex teaching skills such as questioning, discussion techniques, and communicating with students about content. Teaching Works (2014) analyzed core capabilities for teachers and developed a set of 19 high-leverage practices (HLPs) of which mastery will likely lead to increased advances in student learning. The HLPs are based on research linking particular practices to student achievement (Loewenberg Ball & Forzani, 2010). The Teaching Works’ HLPs span across content, teacher style, and setting, and include eliciting and interpreting student thinking, and providing oral feedback on students’ work (Loewenberg Ball, Sleep, Boerst, & Bass, 2009), both of which take place in an inquiry-based discussion. Danielson (2011) provided indicators for similar teaching capabilities, including higher-level questioning. Higher-level questions are defined as open-ended questions that allow students to use past experiences, prior knowledge, and previously learned content in order to create a well thought-out answer (i.e. question statements that begin with “How”, “What”, or “Why”) that relates to new content. For science teachers in particular, questioning appears to be the weakest element of instruction, and researchers have proposed a core set of instructional practices for science teachers, including questioning to elicit student thinking (Windschitl et al., 2012).

**Retooling Using Simulation**

As researchers and policy makers converge on a core set of high-quality teaching practices and corresponding professional learning opportunities for teachers to positively impact student outcomes, what are the best professional learning environments for teachers? Computer simulation is taking center stage as a next generation environment for teacher professional learning, allowing teachers to learn both pedagogical and content skills. Dieker, Straub, Hughes, Hynes, and Hardin (2014) described simulated environments for improving teacher practice, such as TeachLivE. TeachLivE is an immersive, mixed-reality classroom simulator that
combines real and virtual worlds to give users a sense of immersion and presence. Teachers interact with student-avatars in real time, holding authentic discussions on varied content areas. Over 50 universities and school district partners currently have TeachLivE classroom simulators for teacher professional learning, and TeachLivE is currently the only mixed-reality classroom simulator of its kind. Simulation can provide many educational experiences and opportunities that may not be available in real-world settings (Dieker, Rodriguez, Lignugaris/Kraft, Hynes, & Hughes, 2014; Nagendran, Pillat, Kavanaugh, Welch, & Hughes, 2014) and allow for safe rehearsal of skills until mastery. Simulated environments provide a safe place to practice teaching behaviors at an accelerated pace and receive rapid corrective feedback (Dieker et al., 2014a; McPherson, Tyler-Wood, McEnturff, & Peak, 2011). A research base is emerging, focusing on the use of TeachLivE with teachers and teacher candidates, and TeachLivE is currently in its fourth generation of student-avatars (see Andreasen & Haciomeroglu, 2009; Dawson & Lignugaris/Kraft, 2013; Elford, Carter, & Aronin, 2013; Elford, James, & Haynes-Smith, 2013; Straub, Dieker, Hynes, & Hughes, 2014; Vince Garland, Vasquez, & Pearl, 2012; Whitten, Enicks, Wallace, & Morgan, 2013).

Simulation provides unique capabilities in that teachers can receive just-in-time professional learning. Simulation that incorporates an after-action-review based on a theoretical model of performance mastery through feedback (e.g., Hattie & Timperley, 2007) has the potential to reduce discrepancies between current performance and a goal. Immediately after professional learning with simulation, teachers take part in an after-action-review process to engage in structured reflection and receive feedback on performance (Baird, Holland, & Deacon, 1999; Smith & Allen, 1994). Hattie and Timperley (2007) suggested that individuals set a goal for learning, receive data on actual performance, and revise learning goals based on performance. This immediate shaping of behaviors cannot happen in a real classroom, as real students would be made to wait while their teacher received corrective feedback. Avatar-students can be “paused” and wait patiently without losing valuable instructional time. Most importantly, unlike in real classrooms, teachers can re-enter the environment to fix instructional errors with student-avatars without affecting real students. Immersive virtual environments have the potential to revolutionize teacher professional learning, but more research is needed to establish the efficacy of the use of simulation for teacher education. One of the purposes of this research study is to evaluate the use of a classroom simulator with high fidelity (TeachLivE) to affect actual classroom instruction. In this study, we gave teachers an opportunity to practice their use of HLPs in TeachLivE and to evaluate the generalization of those practices to the traditional classroom setting.

Impact on Student Learning

Ultimately, our efforts toward a current and rigorous model of teacher PD should have a goal of positive impact on student learning. If shifts in teacher PD do not increase student learning, then the purpose of PD should be re-envisioned to incorporate new findings and practices in the field. Science students whose teachers use HLPs should show increases in student learning; therefore supporting teachers’ HLPs using TeachLivE should result in corresponding improvements in student learning. The core goal of this project is to support teachers as their practices shift to accommodate the new learning standards for students, so they can improve academically. The effects of Phase II professional development on student learning are provided in a companion report.
Theoretical Framework and Overarching Hypotheses

Our work in computer simulation is grounded in Brown, Collins, and Duguid’s (1989) theory of situated cognition, asserting, “what is learned cannot be separated from how it is learned and used” (p. 88). We believe learning occurs in contextually meaningful settings (Dieker et al., 2014) and, consequently, we created a contextualized simulation activity that provided learners with the opportunity to practice HLPs with student-avatars. Our theory of action arises from examining the critical features of professional learning for teachers that are related to increased student outcomes (e.g., active learning opportunities based on specific teaching practices, such as HLPs). Based on results from Phase I of a national research study investigating simulation in middle school mathematics classrooms, coupled with finding from earlier studies related to using virtual environments for teacher preparation (see Straub Dieker, Hynes, & Hughes, 2014), our overarching hypothesis is that teachers who engage in virtual environment simulations in high school science, specifically biology, will improve their practice in the simulator and this practice will transfer back to their classroom. Specifically, we hypothesized that four 10-minute sessions of virtual rehearsal (i.e., practicing the same lesson and HLPs in TeachLivE) would significantly increase teachers’ frequency of open-ended questions and content-related affirmation to students in both simulated and real classroom instruction, compared to their colleagues who did not engage in TeachLivE.

Changes in Research and Context from Phase I to Phase II

This national research study spanned two years, and in Phase II, methods were revised to incorporate findings from Phase I as well as test new technology components. While the same overarching hypotheses and questions framed the research, innovations in technology enabled researchers to utilize new high school avatars and computer-based feedback. In Phase II a new set of avatars was developed specifically to use TeachLivE with high school versus middle school teachers. The high school avatars used in Phase II were the same characters with the names and personality types from Phase I, but the nuances of their behaviors and look were that of high school students (see Figure 1).

Evolution of System to Include High School Avatars

The creation of the high school version of our five virtual students was an opportunity to improve our processes, both in the artistic development and in the pipeline that brings the artistic assets into the AMITIES Framework (the system that operates TeachLivE). The artistic goal was to create models that could be easily adapted to accommodate different clothing, hairstyles and other cosmetic changes. The pipeline enhancements included automating the process of exporting the artistic assets (models, textures, animations, etc.) so that naming and folder structure conventions are adhered to in a manner that guarantees ease of importing these assets into the framework that provides interaction paradigms, game mechanics, and visual rendering.

The character modeling process employed for the high school avatars is significantly different and more scalable than the process used to produce the middle school avatars. Although both are based on skeletal models, which are needed to support animation, after that there is a
major improvement from the latter to the former. In the case of the middle school students, the clothing, accessories and hair are all part of their “skins” and are not removable, because originally it was determined that they did not need to be changed. This convenience greatly reduced modeling time but created a set of “one-off” models. In contrast, each high school avatar has clothing, accessories and hair that are independent of his or her body (skin). This change makes it possible to personalize appearances to match the cultures of different clients or even to have dress-up or dress-down days.

Beyond the above issues of clothing serving as skin, each middle school student’s desk and chair are actually a part of the student’s model. This means that the student cannot be separated from the desk and so cannot walk to the whiteboard. It even means that removal of a student from the classroom results in removal of his or her desk and chair. The high school students, in contrast, are independent of environmental objects and so can be moved from chair to chair, walk to the whiteboard, fall to the floor or even be replaced by other students. Figure 2 shows Kevin (student at far left) replaced by a new student, Bailey, and the empty desk between Sean and Maria occupied by another new student, Martin, who has just gone to the whiteboard. These changes were more about the commercialization and building a pipeline process for avatars, but also added a new level of movement and authenticity to the study in Phase II.
Evolution of Research Tools, Data Collection and Learning Evaluation

The method for delivering feedback also was changed substantially in Phase II based up on lessons from the study. While the same type of information was shared with teachers after their PD sessions in TeachLivE, the format changed and results were no longer delivered on paper by a human facilitator, but by a computer on a screen immediately after the simulation. Anecdotal evidence from Phase I of the study yielded concerns from teachers that data presented during the after-action-review process was not valid, because researchers had scored teachers incorrectly. A search of literature indicated mixed findings related to computer versus human-delivered feedback. Differential performance resulted when participants received feedback from a computerized source (Earley, 1988) and participants were more likely to seek feedback from a computer than from another person (Kluger, 1993). However, supervisor-mediated feedback has been associated with higher levels of perceived fairness than computer-mediated feedback (Alder & Ambrose, 2005). In light of mixed findings related to format for delivering feedback, Hattie and Timperley’s (2007) meta-analysis of effect sizes indicated stronger effects for computerized feedback than non-computerized feedback. This finding guided the decision to deliver performance feedback from a computer in Phase II of the research on TeachLivE.

Revisions also were made to data collection methods and tools for both teachers and students. Teachers in Phase I were not required to teach a uniform lesson at all sites during classroom observations, and this decision resulted in a wide variance among teachers’ lesson formats and content. However, in Phase II, teachers all taught the same lesson, and this change was a significant improvement in limiting construct irrelevant variances while attempting to unify observable behaviors that guided teachers’ philosophical approach to teaching. The lesson
plans in Phase II were explicitly designed to enhance science literacy and aligned with Disciplinary Core Ideas and Cross-cutting Concepts from the Next Generation Science Standards, as well as the Common Core Standards for Literacy in Science. Lessons in Phase II incorporated Literacy Design Collaborative (LDC) mini-tasks – small, scorable assignments that address a particular literacy skill that a teacher has selected to target based on assessing the needs of students (LDC, 2015). Concept maps were embedded as mini-tasks just before and after each lesson. The lesson plans, based on the 5E Instructional Model, were validated and field-tested in high school Biology classrooms as part of a larger module from the NIH Curriculum Supplement Series (National Institutes of Health, 2005). All of these efforts were aimed at creating a uniform basis for observing teachers and comparing results of PD to support an inquiry-based approach to teaching. Also, while teachers were still observed through the lens of HLPs, researchers refocused to accommodate new types of discussion questions, including questions that challenged students to think at the highest levels of Bloom’s taxonomy (Anderson et al., 2001; Bloom, 1956), creating new ideas and knowledge. As such, student data collection tools were revised to capture the number of ideas students had related to content before and after their teachers received TeachLivE.

Finally, more emphasis was placed on the collection and evaluation of student learning outcomes. Because the focus of this report is on teacher performance, the results as they relate to student learning will be examined in a supplemental report; however, a general overview of the changes from Phase I to II follows. In Phase I of the study, no significant difference was found between student learning based on intervention groups using a 10-item instrument based on the NAEP for eighth grade mathematics. To tease out possible impact on student learning, the researchers extended their data collection in Phase II, by adding an additional instrument. Data were collected on: (1) a 10-item multiple choice assessment based on the College Board (2013) SAT Subject Test in Biology, similar to the format in Phase I and (2) a concept map administered pre- and post-lesson at each observation in order to collect data about students’ number of ideas generated related to the content. The aim was to establish if students not only increased content knowledge as demonstrated on a general measure, but also increased their number of ideas related to the content in response to teachers’ increased questioning as supported by TeachLivE. Specific procedures for collecting and analyzing the large volume of student data generated in this project are presented in detail in an upcoming complementary report.

Research Questions

Phase II of this research study focused on replicating findings from Phase I. In the TeachLivE classroom simulator, we had the following research questions:

**Research Question 1:** Are there differences in performance over four 10-minute sessions of TeachLivE on teacher practice (i.e., questions) during a 10-minute simulation when performance feedback is given?

The ability to receive feedback in a timely, objective manner is a perceived benefit of simulation, so it was important to investigate the impact of providing after-action-review of teacher performance. We hypothesized withholding feedback on a specific teacher practice (i.e., frequency of content-related affirmation) would result in no difference in performance:

**Research Question 2:** Are there differences in performance over four 10-minute sessions of
TeachLivE on teacher practice (i.e., frequency of content-related affirmation) during a 10-minute simulation when no performance feedback is given?

In order to be valuable, change in teacher practice should occur not only in a virtual classroom, but also in a real classroom with students present. Our next research questions were designed to evaluate teacher performance in the classroom.

*Research Question 3:* What are the effects on teaching practice (i.e., questions) in a classroom after four 10-minute sessions of TeachLivE?

*Research Question 4:* What are the effects on teaching practice (i.e., frequency of content-related affirmation) in a classroom after four 10-minute sessions of TeachLivE?

In terms of social validity of the intervention, it is important to know whether or not teachers perceived that the avatars were realistic and the PD was of value (Wolf, 1978). Our next research question was designed to collect information on teachers’ perceptions related to TeachLivE.

*Research Question 5:* What are the perceptions of practicing teachers related to presence and perceptions of preparedness after completing TeachLivE?

**Method**

**Participant Characteristics**

Data analyzed in this study were collected during the second year of a three-year project at 11 separate research locations comprised of university and school district partners. Participants were the primary teachers of record and were primarily high school biology teachers. No restrictions were made based on education level of a teacher, number of years teaching, level of class taught, subject area within science taught, or any other demographic characteristics. Overall, 104 teachers completed the study. Demographic data for all participating teachers are presented in Table 1.
Table 1. Teacher Demographic Data.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Comparison ((n = 51))</th>
<th>TeachLivE ((n = 53))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professional licensure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>6 (12)</td>
<td>48 (91)</td>
</tr>
<tr>
<td>No</td>
<td>40 (78)</td>
<td>2 (4)</td>
</tr>
<tr>
<td>No response</td>
<td>5 (10)</td>
<td>3 (6)</td>
</tr>
<tr>
<td>If licensed, is license in science?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>35 (69)</td>
<td>37 (70)</td>
</tr>
<tr>
<td>No</td>
<td>1 (2)</td>
<td>4 (8)</td>
</tr>
<tr>
<td>No response</td>
<td>15 (29)</td>
<td>12 (23)</td>
</tr>
<tr>
<td>Area of certification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grades 6-12</td>
<td>26 (51)</td>
<td>26 (49)</td>
</tr>
<tr>
<td>Other</td>
<td>6 (12)</td>
<td>8 (15)</td>
</tr>
<tr>
<td>No response</td>
<td>19 (37)</td>
<td>19 (36)</td>
</tr>
<tr>
<td>Highest academic level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bachelor's</td>
<td>18 (35)</td>
<td>27 (51)</td>
</tr>
<tr>
<td>Master's</td>
<td>27 (53)</td>
<td>21 (40)</td>
</tr>
<tr>
<td>Doctorate</td>
<td>2 (4)</td>
<td>1 (2)</td>
</tr>
<tr>
<td>No response</td>
<td>4 (8)</td>
<td>4 (8)</td>
</tr>
<tr>
<td>Area of masters degree</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biology</td>
<td>7 (14)</td>
<td>6 (11)</td>
</tr>
<tr>
<td>Other</td>
<td>19 (37)</td>
<td>18 (34)</td>
</tr>
<tr>
<td>Not applicable or No response</td>
<td>25 (49)</td>
<td>29 (55)</td>
</tr>
<tr>
<td>Years teaching science</td>
<td></td>
<td></td>
</tr>
<tr>
<td>One year</td>
<td>3 (6)</td>
<td>7 (13)</td>
</tr>
<tr>
<td>Two years</td>
<td>10 (20)</td>
<td>5 (9)</td>
</tr>
<tr>
<td>Three years</td>
<td>1 (2)</td>
<td>3 (6)</td>
</tr>
<tr>
<td>Four years</td>
<td>3 (6)</td>
<td>4 (8)</td>
</tr>
<tr>
<td>5-10 years</td>
<td>13 (25)</td>
<td>12 (23)</td>
</tr>
<tr>
<td>More than 10 years</td>
<td>17 (33)</td>
<td>17 (32)</td>
</tr>
<tr>
<td>No response</td>
<td>4 (8)</td>
<td>5 (9)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-29</td>
<td>14 (27)</td>
<td>19 (36)</td>
</tr>
<tr>
<td>30-39</td>
<td>12 (24)</td>
<td>14 (26)</td>
</tr>
<tr>
<td>40-49</td>
<td>13 (25)</td>
<td>8 (15)</td>
</tr>
<tr>
<td>50 or above</td>
<td>8 (16)</td>
<td>9 (17)</td>
</tr>
<tr>
<td>No response</td>
<td>4 (8)</td>
<td>3 (6)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>10 (20)</td>
<td>9 (17)</td>
</tr>
<tr>
<td>Female</td>
<td>37 (73)</td>
<td>41 (77)</td>
</tr>
<tr>
<td>No response</td>
<td>4 (8)</td>
<td>3 (6)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>5 (10)</td>
<td>3 (6)</td>
</tr>
<tr>
<td>Black</td>
<td>6 (12)</td>
<td>4 (8)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>4 (8)</td>
<td>3 (6)</td>
</tr>
<tr>
<td>White</td>
<td>27 (53)</td>
<td>39 (74)</td>
</tr>
<tr>
<td>Other</td>
<td>5 (10)</td>
<td>1 (2)</td>
</tr>
<tr>
<td>No response</td>
<td>4 (8)</td>
<td>3 (6)</td>
</tr>
</tbody>
</table>

Teachers indicated having taught the following grade levels:

<table>
<thead>
<tr>
<th>Grade Level</th>
<th>Control Group</th>
<th>TeachLivE Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>K - grade 5</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Grades 6 - 8</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>Grades 9 - 12</td>
<td>46</td>
<td>43</td>
</tr>
</tbody>
</table>
Sampling Procedures

Participants were identified via a convenience-sampling plan. Approximately 150 high school biology teachers were initially recruited across 11 separate research locations. All teachers agreed to teach a lesson plan based on science content. At each partnership site, teachers were self-nominated or nominated by their supervisors with the intent of receiving lesson resources. Of those teachers, 129 completed a half-hour orientation and began participation in the research project, resulting in about 86% participation of the sample approached. Participation was entirely voluntary with minimal to no compensation provided.

Data were collected in two settings: the teachers’ real classrooms and in the classroom simulator. Teachers were observed in secondary classrooms located in 11 sites across the following states: Florida, Georgia, Idaho, Illinois, Louisiana, Maryland, New York, Texas, and Washington, D.C. School settings ranged over urban, suburban, and rural with public or private enrollment. Classroom simulators were located near teachers’ classrooms at university or school district partner sites.

Teachers voluntarily participated and in most cases were motivated to do so because of access to professional learning resources. The research team leading the study did not offer direct compensation; however, each site was free to offer incentives for participation based on local conventions. All incentives were valued at less than $200 per teacher and came in the form of a stipend or points awarded for PD supplied by the district. For recruiting purposes, the project was described as Next Generation PD, including the potential of working in a computer-simulated environment with innovative technology. Institutional review boards at each site and within each school district examined and approved all procedures, granting permission to conduct research.

Sample Size, Power, and Attrition

The intended sample size was 200 participants, with the goal of targeting high school biology teachers only; however, respondent participation numbers were low, so teaching content area was expanded to include all teachers at secondary level. Consequently, the bulk of teachers taught in science content areas, with a minority of teachers in subjects such as English and special education. As with Phase I, multiple districts were approached in Phase II, but chose not to participate because of concerns that the professional learning would duplicate their own activities or conflict with their district initiatives. Phase II research methods required that all teachers deliver the same lesson plan for Observation 1 and another lesson for Observation 2, which resulted in many logistical challenges for districts and participants who followed pre-determined curriculum pacing guides and for many sites this study occurred during a very difficult winter that greatly impacted the ability to retain all participants. While the content and format of the lesson afforded flexibility, many teachers were reticent to carve out valuable instructional time, and did not participate in the study or were not longer available due to non-ending cancellation of days for schools in the north and northeast due to weather related events.

Power analysis for sample size. In a review of literature, the only similar study identified using a large group design for practicing teachers’ PD in a classroom simulator was research conducted in Phase I. Effect sizes were reviewed to offer an estimate of the desirable
effect size, yielding a range from small to large of $\eta^2_p = .025$ to .149. An a priori power analysis was conducted (Cohen, 1988) using a medium sized effect (0.25). Power analysis for an F-test Analysis of Variance (ANOVA) within-between interactions resulted in a total sample size of 48 participants to have 80% power for detecting a medium sized effect (0.25) when employing a 0.10 criterion of statistical significance. As with Phase I, a 0.10 criterion was selected due to the new research in the field with a low risk to humans; therefore, a larger Type Two error was acceptable in considering the overall findings. The projected number of participants was 200, based on funding allocated for the research project. The anticipated number of participants exceeded the suggested number of 48 participants for a medium sized power effect.

**Research Design**

The research design was a randomized controlled trial, consisting of two groups of teachers measured pre-post in the classroom, half of whom also were measured four times in the classroom simulator. The random assignment procedure took place at all 11 partnership sites, resulting in two experimental groups.

**Interventions**

Teachers were assigned to one of two groups, and both groups received the same lesson plan resources. The lesson plans were designed to enhance science literacy and aligned with Disciplinary Core Ideas and Cross-cutting Concepts from the Next Generation Science Standards, as well as the Common Core Standards for Literacy in Science. The lesson plans were based on the 5E Instructional Model and were validated and field-tested in high school biology classrooms as part of a larger module from the NIH Curriculum Supplement Series “Using Technology to Study Cellular and Molecular Biology” (National Institutes of Health, 2005). Lesson 1, entitled “What is Technology?” was the basis for the first observation, while Lesson 2, entitled “Modeling Issues” was the basis for the second observation (see Appendices A and B). In Phase I, teachers had taught a variety of content and lesson formats, resulting in construct irrelevant variance. As a means of removing potentially confounding variables, all teachers taught Lesson 1 at pre-treatment observation and Lesson 2 at post-treatment observation. In this way, lesson, content, and format were standardized at the 11 sites across the country. Lesson 1 also had an accompanying video model of a teacher teaching Lesson 1 using HLPs in a real classroom.

Both lessons had the same structure and parallel content. The lesson began with a concept map, which served as a mini-task literacy activity (LDC, 2015). Teachers showed a model concept map and explained the components to students (see Appendix C), then they gave their students five minutes to generate as many ideas as they could related to the prompt in the center of the map. The prompts corresponded directly to the lesson with the purpose of eliciting student thinking related to the content just before and after the lesson. During the lesson, teachers facilitated a whole class discussion focused on interpreting, inferring and deducing from data, and integrating information to form conclusions. At the close of the discussion, students completed an identical concept map, responding to the same prompt. In this way, each student created two concept maps on the same topic and learning was assessed by changes between concept maps.
Students responded to the 10-item curriculum-based assessment prior to beginning Lesson 1 and after completing Lesson 2. All teachers received the lesson plans and accompanying video resources via email after orientation and prior to the first observation. Two observations were scheduled at least three weeks apart (one for each lesson) and treatment took place in between observations.

Group 1 teachers served as a comparison group and received lesson plan resources while Group 2 teachers received the same lesson plan resources plus four, 10-minute sessions of TeachLivE. See Figure 3 for an overview of both treatment groups.

**Group 1: Comparison.** Comparison teachers received Lessons 1 and 2 and the accompanying video for Lesson 1 via email. They were given no other intervention as a course of this study, but did receive any PD provided by their district throughout the course of the school year. They taught Lesson 1 at pre-treatment observation and Lesson 2 at post-treatment observation.

**Group 2: Simulation.** Simulation teachers received both lessons and resources (like the comparisons teachers), as well as four 10-minute virtual rehearsal sessions in the TeachLivE classroom simulator. In the simulator, teachers attended individually and interfaced with a computer-generated, animated student population of five high school avatars digitally controlled by a professional who enacted a highly interactive, authentic simulation of a high school classroom. The software is programmed to react to certain commands of the teacher and the interactor, with the purpose of increasing the teacher’s aptitude in the classroom. Classroom simulators at 10 client sites across the country were connected via a secure server to the Synthetic Reality Laboratory at the University of Central Florida, which served as the central distribution point for TeachLivE and provided fidelity of treatment that all sessions were controlled at the primary research site. For operation at the teacher client sites, the simulator required a computer with TeachLivE software, large display monitor, webcam, lavaliere microphone, speakers, system for tracking movement, and an Internet connection. A session facilitator, trained on how to use the software and enact the research procedures, facilitated the
sessions and collected the data. At the server where the interactor was located, a computer with TeachLivE software, monitor, and motion tracking devices were needed to operate the system. The teachers experienced computer-simulated classroom activities with the student-avatars as they would with human students in a traditional classroom. Visits to the simulator took place over the course of three to four weeks following the first classroom observation.

Simulation teachers participated in one 10-minute session to orient them to the simulation system. Data were not collected during the orientation session, as users were not teaching content but interacting with the student-avatars with the objective of learning about their class. After orientation, teachers received four 10-minute sessions (sessions 1 through 4) to take part in virtual rehearsal (i.e., targeted practicing of a skill in a virtual environment), with data on targeted behaviors gathered during each session. After orientation, teachers typically took part in two 10-minute sessions and returned within a month for another two 10-minute sessions. Teachers were instructed to follow the first 10-minutes of Lesson 1 (beginning immediately after students had completed the concept maps) for each session.

**After-Action-Review.** At the close of each session, teachers took part in an after-action-review of their performance led by a facilitator using a digital chart displayed on a large video monitor (approximately 60”). This process of computerized feedback was a complete departure from Phase I methods, in which facilitators had written performance data on a piece of paper and handed it to participants. Methods were refined based on participant feedback from Phase I, so that results were now displayed digitally. After-action-review consisted of four parts: (1) teachers were presented with frequency of observed behaviors (i.e., closed-ended questions (CE), open-ended questions (OE), and open-ended plus questions (OE+): detailed definitions of each behavior is provided below) during the session verbally and on a large display; (2) teachers read examples of question types CE, OE, and OE+ on a large display; (3) teachers were asked to set a goal for their performance in the next session on OE questions; and (4) just prior to commencing the session, teachers stated their target goal of OE questions. Performance goals were not set for OE+ questions, because they are part of the larger category of OE questions. Upon completion of the after-action-review, teachers returned to the simulation for another session. Performance on content-related affirmation (CRA) was withheld from participants, in order to evaluate simulation with no after-action-review. Table 2 describes the after-action-review process.

Table 2. After-Action-Review Process.

<table>
<thead>
<tr>
<th>Step</th>
<th>CE</th>
<th>OE</th>
<th>OE+</th>
<th>CRA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Data display of observed behaviors</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2: Read examples of question types</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>2: Teachers set performance goal</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4: Teachers stated performance goal</td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>
Data were collected on a variety of measures from teachers and their students, including qualitative and quantitative measures. To ensure that all 11 sites nationally had high reliability in data collection, researchers employed methods to enhance the quality of measurements. As with Phase I, due to the national nature of the study, training observers across the country presented challenges in terms of training and reliability of observations. Therefore, all data collectors were trained online using a combination of asynchronous assessment and synchronous data collection training on the constructs (e.g., Danielson sub-constructs and HLPs) and methods (e.g., frequency counts during rotating intervals as described above) for data collection. Data collectors used the asynchronous online modules to demonstrate proficiency with the content of observations. Each practice was defined and a case example was provided. Observers had to pass a multiple-choice content assessment with 90% accuracy for the asynchronous portion of the training. The synchronous online training was enhanced from Phase I, and consisted of a series of rigorous activities delivered via a video conferencing platform that exposed observers to operational definitions and required the collection of frequency counts in real time while watching a video online as a group to simulate classroom observations. Each observer was checked for reliability during the online training and required to complete a synchronous online activity with 90% accuracy.

**Data Collection.** As with Phase I, in Phase II quantitative and qualitative observations on the Teacher Practice Observation Tool (TPOT; see Appendix D) were used to collect data on teachers in their classrooms pre- and post- treatment. See Table 3 for an overview of data sources.

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Individuals</th>
<th>Instrument</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher demographics</td>
<td>Teachers</td>
<td>Professional Development Questionnaire</td>
</tr>
<tr>
<td>Teaching practice in TeachLivE classroom simulator</td>
<td>Teachers</td>
<td>ReflectLivE After-Action-Review System</td>
</tr>
<tr>
<td>Teacher perceptions of TeachLivE experience</td>
<td>Teachers</td>
<td>TeachLivE Presence Questionnaire</td>
</tr>
<tr>
<td>Teacher perceptions of preparation after TeachLivE</td>
<td>Teachers</td>
<td>TeachLivE Perceptions Questionnaire</td>
</tr>
<tr>
<td>Teaching practice in classroom</td>
<td>Teachers</td>
<td>Teacher Practice Observation Tool</td>
</tr>
<tr>
<td>Student academic performance</td>
<td>Students</td>
<td>Curriculum-based measure</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cognitive Maps Scores</td>
</tr>
<tr>
<td>Student demographics</td>
<td>Students</td>
<td>Cross-reference Demographic Sheet</td>
</tr>
</tbody>
</table>

Teachers responded to demographic questions prior to treatment. During classroom simulator observations, data were collected on the frequency of HLPs determined to increase the likelihood that these teaching behaviors would have a positive effect on students’ learning outcomes (Teaching Works, 2014). For the teachers who experienced the classroom simulator,
data also were collected on their sense of presence and preparedness after the four sessions of virtual rehearsal. Observers focused simultaneously on two variables: questioning and feedback throughout the 10-minute session. In the real classroom, observers collected data on frequency of HLPs as well as on modified sub-constructs from the 2011 Danielson Framework for Teaching Evaluation Instrument (Danielson, 2011). Qualitative field notes also were taken regarding teacher practice during observations. Data were collected in nine-minute intervals (five intervals for a total of 45-minutes), rotating across constructs such that observers focused on one construct at a time, except in the case of Affirmation and Student Talk variables (see Appendix D). Table 4 describes the variables observed in the real classroom and the classroom simulator.

Table 4. Variables Observed in the Classroom and Classroom Simulator.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Classroom</th>
<th>Classroom Simulator</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>One period 45 minutes (2 observations)</td>
<td>10-minute session (4 observations)</td>
</tr>
<tr>
<td>High-leverage Practices</td>
<td>Questioning (Two Types) Content-related Affirmation</td>
<td>Questioning (Two Types) Content-related Affirmation</td>
</tr>
<tr>
<td>Type of Data Collected</td>
<td>Questioning (3-minute intervals) Content-related Affirmation</td>
<td>Frequency or Percentage per 10-minute session</td>
</tr>
<tr>
<td>Sub-constructs from 2011</td>
<td>8 sub-constructs modified</td>
<td></td>
</tr>
<tr>
<td>Danielson Framework for</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teaching</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of Data Collected</td>
<td>-Sum of observer ratings at the end of the observation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-Structured protocol of field notes for 2 minutes every 9 minutes</td>
<td></td>
</tr>
</tbody>
</table>

**High-leverage practices.** Using research from the Measures of Effective Teaching project, descriptions of HLPs, and other empirically-based research in the field, operational definitions for observation were created (See Appendix E for Operational Definitions Quick Reference Sheet used by observers).

Data were collected on teachers’ frequency and type of eliciting and interpreting individual students’ thinking (HLP #3). These same measures occurred in the classroom simulator and pre-post in the real classroom. Each simulator session lasted 10 minutes so the
frequency and type of instances for each behavior were noted. In the teachers’ classrooms, the prescribed lesson was 45 minutes in length:

- closed-ended questions (CE): content questions that have restricted parameters, expecting one possible response as its only acceptable answer; constrains a student’s response, such as test questions, yes–no questions and forced choice questions.

- open-ended questions (OE): content questions to which a number of different answers would be acceptable; content questions that have no parameters and do not constrain student’s response. Open-ended plus questions (see definition below) were included in this category for data collection purposes.

- open-ended plus questions (OE+): content questions that ask a student to extend, produce, or combine ideas to generate new ideas (related to Bloom’s highest cognitive domain – creating). Open-ended plus question were included within the open-ended questions category.

Observers also looked for content-related affirmation (e.g., feedback) that teachers gave students. Effective feedback is specific, not overwhelming in scope, focused on the academic task, and supports students’ perceptions of their own capability (HLP #12 and 16). As with questions, in the classroom simulator, frequency data of CRA were collected in each 10-minute session. In the teachers’ real classrooms, the prescribed lesson was 45 minutes in length. For the purposes of this study, the focus was on affirmation related to content only and it was defined as:

- content-related affirmation (CRA): teacher’s positive verbal affirmation about what a student or group of students did or said related to content in a single episode within class (multiple statements about the same episode count as one occurrence of affirmation).

**Sub-constructs from 2011 Danielson Framework for Teaching.** As with Phase I, eight sub-constructs that correlated with student achievement from the 2011 Danielson Framework for Teaching Evaluation Instrument (Measures of Effective Teaching Project, 2010) were identified. Key words from Danielson’s indicators were chosen to create an abbreviated version for classroom observations, and combined with the collection of frequency data in relation to describe/explain questions, specific feedback, and wait time. Danielson’s four levels of performance (i.e., unsatisfactory, basic, proficient, distinguished) were the basis for a four-point scale for each sub-construct: establishing a culture for learning, engaging students in learning, managing student behavior, managing classroom procedures, communicating with students, using questioning and discussion techniques, creating an environment of respect and rapport, and using assessment in instruction. Further, qualitative data were collected during the classroom observation on each sub-construct listed above using a field notes method. See Appendix D for the TPOT that includes each sub-construct and associated scale. For a description of TPOT development, see Straub, et al. (2014). Reliability estimates related to each variable are provided in the results section.

**ReflectLivE: After-Action-Review System.** During each TeachLivE session, the teacher’s virtual rehearsal was transmitted via secure Skype video and audio connection. The transmissions were recorded and coded for pedagogical strategy analysis using ReflectLivE
software. ReflectLivE is a video tagging software integrated with the TeachLivE classroom simulator that records sessions, compresses the video to a smaller format, storing all data (video and tags) on the observer’s workstation. These data can then be sent over a secure network to be stored at the originating research site computer containing the TeachLivE software. During each session, videos were tagged for frequency of questions and content-related affirmation. As with Phase I, a beta version of ReflectLivE was used in Phase II of the project, and brought about intermittent issues with recording and exporting of tags, so data also were collected using a paper and pencil backup to maintain integrity. A new tool was developed for Phase II to automatically gather data related to student talk time in the simulator. As the avatars mouths moved, the system logged the time spent “talking” and a percentage of student talk to total time in the session was automatically calculated. Because this tool was in beta version, the measurements were not yet validated; moreover, observers were not always consistent in when they started the recordings. As a result of these inconsistencies, these data were not analyzed.

**TeachLivE questionnaires.** Each teacher that entered the classroom simulator was administered two researcher-created questionnaires in order to collect data on the social validity of the intervention:

- TeachLivE Presence Questionnaire (Hayes, Hardin, & Hughes, 2013): Teachers responded to questions about their simulation experience related to suspension of disbelief, presence, fidelity, and immersion.
- TeachLivE Perceptions Questionnaire (Hayes, Hardin, & Hughes, 2013): Teachers also responded to items about how virtual rehearsal in the classroom simulator prepared them for teaching in their own classrooms.

**Results**

At the beginning of the research study, 129 teachers completed orientation and were randomly assigned to four groups in a randomized controlled trial design (see Table 5); however, teacher requests to change treatment groups necessitated a modification resulting in the final quasi-experiment design (Step 1 of participant flow through the quasi-experiment). Once scheduling began, seven teachers requested to be removed from the study or to be changed to a different treatment group due to scheduling restrictions (Step 2 of participant flow). Four teachers wanted to continue participation, but could not complete the activities due to prior commitments, so they were moved from the simulation to the comparison group. Three teachers requested to be moved from comparison to simulation to increase their level of treatment to potentially receive benefits of simulation. While changes in treatment group did violate random assignment procedures, all changes occurred prior to interventions so that no teachers received a partial intervention and then switched to another group midway through an intervention. Teachers attended events individually; therefore, group assignment could occur prior to the intervention. Table 5 outlines the participant flow through each stage of the study. First, teachers were randomly assigned to treatment groups; then teachers made requests to be changed from their group and some teachers were lost to attrition (Step 3 of participant flow), resulting in the final number of teachers per treatment group prior to beginning treatment.
Table 5. Participant Flow through the Quasi-Experiment

<table>
<thead>
<tr>
<th>Step</th>
<th>Comparison</th>
<th>Simulation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1: Teachers oriented and randomly assigned</td>
<td>56</td>
<td>73</td>
<td>129</td>
</tr>
<tr>
<td>Step 2: Changes in groups per teacher request</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>From Comparison to Simulation</td>
<td>3</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>From Simulation to Comparison</td>
<td>N/A</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Step 3: Teachers lost to attrition</td>
<td>-6</td>
<td>-19</td>
<td>-25</td>
</tr>
<tr>
<td>Step 4: Final number of teachers per group</td>
<td>51</td>
<td>53</td>
<td>104</td>
</tr>
</tbody>
</table>

**Treatment Fidelity**

Fidelity checks were in place throughout the study. All teachers received the lesson plan in digital format, as evidenced by a checklist of teacher contact information at each site. The professional who controlled the five computer-generated high school avatars was trained to follow five distinct patterns of behavior aligned to common student perceptions related to Lesson 1 content, and to maintain consistent, authentic responses through sessions that would reset with each interaction. However, fidelity of implementation data were not collected on the avatars’ performance to measure consistency across participants during treatment. During the TeachLivE sessions, the facilitator followed a detailed procedural checklist to turn on and operate the software for the simulation, ensuring fidelity of implementation.

**Teacher Results**

Teaching practices were defined on five distinct dimensions pre- and post-intervention: (a) close-ended questions (CE), (b) open-ended questions (OE), (c) open-ended plus questions (OE+), (d) content-related affirmation (CRA), and (e) summary score on the TPOT (TPOT Sum). Maxwell’s (2001) recommendation of moderate correlation (0.3 – 0.7) was used as a threshold for all variables to determine if it was appropriate to conduct a multivariate analysis of variance. Content-related affirmation was excluded from the analysis, because the researchers predicted no significant findings. In the case of the variables under investigation, the majority did not meet correlation thresholds, so analysis of variance (ANOVA) tests were more appropriate. See Table 6 for correlations of dependent variables.
Specific statistical tests used and variables under consideration are described in detail in the results section. The results are divided by classroom simulator data and classroom data, and then further subdivided by research question.

### Classroom Simulator Results

**Research Question 1: Differences in performance over time with simulation and performance feedback.** To examine performance of teachers over four 10-minute sessions, a within-subjects ANOVA was performed. Time (four sessions) was cast as a within-subjects factor with dependent variables of CE for question 1.1, OE for question 1.2, and OE+ for question 1.3. One observer collected data during all of the TeachLivE sessions. Due to the novel nature of the intervention (e.g., dearth of group design research identified on simulation in teacher education), an alpha level of .10 was established to judge statistical significance. Partial eta squared was used to interpret effect size rather than eta squared because a multifactor design was used (Pierce, Block, & Aguinis, 2004) to be able to compare effects across different factorial designs used in the study (Levine & Hullet, 2002).

**Question 1.1: CE questions in simulator.** After each session, teachers were presented with data verbally and on a large display on CE questions, but no performance goals were set for subsequent sessions. Analysis was conducted with a within-subjects design ANOVA. Mauchly’s test of sphericity indicated that the assumption of sphericity had not been violated, \( \chi^2(5) = 6.772, p = .238 \). Results indicated a significant time effect \((F(3,87) = 3.710, p = .015, \eta^2_p = .113)\). Pallant (2007) recommends interpreting partial eta squared using Cohen’s (1988) guidelines for eta squared effect size: small (.01), medium (.06), or large (.14). Mean scores decreased significantly at each session, which was expected, because: (a) although feedback on performance was given, no performance goals were set for CE questions, and (b) teachers were focusing on increasing an opposing behavior of OE questions. See Table 7 for mean CE questions across 10-minute TeachLivE sessions.
Table 7. Mean CE Questions across 10-minute TeachLivE Sessions

<table>
<thead>
<tr>
<th>TeachLivE Sessions</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>10.23 (6.1)</td>
<td>9.6 (6.2)</td>
<td>8.7 (3.9)</td>
</tr>
</tbody>
</table>

Figure 4. Frequency of CE Questions over Time.

**Question 1.2: OE questions in simulator.** Teachers were primarily attempting to increase their frequency of OE questions. Analysis was conducted with a within-subjects design ANOVA. Mauchly’s test of sphericity indicated that the assumption of sphericity was violated, \( X^2(5) = 93.798, p = .000 \). Epsilon (\( \varepsilon \)) was 0.387, as calculated according to Greenhouse and Geisser (1959), and was used to correct the ANOVA. Results indicated no significant time effect \( (F(1.162, 33.694) = .320, p = .609, \eta^2_p = .011) \). Mean scores are displayed in Table 8 and frequency over time is displayed in Figure 5.

Table 8. Mean OE Questions across 10-minute TeachLivE Sessions

<table>
<thead>
<tr>
<th>TeachLivE Sessions</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>16.65 (17.7)</td>
<td>15.37 (6.4)</td>
<td>17.5 (7.4)</td>
</tr>
</tbody>
</table>
Question 1.3: OE+ questions in simulator. As a specific subset of OE questions, OE+ questions also were measured. After each session, teachers were presented with OE+ data verbally and on a large display, and a definition for OE+ questions was read aloud; however, performance goals were not set for subsequent sessions because OE+ questions are part of a larger category of OE questions. Analysis was conducted with a within-subjects design ANOVA. Mauchly’s test of sphericity indicated that the assumption of sphericity was violated, $X^2(5) = 86.024, p = .000$. Epsilon ($\varepsilon$) was 0.795, as calculated according to Greenhouse and Geisser (1959), and was used to correct the ANOVA. Results indicated a significant time effect ($F(2.385, 69.178) = 4.789, p = .008, \eta^2_p = .142$). Mean scores are displayed in Table 9 and mean frequency of OE+ questions over time is displayed in Figure 6.

Table 9. Mean OE+ Questions across 10-minute TeachLivE Sessions

<table>
<thead>
<tr>
<th>TeachLivE Sessions</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
<td>.73 (1.7)</td>
<td>.77 (1.0)</td>
<td>1.27 (2.1)</td>
</tr>
</tbody>
</table>
Figure 6. Frequency of OE+ Questions over Time.

**Research Question 2: Differences in performance over time with simulation and no feedback on performance.** All of the above research questions were designed to investigate how providing performance feedback in after-action-review of simulation would impact teacher practice in a classroom simulator. Researchers for question 2 evaluated the effects of withholding feedback on a specific teacher practice (i.e., frequency of content-related affirmation) in a classroom simulator. To examine performance of teachers over four 10-minute sessions, a within-subjects ANOVA was performed. Time (four sessions) was cast as a within-subjects factor with a dependent variable of CRA. After each session, teachers were not presented with any data related to CRA. One observer collected data on frequency of OE questions asked by teachers per TeachLivE session. Analysis was conducted with a within-subjects design ANOVA. Mauchly’s test of sphericity indicated that the assumption of sphericity was violated, $X^2(5) = 16.138, p = .006$. Epsilon ($\varepsilon$) was 0.718, as calculated according to Greenhouse and Geisser (1959), and was used to correct the ANOVA. Results indicated no significant time effect ($F(2.153, 62.43) = .455, p = .651, \eta^2_p = .015$), which was expected, because no feedback had been provided. Mean scores are displayed in Table 10, and mean frequency of CRA over time is displayed in Figure 7.

Table 10. Mean CRA across 10-minute TeachLivE Sessions

<table>
<thead>
<tr>
<th>TeachLivE Sessions</th>
<th>Session 1</th>
<th>Session 2</th>
<th>Session 3</th>
<th>Session 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>30</td>
<td>30</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>M (SD)</td>
<td>5.13 (4.0)</td>
<td>5.33 (3.2)</td>
<td>5.73 (5.7)</td>
<td>4.57 (3.6)</td>
</tr>
</tbody>
</table>
Classroom Results

To examine impact of simulation of teachers in a real classroom with students present, the next research questions were designed to evaluate teacher performance on variables that had been part of the after-action-review in the simulator. Research question 3.1 evaluates teacher performance on a general measure of teacher practice (TPOT), while more specific practices were evaluated in questions 3.2 (CE), 3.3 (OE), and 3.4 (OE+).

Research question 3: Classroom results of simulation with feedback performance.

Question 3.1: TPOT Sum. An observer collected data on the TPOT Sum and two observers observed 30% of classes to establish inter-rater reliability. While performance feedback was not given using the TPOT Sum score as the measurement instrument, the score is considered to be a general measure of teacher performance. Reliability of scores between observers during both observations was calculated (pre-intervention, $r = .932$; post-intervention, $r = .882$). Results from a mixed ANOVA indicated there were not statistically significant changes in TPOT Sum scores between Observation 1 and 2 based on treatment group ($F(1,94) = .097, p = .757, \eta^2_p = .001$). For main effects, there was neither a statistically significant difference between the first and second observation collapsed across treatment groups ($F(1,94) = 1.460, p = .230, \eta^2_p = .015$), nor between groups collapsed across observations ($F(1,94) = .555, p = .458, \eta^2_p = .006$). Mean TPOT scores are displayed in Table 11 and TPOT scores over time are displayed in Figure 8.
Table 11. Mean TPOT Scores over Time

<table>
<thead>
<tr>
<th></th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>M (SD)</td>
</tr>
<tr>
<td>Comparison</td>
<td>46</td>
</tr>
<tr>
<td>TeachLivE</td>
<td>50</td>
</tr>
<tr>
<td>Total</td>
<td>96</td>
</tr>
</tbody>
</table>

Figure 8. TPOT Scores over Time.

**Question 3.2: CE questions.** An observer collected data on CE questions asked by the teacher, and two observers observed 30% of classes to establish inter-rater reliability. Reliability of scores between observers during both observations was calculated (pre-intervention, $r = .929$; post-intervention, $r = .798$). Results from a mixed ANOVA indicated there were no statistically significant changes in frequency of CE questions between Observation 1 and 2 based on treatment group ($F(1,100) = .796$, $p = .374$, $\eta^2_p = .008$). See Table 12 for mean frequency of CE questions over time by group and Figure 9 for mean frequency of CE questions over time.

Table 12. Mean CE Questions over Time by Group

<table>
<thead>
<tr>
<th></th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>M (SD)</td>
</tr>
<tr>
<td>Comparison</td>
<td>50</td>
</tr>
<tr>
<td>TeachLivE</td>
<td>52</td>
</tr>
<tr>
<td>Total</td>
<td>102</td>
</tr>
</tbody>
</table>
Figure 9. Frequency of CE Questions over Time.

**Question 3.3: OE questions.** An observer collected data on the number of OE questions asked by the teacher, and two observers 30% of classes to establish inter-rater reliability. Reliability of scores between observers during both observations was calculated (pre-intervention, \( r = .864 \); post-intervention, \( r = .954 \)). Results from a mixed ANOVA indicated there were not statistically significant changes in frequency of questions between Observation 1 and 2 based on treatment group \( (F(1,100) = 1.299, \ p = .257, \ \eta^2_p = .013) \). To determine the difference between groups at each level of time and vice versa, separate ANOVAs were run. There was no significant difference between treatment groups at Observation 1 \( (F(1,102) = 1.079, \ p = .301, \ \eta^2_p = .010) \) or Observation 2 \( (F(1,100) = .194, \ p = .661, \ \eta^2_p = .002) \). When comparing main effects over time by group, for the Comparison group, OE questions were not statistically significantly different between observations \( (F(1,49) = .282, \ p = .598, \ \eta^2_p = .006) \). However, for the TeachLivE group, OE questions were statistically significantly different between observations \( (F(1,51) = 4.403, \ p = .041, \ \eta^2_p = .079) \), with teachers increasing OE questions from Observation 1 \( (M = 9.81, \ SD = 5.83) \) to Observation 2 \( (M = 12.35, \ SD = 7.62) \). See Table 13 for mean frequency of OE questions over time by group and Figure 10 for mean frequency of OE questions over time.

<table>
<thead>
<tr>
<th>Group</th>
<th>Observations 1</th>
<th>Observations 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td><strong>Comparison</strong></td>
<td>11.12 (7.3)</td>
<td>11.74 (6.2)</td>
</tr>
<tr>
<td><strong>TeachLivE</strong></td>
<td>9.81 (5.8)</td>
<td>12.35 (7.6)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>10.45 (6.6)</td>
<td>12.05 (6.9)</td>
</tr>
</tbody>
</table>
**Question 3.4: OE+ questions.** An observer collected data on OE+ questions asked by the teacher, and two observers observed 30% of classes to establish inter-rater reliability. Reliability of scores between observers during both observations was calculated (pre-intervention, $r = .586$; post-intervention, $r = .792$). Results from a mixed ANOVA indicated there were statistically significant changes in frequency of OE questions between Observation 1 and 2 based on treatment group ($F(1,100) = 2.223$, $p = .030$, $\eta^2_p = .046$). To determine the difference between groups at each level of time and vice versa, separate ANOVAs were run. There was no significant difference between treatment groups at Observation 1 ($F(1,102) = 2.402$, $p = .124$, $\eta^2_p = .023$) or Observation 2 ($F(1,100) = 1.699$, $p = .195$, $\eta^2_p = .017$). For the Comparison group, OE+ was statistically significantly different between observations ($F(1,49) = 5.512$, $p = .023$, $\eta^2_p = .101$), with teachers significantly decreasing their OE+ from Observation 1 ($M = .96$, $SD = 1.93$) to Observation 2 ($M = .36$, $SD = .69$). For the TeachLivE group, OE+ was not statistically significantly different between observations ($F(1,51) = .323$, $p = .572$, $\eta^2_p = .006$), although teachers increased from Observation 1 ($M = .50$, $SD = .90$) to Observation 2 ($M = .62$, $SD = 1.21$). See Table 14 for mean OE+ questions over time by group and Figure 11 for mean frequency of OE+ questions over time.
Table 14. Mean OE+ Questions over Time by Group

<table>
<thead>
<tr>
<th>Observations</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>M (SD)</td>
<td>M (SD)</td>
</tr>
<tr>
<td>Comparison</td>
<td>50 (.96 (1.9)</td>
<td>.36 (.7)</td>
</tr>
<tr>
<td>TeachLivE</td>
<td>52 (.50 (.9)</td>
<td>.62 (1.2)</td>
</tr>
<tr>
<td>Total</td>
<td>102 (.73 (1.5)</td>
<td>.49 (1.0)</td>
</tr>
</tbody>
</table>

**Frequency of OE+ over Time**

Figure 11. Frequency of OE+ over Time.

**Research question 4: Classroom results of simulation without feedback performance.** An observer collected data on CRA asked by the teacher, and two observers observed 30% of classes to establish inter-rater reliability. Reliability of scores between observers during both observations was calculated (pre-intervention, $r = .931$; post-intervention, $r = .385$), and results should be interpreted with caution due to low reliability scores at post-observation. Results from a mixed ANOVA indicated there were no statistically significant changes in frequency of CRA between Observation 1 and 2 based on treatment group ($F(1,100) = .127$, $p = .722$, $\eta^2_p = .001$). See Table 15 for mean CRA over time by group and Figure 12 for mean frequency of CRA over time.
Table 15. Mean CRA over Time by Group

<table>
<thead>
<tr>
<th></th>
<th>Observations</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Comparison</td>
<td>50</td>
<td>5.12 (5.4)</td>
<td>4.78 (3.9)</td>
</tr>
<tr>
<td>TeachLivE</td>
<td>52</td>
<td>5.42 (5.2)</td>
<td>4.65 (5.4)</td>
</tr>
<tr>
<td>Total</td>
<td>102</td>
<td>5.27 (5.3)</td>
<td>4.71 (4.7)</td>
</tr>
</tbody>
</table>

**Frequency of CRA over Time**

Figure 12. Frequency of CRA over Time.

**Research Question 5: What are the perceptions of practicing teachers related to presence and perceptions of preparedness after completing TeachLivE?**

Teachers who received TeachLivE were administered a questionnaire that gathered information about their perceptions of authenticity in the simulator. As with Phase I, the large majority (76.9%) of teachers in Phase II agreed that the classroom simulator felt like a real classroom. Also in alignment with results from Phase I, teachers agreed that the virtual students accurately represented the kinds of people that existed in the real world (82.9%).

**Discussion**

This study was the second phase of a two-phase project. In Phase I middle school mathematics teachers were grouped by school and randomly assigned to one of four treatment conditions with varying levels of professional learning. Teachers were observed in their classrooms teaching the lesson and content of their choice, and those teachers who received TeachLivE were observed in the classroom simulator. Results indicated that four 10-minute sessions in the TeachLivE significantly improved targeted teaching behaviors in the simulator.
scenarios, and those improvements transferred into the teachers’ original classroom settings. In Phase II participants were primarily high school science teachers, and various refinements to methods occurred, including a reduction from four treatment groups to two (teachers who received lesson resources and teachers who received simulation) and uniform delivery of lesson plans.

As with Phase I, teachers overwhelmingly agreed that the classroom simulator felt like a real classroom and that the students also represented the kinds of students that existed in the real world. In terms of teacher practice, results from Phase I were replicated, in that teachers significantly increased targeted teaching practices in the simulator (OE+ questions) and improvements transferred into the teachers’ original classroom settings for OE questions. For OE+ questions, while there was a significant increase in the simulator, those effects did not carry over to the real classroom, and this was unanticipated. In the classroom, although teachers who received simulation increased their frequency of OE+ questions, they did not do so significantly. It is interesting to note that their colleagues who did not receive simulation significantly decreased their frequency of OE+ questions from observation 1 to 2. Both groups were observed teaching the same lesson, so it is possible that the differences in performance can be attributed to practice in the simulator. Hattie and Timperley (2007) indicated that the impact of feedback was largest when given relative to performance on a specific task with low complexity. It is possible that the feedback model resulted in less impact on OE+ questions because of the level of complexity. It is also possible that teachers who received simulation focused on too many performance objectives (OE and OE+ questions) and this resulted in a challenge to learning, reflected in the classroom when they only significantly changed practice related to one variable (OE questions). Consequently, the feedback model should be investigated to determine the best approach for impacting performance.

Consistent with Phase I, no significant difference in performance was observed in teachers’ classrooms when participants did not set performance objectives on a variable (CE questions). Looking back to teacher performance in the simulator, CE questions decreased significantly, which was expected, because teachers were focusing on increasing an opposing behavior of OE questions. Finally, consistent with Phase I, when teachers were not provided with feedback on a variable (CRA), no significant difference in performance was observed. This is not surprising, as many researchers indicate setting objectives and providing feedback are essential components to improving teacher performance (Hattie & Timperley, 2007). Our work underscores the importance of providing a structured after-action-review that takes into account best practices for providing feedback on performance. As we saw in Phase I and II, when teachers did not receive data on their performance they did not change their practice.

As a whole, results validated emerging research in the field that suggests that professional learning in mixed-reality simulated classrooms can be effective. We found support for our hypothesis that simulation would increase teachers’ frequency of OE questions and that this increase also would be observed in their classrooms. Teachers who took part in a series of sessions significantly increased their instances of OE+ questions in the simulator, similar to studies conducted earlier (e.g., Dawson & Lignugaris/Kraft, 2013; Elford et al., 2013; Garland et al.; 2012), and their performance in OE questions also increased significantly in comparison to colleagues who did not receive simulation. Overall, findings from Phase I of this study were replicated in Phase II, providing support for our overarching hypothesis that teachers who engage
in TeachLivE professional development can improve their pedagogical knowledge.

Limitations

In Phase I, limitations to internal validity resulted from the nested design in which teachers were grouped by school, because teachers within one school may be more similar than teachers across schools. However, in Phase II, random assignment occurred at the teacher level, rather than the school level, based on the idea that performance in a simulator was individualized and that threats to validity as a result of treatment diffusion (i.e., treatment effects spreading from one group to another) were unlikely. This random assignment at the teacher level allowed for balancing of similarities within each school. However, as with Phase I, the original research design was a randomized trial, yet the nature of the design changed as a consequence of teacher requests to change initial treatment conditions. In each case, teachers remained in the study but reported that they would only participate if allowed to change treatment groups ($n = 7$). This phenomenon violated random assignment and changed the research design to a quasi-experiment.

Other threats to data reliability and confounding factors also existed. In Phase I classroom observation data had low reliability, so operational definitions were revised and data collection training was improved to increase reliability of results. Data collectors engaged in more rigorous group coding activities, and data met reliability thresholds (75% or better) in almost all variables at observation 1 and 2. Most significantly, in Phase I, classroom instruction was not standardized by a common lesson, and so content and format varied widely. In Phase II, this confounding variability was removed in that all teachers taught a common lesson at each observation, providing a stronger basis for comparison.

Of the 11 research sites, only 1 was delivered in the school setting. At this site, a researcher brought the simulation equipment to the teachers’ school using a mobile unit. This took a significant amount of coordination between technology staff on the research team and with the district, as the software requires specific network settings. At the other 10 sites, teachers traveled to the simulation sites located at institutes of higher education. Teachers receiving simulation were required to visit the simulator three times, which required significant scheduling efforts in the cases of last minute cancellations or delays resulting from technology issues. Cancellations due to travel were not an issue for the mobile lab; however, new barriers to scheduling arose, as teachers were more likely to run late to sessions as they tried to juggle on-site job duties. Future research should explore the idea of school district-level coordination for professional learning so teachers do not have competing demands for their attention.

Future Research and Implications

Findings from this study can be generalized to other high school science teachers who receive four 10-minute sessions of TeachLivE with after-action-review. Because results were replicated from Phase I to II, the next important step is to evaluate the impact of varying durations, frequency, and total number of simulations to determine the optimal level of treatment needed to produce the desired results. This question of dosage is critical to unlocking the benefits of simulation for busy teachers and school districts with limited financial resources; simulation has the potential to deliver professional learning in an accelerated format, on site at schools, and in a compressed amount of time. Identifying the components of effective
simulation will save valuable time and money. In light of findings from Phase I and II, the team currently has three areas of unanswered questions related to time. First, if four 10-minute sessions impact practice, how long does this change in practice continue? Second, what is the optimal frequency of practice over time to ensure retention of new skills acquired in the simulator? Or, in other words, do teachers need to re-visit the simulator to practice once a month, a semester, or a year? Third, in both Phase I and II, a pattern was observed that after about five to seven minutes of working on a new skill, teachers tend to fall back to patterns of old behavior. Early work is being done to establish if shorter sessions result in a significant increase in practice as we explore the motivational value of adding micro-credentials for reaching goals with specific high-level practices.

Our findings have important implications for researchers and educators designing simulation activities, as simulation with no feedback is likely a waste of resources. The simulation after-action-review model should be examined in a component analysis to identify best practices for simulation integrated with feedback. In both Phase I and II, the researchers found that teachers did not change their performance when no feedback was provided. A basic understanding is emerging that feedback in simulation is essential, but future research should explore the aspects of after-action-review to determine the most effective model.

Just as models of feedback need further investigation, so do methods of grouping teacher participants in the simulator. For this project teachers attended sessions individually, resulting in the most costly use of the simulator. Would it be more effective for teachers to attend simulation activities in small groups (lesson studies or even Professional Learning Communities) with the twofold benefit of capitalizing on the social nature of learning and cost-savings? School districts and teacher preparation institutes seeking to save on simulation budgets are interested in learning if grouping individuals enhances or degrades the simulation experience, with the idea that more people can be trained if group models are beneficial. While questions surrounding group versus individual models of training were not explored in this project, there is much interest in the impact of these competing formats since understanding their relative benefits can further inform the field as more simulation technology is used for teacher learning.

As we continue to collect data to support the use of simulation in teacher education, we also are interested in the impact of simulation in other domains. Our future work for Phase III of this project explores the use of simulation with adult avatars for other educational professionals such as administrators, guidance counselors, psychologists, and counselor educators. Use of simulation for parent-teacher conferences can provide invaluable experiences for professionals who likely did not engage with parents during their clinical field experiences or at any time during their preparation programs. Simulation can provide a safe practice ground so individuals can learn from mistakes without harming relationships with parents. In the area of counselor education, techniques for counseling can be provided with adult avatars to practice challenging, difficult interactions. Because teachers report that the mixed-reality simulation feels authentic to teaching, the same feelings of authentic presence will likely be found with the adult avatar interactions. This expansion to the adults space is a large potential area of exploration, as more and more sites are expressing interest in using adult avatars for difficult conversations.

Beyond educational professionals’ use of simulation, we are highly interested in how simulation might also be used to impact student learning. As part of Phase III of our research,
we are investigating student social skill interactions for students with autism and problem-solving skills for students with intellectual disabilities.

Our findings support the theory that simulated learning environments provide an efficient tool for learning and practicing new teaching strategies. We have found that four, 10-minute simulator sessions on a specific teaching skill can change teacher behavior not only in the simulator with student-avatars, but also in the classroom with real students. Teachers have the opportunity to practice, make mistakes, and try new approaches to retool their teaching, all in a safe place for teachers and students. As we explore the next generation of tools for teachers’ professional and personalized learning, we keep in mind that the ultimate goal of the research team is that the simulator does not replace “real” teaching, but instead allows for safe and fun practice that is targeted and personalized. Our vision is that simulators can be used to prepare and retool the skills of teachers at all levels from pre-service to in-service, as well as individuals in other domains whose interactions are complex and nuanced like the act of teaching.
References


What is Technology?

First Observation

Overview
This lesson involves classroom discussion and a short scenario to allow students to develop a sense of what technology is, dispel the notion that technology relates mostly to computers, and examine the impact of technology. The lesson is designed to enhance science literacy and is aligned with Disciplinary Core Ideas and Cross-cutting Concepts from the Next Generation Science Standards, as well as the Common Core Standards for Literacy in Science. The lesson is based on the 5E Instructional Model and has been validated and field-tested in high school Biology classrooms as part of a larger module from the NIH Curriculum Supplement Series “Using Technology to Study Cellular and Molecular Biology” which can be found in its entirety online at http://science.education.nih.gov-supplements/nih4/technology-guide/nih_technology_curr-sup.pdf

Major Concepts
Technology is a body of knowledge used to create tools, develop skills, and extract or collect materials. It is also the application of science (the combination of the scientific method and material) to meet an objective or solve a problem.

Standards-based Objectives
- be able to explain what technology is
- recognize that human intervention is the common bond among technological ecosystems
- use mathematical representations to support and revise explanations based on evidence about factors affecting biodiversity and populations in ecosystems of different scales
- evaluate a solution to a complex real-world problem based on prioritized criteria and trade-offs that account for a range of constraints, including cost, safety, reliability, and aesthetics as well as possible social, cultural, and environmental impacts

Standards
See Appendix A for detailed list of standards.
**Before Class**

1. Follow instructions to fill out the *Cross-Reference Demographics Sheet*

2. Administer the *Student Assessment*

3. For the lesson, make sure you have a copy for each student:
   - *Handout 1.1: BEFORE Lesson Concept Map* (white paper)
   - *Handout 1.2: Estimated No. of Polio Cases per Year & Timeline* (white paper)
   - *Handout 1.3: AFTER Lesson Concept Map* (yellow paper)

4. Be prepared to give the following to the observer at the end of the first observation:
   - *Cross-Reference Demographics Sheet*
   - *Handout 1.1: BEFORE Lesson Concept Map* for each student
   - *Handout 1.3: AFTER Lesson Concept Map* for each student
   - *Student Assessment* for each student

**Materials**

Have a set of all three handouts for each student:
- *Handout 1.1: BEFORE Lesson Concept Map* (white paper)
- *Handout 1.2: Estimated No. of Polio Cases per Year & Timeline* (white paper)
- *Handout 1.3: AFTER Lesson Concept Map* (yellow paper)

This lesson includes science literacy mini-tasks embedded throughout the lesson. Look for the Literacy Design Collaborative (LDC) logo to see the location of science literacy mini-tasks. For more information about increasing science literacy, visit the LDC at http://www.literacydesigncollaborative.org/. The LDC offers a framework for building the college-and-career-ready literacy skills specified by the Common Core State standards.

**Procedure**

*Pre-Lesson: 10 minutes*

Before the lesson, pass out *Handout 1.1: BEFORE Lesson Concept Map* (white paper). Have students follow instructions to create student codes.

While the students are completing their codes, draw a model concept map on the board that looks like this:

![Concept Map Diagram](image)

After you have drawn the model and have students’ attention, say:

“This is a concept map. It is used for understanding concepts and relationships. A general topic/question is in the circle at the top of the map. Write the more specific concepts that relate in some way to the general concept. Tie the general and specific concepts together with linking words in some way that makes sense or has meaning to you. Look for cross-linkages between the general and more specific concepts. You have 5 minutes to generate as many ideas as you can.”

After 5 minutes, collect the maps and then begin the lesson.
Part 1: 10 minutes

1. Begin by asking the class, “How do you define technology?”

Accept all answers and write student responses on the board. Do not attempt to have students refine their definitions of technology at this point. They will revisit their definitions and refine them later.

Students, like older individuals, may harbor the preconception that technology relates mostly to computers. Through advertisements and media articles, they are familiar with the terms information technology and computer technology.

Teacher note: Asking this question requires students to call on their prior knowledge, and it engages their thinking. At this point, do not critique student responses. Appropriate teacher comments are short and positive, such as “good” and “what else?” Other appropriate teacher responses include, “Why do you believe that?” or “How do you know that?” Questions such as these allow the teacher to assess students current knowledge about the subject and to adjust lessons accordingly. They also provide a springboard to “Let’s find out” or “Let’s investigate.” In general, it is time to move forward when you see that thinking has been engaged.

2. Ask students, “In general, what does technology do for us?”

This question may help students understand that technology helps us solve problems, makes our lives easier, and extends our abilities to do things. Technology is used to develop skills or tools, both in our daily lives and in our occupations.

Enrichment: If students bring up the term ecosystem, as it may pertain to past biology concepts, it is appropriate to discuss ecosystems.

3. Focus discussion on technologies that are relevant to each student’s life. Ask students to look around the room. What technologies do they see? How do these technologies solve problems and make their lives easier in society, culture, and the environment?

Accept all responses and write them on the board. Students may mention any number of items. Some may be school-related, such as binders, backpacks, pens, pencils, paper, and paper clips. Other items may be more personal, such as water bottles, personal stereos, and hair clips. Students may neglect items such as shoelaces, zippers, buttons, fabric, eyeglasses or contact lenses, makeup, and bandages. Discussion should reinforce the notion that humans develop technology with a specific objective in mind. A related concept is that a given task requires the right tool or tools.

4. Turn the discussion to how technology has impacted major world problems such as disease. Ask “What diseases have been impacted by technology?” and have students talk in small groups, before discussing as a whole class.

After a brief discussion as a class, tell students they will now use mathematics to support explanations using data from the Bill & Melinda Gates Foundation about technology and the number of cases of polio worldwide.

Misconception Alert: At this point in the discussion, you may need to clarify the concept of viruses and bacteria for students.
Part 2: 15 minutes

5. Pass out Handout 1.2: Estimated Number of Polio Cases per Year & Timeline (white paper copies provided). Explain to students you are now going to talk about a specific disease and theorize how technology impacted polio.

6. Ask students to read Handout 1.2 and ask, “How do you interpret the trend in estimated number of polio cases per year?”

Students first should recognize that the estimated number of polio cases per year is decreasing. Bring attention to the embedded graph. Ask students, “Why do you think the creators of the graph decided to show the data in this way?” Students may have the misconception that the number of estimated cases in 2000 was similar to that of 1989, because of how the data are displayed and the scale that is used. Have students pay attention to the y-axis.

7. Ask students to read Handout 1.2. Ask, “How can evidence from the graph or timeline be used to make claims about technology’s impact on polio?”

Student responses will vary, and some students may want to look for events which correspond exactly to years on the graph when polio cases showed a dramatic decrease. Students may also note that the data begins in 1988 when the Global Polio Eradication Initiative was launched. They may note that the years in the graph are not the same as the years in the timeline, providing the opportunity to discuss the limitations of data sets. Slow them down and have them consider cumulative changes in knowledge and technologies. Advances in microbiology and chemistry contributed to development of vaccines over time. Innovations such as the iron lung were early technological advances for polio. Students should derive and understanding that a relationship exists between problems and the technology to solve it.

8. On the basis of previous discussion, ask the students to rethink and refine their definition of technology (from Step 1), including its impact on society, culture, and environment. Students should mention that technology is a way of solving problems through the application of knowledge from multiple disciplines.

Post-Lesson: 10 minutes

LDC Mini-task: Lesson Reflection. The concept map activity at the end of the lesson facilitates review, reflection, elaboration, and consolidation of student understanding.

After the lesson, pass out Handout 1.3: AFTER Lesson Concept Map (yellow paper). Have students follow instructions to create student codes. While the students are completing their codes, draw a model concept map on the board that looks like this:

After you have drawn the model and have students’ attention, say:

“This is a concept map. It is used for understanding concepts and relationships. A general topic/question is in the circle at the top of the map. Write the more specific concepts that relate in some way to the general concept. Tie the general and specific concepts together with linking words in some way that makes sense or has meaning to you. Look for cross-linkages between the general and more specific concepts. You have 5 minutes to generate as many ideas as you can.”

After 5 minutes, collect the maps and give all documents to the observer.
Appendix A: Standards

Students who demonstrate understanding can:
Use mathematical representations to support and revise explanations based on evidence about factors affecting biodiversity and populations in ecosystems of different scales (Biology Domain and Biology Repeat HS-LS2-2).
Evaluate a solution to a complex real-world problem based on prioritized criteria and trade-offs that account for a range of constraints, including cost, safety, reliability, and aesthetics as well as possible social, cultural, and environmental impacts (Biology Domain and Crosscutting Concept with Chemistry, Physics, and Earth & Space Science HS-ETS1-3).

Students who demonstrate understanding can:
CCSS.ELA-Literacy.RST.11-12.9 Synthesize information from a range of sources (e.g., texts, experiments, simulations) into a coherent understanding of a process, phenomenon, or concept, resolving conflicting information when possible.
CCSS.ELA-Literacy.SL.9-10.1c Propel conversations by posing and responding to questions that relate the current discussion to broader themes or larger ideas; actively incorporate others into the discussion; and clarify, verify, or challenge ideas and conclusions.
CCSS.ELA-Literacy.SL.9-10.1d Respond thoughtfully to diverse perspectives, summarize points of agreement and disagreement, and, when warranted, qualify or justify their own views and understanding and make new connections in light of the evidence and reasoning presented.
CCSS.ELA-Literacy.SL.9-10.4 Present information, findings, and supporting evidence clearly, concisely, and logically such that listeners can follow the line of reasoning and the organization, development, substance, and style are appropriate to purpose, audience, and task.
CCSS.ELA-Literacy.CCRA.SL.1 Prepare for and participate effectively in a range of conversations and collaborations with diverse partners, building on others’ ideas and expressing their own clearly and persuasively.
CCSS.ELA-Literacy.CCRA.SL.2 Integrate and evaluate information presented in diverse media and formats, including visually, quantitatively, and orally.
CCSS.ELA-Literacy.CCRA.SL.3 Evaluate a speaker’s point of view, reasoning, and use of evidence and rhetoric.
What does technology do for us?

Teacher Code SAMPLE:

1. Write your three-letter initials (if no middle initial, use the second letter of your first name):

2. Circle the day of the month you were born: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

Follow the instructions below to create student codes:

Handout 1.1: BEFORE Lesson Concept Map
Handout 1.2: Estimated Number of Polio Cases per Year & Timeline

**Polio Timeline**

**3000 BC**
Egyptian paintings and carvings depict people with withered limbs and walking with canes

**1928**
First iron lung used at Children’s Hospital in Boston

**1952**
Worst polio outbreak in the United States history, with 658,000 reported cases

**1955**
Jonas Salk’s injected polio vaccine introduced

**1963**
Albert Sabin’s oral polio vaccine licensed

**1979**
Last case of naturally occurring polio in the United States

**1970s**
National immunization programs launched leading to control of the disease in many developing countries

**1988**
Polio still exists in 125 countries and paralyzes an estimated 350,000 children; Global Polio Eradication Initiative created

**2007**
The World Health Organization declares polio eradicated in the Americas, Europe, and the Western Pacific.

**2010**
Sustained transmission of polio in four countries, but outbreaks in 16 countries are reminders that polio anywhere is a threat everywhere

Source: Bill & Melinda Gates Foundation; Polio Infographics. [http://www.gatesfoundation.org/What-We-Do/Global-Development/Polio/Infographics](http://www.gatesfoundation.org/What-We-Do/Global-Development/Polio/Infographics)
Handout 1.3: AFTER Lesson Concept Map

What does technology do for us?

1. Write your three-letter initials (if no middle initial use the second letter of your first name): 

2. Circle the DAY of the month you were born: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

Teacher Code: SAMPLE
Overview
This lesson involves classroom discussion and examination of a biological modeling system using technology. The lesson is designed to enhance science literacy and is aligned with Disciplinary Core Ideas and Cross-cutting Concepts from the Next Generation Science Standards, as well as the Common Core Standards for Literacy in Science. The lesson is based on the 5E Instructional Model and has been validated and field-tested in high school Biology classrooms as part of a larger module from the NIH Curriculum Supplement Series “Using Technology to Study Cellular and Molecular Biology” which can be found in its entirety online at http://science.education.nih.gov/supplements/nih4/technology/guide/nih_technology_curr-supp.pdf

Major Concepts
The process of modeling biological systems involves the use of computer simulations to analyze and visualize complex processes. Computers are widely used today to simulate a system’s response to internal and external stimuli. The goal of modeling is to create accurate, real-time models.

Standards-based Objectives
After completing this lesson, students will:
• be able to explain what modeling is
• recognize that as technologies change, so do our modeling capabilities
• use mathematical representations to support and revise explanations based on evidence about factors affecting biodiversity and populations in ecosystems of different scales
• evaluate a solution to a complex real-world problem based on prioritized criteria and trade-offs that account for a range of constraints, including cost, safety, reliability, and aesthetics as well as possible social, cultural, and environmental impacts

Standards
See Appendix A for detailed list of standards.
**Before Class**

1. Follow instructions to fill out the *Cross-Reference Demographics Sheet*

2. Administer the *Student Assessment*

3. For the lesson, make sure you have a copy for each student:
   - Handout 2.1: *BEFORE Lesson Concept Map* (blue paper)
   - Handout 2.2: *Annual U.S. Flu Activity* (white paper)
   - Handout 2.3: *AFTER Lesson Concept Map* (pink paper)

4. Be prepared to give the following to the observer at the end of the first observation:
   - Cross-Reference Demographics Sheet
   - Handout 2.1: *BEFORE Lesson Concept Map* for each student
   - Handout 2.3: *AFTER Lesson Concept Map* for each student
   - *Student Assessment* for each student

**Materials**

Have a set of all three handouts for each student:
- *Handout 2.1: BEFORE Lesson Concept Map* (blue paper)
- *Handout 2.2: Annual U.S. Flu Activity* (white paper)
- *Handout 2.3: AFTER Lesson Concept Map* (pink paper)

This lesson includes science literacy mini-tasks embedded throughout the lesson. Look for the Literacy Design Collaborative (LDC) logo to see the location of science literacy mini-tasks. For more information about increasing science literacy, visit the LDC at http://www.literacydesigncollaborative.org/. The LDC offers a framework for building the college-and-career-ready literacy skills specified by the Common Core State standards.

**Procedure**

**Pre-Lesson:**

- **10 minutes**

  Before the lesson, pass out *Handout 2.1: BEFORE Lesson Concept Map* (blue paper). Have students follow instructions to create student codes. While the students are completing their codes, draw a model concept map on the board that looks like this:

  ![Concept Map Diagram](image)

  **LDC Mini-task:**
  
  *Activate Prior Knowledge.*

  The concept map activity is designed to engage students in learning about technology by activating their prior knowledge.

  After you have drawn the model and have students’ attention, say:

  “This is a concept map. It is used for understanding concepts and relationships. A general topic/question is in the circle at the top of the map. Write the more specific concepts that relate in some way to the general concept. Tie the general and specific concepts together with linking words in some way that makes sense or has meaning to you. Look for cross-linkages between the general and more specific concepts. You have 5 minutes to generate as many ideas as you can.”

  After 5 minutes, collect the maps and then begin the lesson.
**Part 1: 10 minutes**

1. Begin by writing on the board, “Technology is a means of extending human potential or of extending human senses.” Ask students, “Who agrees or disagrees with this statement and why?” Ask students, “What specific technologies can you relate to the extension of human attributes or senses?”

   Students will generally agree that technology extends human potential. Obvious examples include the wheel and other transportation innovations that extend our ability to communicate. Microscopes, telescopes, eyeglasses, and contact lenses extend and enhance our sense of vision. Computers and written materials can be seen as ways to extend memory. Appropriate teacher comments are short and positive, such as “good” and “what else?” Others include, “Why do you believe that?” or “How do you know that?”

2. Ask students to consider only technologies that have increased our understanding of living systems. Ask, “What human attributes are extended?”

   Students will probably focus on those that extend vision, since they are the easiest to recognize. Examples could include radar, eyeglasses, contact lenses, and telescopes. Other technologies might be mentioned. Accept all responses and write them on the board. This is an opportunity to identify students’ current understanding of these technologies.

3. Ask students, “How do data and technology extend what we know?” In steps 1 and 2, students will likely focus on technologies that extend human senses since they are the easiest to recognize.

4. Move the discussion to physical or computer models used to explain or detect phenomena. Ask “What physical or computer models have we discussed in class?” and have students talk in small groups before discussing as a whole class.

   The process of modeling biological systems involves the use of computer simulation to extend what we know by analyzing and visualizing complex processes. Computers are widely used today to simulate a system’s response to internal and external stimuli. The goal of modeling is to create accurate, real-time models.

**Part 2: 15 minutes**

5. Pass out Handout 2.2: Annual U.S. Flu Activity (blue paper copies provided) and explain that it shows two different methods of gathering data using technology in order to detect influenza: the Center for Disease Control (CDC) and Google search engine terms (Google).

   Explain CDC publishes national and regional data reported to them from physician’s offices on a weekly basis, typically with a 1-2 week reporting lag. In an attempt to provide faster detection, innovative surveillance systems have been created to monitor indirect signals of influenza, such as call volume to telephone triage advice lines and over-the-counter-drug sales. Google web search logs provide estimates that are current each day.
6. Ask students, “How do you interpret the trend lines of CDC versus Google?” and ask them to justify their answers: “What evidence do you see in the graphs?” Student responses will vary. Student should identify that the influenza-like illnesses (ILI) percentages are similar for CDC and Google from Figure 1. From Figure 2, they should note that the Google model received data at a faster pace. Ask the students, “How did technologies impact those findings?”

7. Turn the discussion to the importance of being able to detect illness such as influenza. Begin by asking students, “Why is using technology from early detection of a disease like influenza important to the community?”

Students should identify the benefits of early detection of influenza disease. Early detection of disease activity, when followed by a rapid response, can reduce the impact of both seasonal and pandemic influenza. Seasonal influenza epidemics are a major public health concern, causing tens of millions of respiratory illnesses and 250,000 to 500,000 deaths worldwide each year. Early detection of disease activity, when followed by a rapid response, can reduce the impact of both seasonal and pandemic influenza.


8. On the basis of previous discussion, ask the students to rethink how technology extends our human potential and senses, socially, culturally, and environmentally. Students should mention that models provide a way to detect relationships in systems.

Post-Lesson: 10 minutes

LDC Mini-task:
Lesson Reflection.
The graphic organizer activity at the end of the lesson facilitates review, reflection, elaboration, and consolidation of student understanding.

After the lesson, pass out Handout 2.3: AFTER Lesson Concept Map (pink paper). Have students follow instructions to create student codes. While the students are completing their codes, draw a model concept map on the board that looks like this:

After you have drawn the model and have students’ attention, say: “This is a concept map. It is used for understanding concepts and relationships. A general topic/question is in the circle at the top of the map. Write the more specific concepts that relate in some way to the general concept. Tie the general and specific concepts together with linking words in some way that makes sense or has meaning to you. Look for cross-linkages between the general and more specific concepts. You have 5 minutes to generate as many ideas as you can.” After 5 minutes, collect the maps and give all documents to the observer.
Appendix A: Standards

Students who demonstrate understanding can:
Use mathematical representations to support and revise explanations based on evidence about factors affecting biodiversity and populations in ecosystems of different scales (Biology Domain and Biology Repeat HS-LS2-2).
Evaluate a solution to a complex real-world problem based on prioritized criteria and trade-offs that account for a range of constraints, including cost, safety, reliability, and aesthetics as well as possible social, cultural, and environmental impacts (Biology Domain and Crosscutting Concept with Chemistry, Physics, and Earth & Space Science HS-ETS1-3).

Students who demonstrate understanding can:
CCSS.ELA-Literacy.RST.11-12.9. Synthesize information from a range of sources (e.g., texts, experiments, simulations) into a coherent understanding of a process, phenomenon, or concept, resolving conflicting information when possible.
CCSS.ELA-Literacy.SL.9-10.1c. Propose conversations by posing and responding to questions that relate the current discussion to broader themes or larger ideas; actively incorporate others into the discussion; and clarify, verify, or challenge ideas and conclusions.
CCSS.ELA-Literacy.SL.9-10.1d. Respond thoughtfully to diverse perspectives, summarize points of agreement and disagreement, and, when warranted, qualify or justify their own views and understanding and make new connections in light of the evidence and reasoning presented.
CCSS.ELA-Literacy.SL.9-10.4. Present information, findings, and supporting evidence clearly, concisely, and logically such that listeners can follow the line of reasoning and the organization, development, substance, and style are appropriate to purpose, audience, and task.
CCSS.ELA-Literacy.CCRA.SL.1. Prepare for and participate effectively in a range of conversations and collaborations with diverse partners, building on others’ ideas and expressing their own clearly and persuasively.
CCSS.ELA-Literacy.CCRA.SL.2. Integrate and evaluate information presented in diverse media and formats, including visually, quantitatively, and orally.
CCSS.ELA-Literacy.CCRA.SL.3. Evaluate a speaker’s point of view, reasoning, and use of evidence and rhetoric.
Handout 2.1: BEFORE Lesson Concept Map

**How does technology extend our potential?**

**Concept Map Instructions:**
Wait for your teacher's signal to begin.

This is a concept map. It is used for understanding concepts and relationships. A general topic/question is in the circle at the top of the map. Write the general and more specific concepts that relate in some way to the general concept. Tie the general and specific concepts together with linking words in some way that makes sense or has meaning to you. Look for cross-linkages between the general and more specific concepts.

### Students

Follow the instructions below to create student codes:

1. Write your three-letter initials (if no middle initial, use the second letter of your first name):

2. Circle the DAY of the month you were born:

   1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27  28  29  30  31

**Teacher Code:** SAMPLE
Handout 2.2: Annual U.S. Flu Activity

Figure 1

Center for Disease Control (CDC) Method: collects physician visits data and publishes weekly

Google Method: collects search terms daily

ILI = influenza-like illness
Google = Black line
CDC = Red Line

Figure 2

Handout 2.3: AFTER Lesson Concept Map

How does technology extend our potential?

**Concept Map Instructions:**

1. Follow the instructions below to create student codes:
   - Write your three-letter initials (if no middle initial use the second letter of your first name): _______ _______.
   - Circle the day of the month you were born: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

2. Look for cross-linkages between the general and more specific concepts. This is a concept map. It is used for understanding concepts and relationships. A general topic/question is in the circle at the top of the map. Write the more specific concepts that relate in some way to the general concept. Look for cross-linkages between the general and more specific concepts. Write linking words in some way that makes sense or has meaning to you.
APPENDIX C: CONCEPT MAP DIRECTIONS
First Observation
Concept Map Directions for Handouts 1.1 & 1.3

Handout 1.1 Instructions – White Paper

Before the lesson, pass out Handout 1.1: BEFORE Lesson Concept Map (white paper). Have students follow instructions to create student codes. While the students are completing their codes, draw a model concept map on the board that looks like this:

After you have drawn the model and have students’ attention, say:
“This is a concept map. It is used for understanding concepts and relationships. A general topic/question is in the circle at the top of the map. Write the more specific concepts that relate in some way to the general concept. Tie the general and specific concepts together with linking words in some way that makes sense or has meaning to you. Look for cross-linkages between the general and more specific concepts. You have 5 minutes to generate as many ideas as you can.”

After 5 minutes, collect the maps and then begin the lesson.

Handout 1.3 Instructions – Yellow Paper

After the lesson, pass out Handout 1.3: AFTER Lesson Concept Map (yellow paper). Have students follow instructions to create student codes. While the students are completing their codes, draw a model concept map on the board that looks like this:

After you have drawn the model and have students’ attention, say:
“This is a concept map. It is used for understanding concepts and relationships. A general topic/question is in the circle at the top of the map. Write the more specific concepts that relate in some way to the general concept. Tie the general and specific concepts together with linking words in some way that makes sense or has meaning to you. Look for cross-linkages between the general and more specific concepts. You have 5 minutes to generate as many ideas as you can.”

After 5 minutes, collect the maps and give all documents to the observer.
APPENDIX D: TEACHER PRACTICE OBSERVATION TOOL (TPOT)
<table>
<thead>
<tr>
<th>Interval 1</th>
<th>Open-ended and Open-ended+</th>
<th>Closed-Ended Questions</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Danielson: 1 min</td>
<td><strong>Managing Student Behavior &amp; Managing Classroom Procedures</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questions: 3 min</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affirmation: and 3 min</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student Talk Duration in seconds (do not reset, continue on next interval):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field Log 2 min: Write at least one note about practice per area then check appropriate box. Provide comments to justify rating.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Managing student behavior:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ no established standards of conduct</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ inconsistent standards of conduct</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ teacher established standards</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ students self-monitor with standards</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Managing classroom procedures:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ much instructional time is lost</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ some instructional time is lost</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ little loss of instructional time</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ time is maximized</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interval 2</td>
<td>Open-ended and Open-ended+</td>
<td>Closed-Ended Questions</td>
<td>Comments</td>
</tr>
<tr>
<td>Danielson: 1 min</td>
<td><strong>Establishing a Culture for Learning &amp; Engaging Students in Learning</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Questions: 3 min</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affirmation: and 3 min</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student Talk Duration in seconds (continue from last interval):</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Field Log 2 min: Write at least one note about practice per area then check appropriate box. Provide comments to justify rating.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Establishing a culture for learning:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ lack of commitment to learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ little commitment to learning</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ high expectations by teacher</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ shared belief in importance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Engaging students in learning:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ few engaged</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ some engaged</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ most engaged</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>☐ virtually all highly engaged</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Interval 3

<table>
<thead>
<tr>
<th>Danielson:</th>
<th>1 min</th>
<th><strong>Communicating with Students &amp; Using Questioning and Discussion</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Questions:</td>
<td>3 min</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Affirmation:</th>
<th>3 min</th>
</tr>
</thead>
</table>

**Student Talk**

**Duration in seconds (continue from last interval):**

**Field Log**

2 min: Write at least one note about practice per area then check appropriate box. Provide comments to justify rating.

**Communicating with students:**

- [ ] explanations confusing or with errors
- [ ] explanations initially confusing
- [ ] explanations clearly communicated
- [ ] explanations clear & anticipate confusion

**Using questioning and discussion techniques:**

- [ ] a few students respond
- [ ] some students discuss
- [ ] teacher engages most students
- [ ] students extend discussion

### Interval 4

<table>
<thead>
<tr>
<th>Danielson:</th>
<th>1 min</th>
<th><strong>Creating and Environment of Respect and Rapport &amp; Using Assessment</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Questions:</td>
<td>3 min</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Affirmation:</th>
<th>3 min</th>
</tr>
</thead>
</table>

**Student Talk**

**Duration in seconds (continue from last interval):**

**Field Log**

2 min: Write at least one note about practice per area then check appropriate box. Provide comments to justify rating.

**Creating an environment of respect and rapport:**

- [ ] mostly negative interactions
- [ ] generally appropriate interactions
- [ ] general caring and respect
- [ ] genuine warmth and caring

**Using assessment in instruction:**

- [ ] little or none
- [ ] used sporadically
- [ ] used regularly
- [ ] fully integrated (formative assessment)

### Interval 5

<table>
<thead>
<tr>
<th>Danielson:</th>
<th>1 min</th>
<th><strong>Communicating with Students &amp; Using Questioning and Discussion</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Questions:</td>
<td>3 min</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Affirmation:</th>
<th>3 min</th>
</tr>
</thead>
</table>

**Student Talk**

**Duration in seconds (continue from last interval):**

**Field Log**

2 min: Write at least one note about practice per area then check appropriate box. Provide comments to justify rating.

**Communicating with students:**

- [ ] explanations confusing or with errors
- [ ] explanations initially confusing
- [ ] explanations clearly communicated
- [ ] explanations clear & anticipate confusion

**Using questioning and discussion techniques:**

- [ ] a few students respond
- [ ] some students discuss
- [ ] teacher engages most students
- [ ] students extend discussion
Final Summary Check:
1. Now that you have completed the full observation, write brief statements of the teacher’s practice in the space provided below.
2. Choose the final level by checking one box within each category and be sure to provide justification for your rating.
3. If you are conducting this observation with another rater for inter-rater reliability, you may then discuss your chosen levels.
4. After your conversation, you may choose to make a change to one of the levels below.
5. If you choose to make a change, please provide a justification.
6. No changes may be made to the frequency counts or the field logs on pages 1-2.

<table>
<thead>
<tr>
<th>Managing student behavior:</th>
<th>□ no established standards of conduct □ inconsistent standards of conduct □ teacher established standards □ students self-monitor with standards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Managing classroom procedures:</td>
<td>□ much instructional time is lost □ some instructional time is lost □ little loss of instructional time □ time is maximized</td>
</tr>
<tr>
<td>Establishing a culture for learning:</td>
<td>□ lack of commitment to learning □ little commitment to learning □ high expectations by teacher □ shared belief in importance</td>
</tr>
<tr>
<td>Engaging students in learning:</td>
<td>□ few engaged □ some engaged □ most engaged □ virtually all highly engaged</td>
</tr>
<tr>
<td>Communicating with students:</td>
<td>□ explanations confusing or with errors □ explanations initially confusing □ explanations clearly communicated □ explanations clear &amp; anticipate confusion</td>
</tr>
<tr>
<td>Using questioning and discussion techniques:</td>
<td>□ a few students respond □ some students discuss □ teacher engages most students □ students extend discussion</td>
</tr>
<tr>
<td>Creating an environment of respect and rapport:</td>
<td>□ mostly negative interactions □ generally appropriate interactions □ general caring and respect □ genuine warmth and caring</td>
</tr>
<tr>
<td>Using assessment in instruction:</td>
<td>□ little or none □ used sporadically □ used regularly □ fully integrated (formative assessment)</td>
</tr>
</tbody>
</table>

Student Talk (final number on stopwatch cumulative across lesson) in seconds: ______________________
APPENDIX E: OPERATIONAL DEFINITIONS QUICK REFERENCE SHEET
**Definitions: Quick Reference Sheet**

**Closed-ended question:** A content question which has restricted parameters, expecting one possible response as its acceptable answers. A question which constrains a student's response such as test questions, yes–no questions and forced choice questions.

**Example:** "Which is more essential: a computer or a phone?"
- "Does anyone know what technology is?"
- "Do you have any ideas?"
- "Do you want to expand on that?" or "Is that right?"
- "Anything else?" or "Anybody else?"
- "How many of you think so?"

**Non-examples:** "What knowledge is required to create a computer?" (open)
- "How do you define technology?" (open)

**Do not code:** "...Ok?" "...Right?" "...Alright?" "...Huh?"
- When a teacher calls on a student and says “Yes...”
- Also, do not code statements made with questioning intonation.

**Open-ended Question:** A content question to which a number of different answers would be acceptable. A content question which has no parameters and does not constrain student's response.

**Example:** "What is your personal definition of technology?"
- "How does technology help us in our everyday lives?"
- "What else?"
- "How else?"

**Non-examples:** Yes-no questions that the teacher intends as open-ended questions are still coded as closed-ended questions: "Can anyone tell me what technology is?" and "Do you have any examples of technology?" are yes/no questions and should be coded as closed-ended.

**Open-ended Question Plus:** A content question that asks a student to extend, produce, or combine ideas to generate new ideas (related to Bloom's highest Cognitive domain –Creating).

**Example:** "How would you prove that, using evidence from our activity in class today?"
- "How could you use today's technology to create transportation for people with limited visibility?"
- "How do you think people will use this tool in 50 years?"

**Questioning - Points to Remember:**
- A question is a question, so count them all, but only content-related questions are considered.
- Wait for the end of question; sometimes questions change midway.
- If the teacher asks a string of questions, code each complete one individually.
- The type of question is determined by the literal response requested by the teacher. Do not attempt to identify teacher's intent (e.g., "Does anyone else have an idea?" is a yes/no question even though the teacher is soliciting more responses. Student can choose to simply say "No").

**Questioning - Background:** Checking for student understanding during a lesson using informal assessment techniques (METEX TW13). Eliciting and interpreting individual students’ thinking (High-leverage Practice #3).

**Student Talk:** Teacher is leading or facilitating a discussion among students; one or more students are contributing and focus is on content. Also occurs during student small-group work in which students have the opportunity to discuss content. A majority of students should be discussing content in their groups.

**Examples:** Student responding to a question. Student asking a question. Student commenting on a peer’s idea. Small group discussions. If half of the groups are discussing content – count it.

**Non-examples:** Students' off-task side conversations. Student discussion about topics not related to content.
- "When is the next school holiday?"

**Student Talk - Points to Remember:**
- If student and teacher talk is happening - count it as student talk (e.g. if teacher is meeting with other groups while students are discussing content),
- You will need a stopwatch to measure the amount of time students spend talking. Use the start/stop function to gather cumulative amount of time students spend talking on task during the interval. At the end of the observation, you should have the total time across all five intervals.

**Student Talk - Background:** In a whole class format, the entire class is talking, listening, or working together (METEX TW01). Teacher is leading or facilitating a discussion among students; several students are contributing, and teacher seeks to involve multiple children in listening and speaking. Discussion involves sustained interaction and is focused on a text, issues, problems, or questions where the goal is to work on collective understanding, analysis, or solutions (METEX TW08). Leading a whole class discussion (High-leverage practice #2). Eliciting and interpreting individual students’ thinking (High-leverage Practice #3).

**Specific Affirmation:** Teacher's positive verbal affirmation about what a student or group of students did or said related to content in a single episode within class (multiple statements about the same episode count as 1 occurrence of affirmation).

**Examples:**
- "I like the way you described cell structure."
- "Yes, Angie said we should add more items here and that is correct."
- "Nice observation."
- "Good work" or "Job well done."
- "You got it!" or "I appreciate that." ("it" or "that" refers to content).

**Non-examples:**
- "Good." "Uh-huh." "Yes. "OK." (not specific)
- "You are so smart" (not related to content). "OK, very good. Thank you" (not specific).
- **Do not code:** Teacher shrugs, teacher nods, or any type of only non-verbal response. We are looking for verbal responses only.

**Specific Affirmation - Background:** Implementing norms and routines for academic discourse and work. The teacher is making explicit, commenting on, reviewing or reinforcing, or teaching students specific norms and routines for academic discourse. The teacher might praise or reinforce the use of an established norm or routine (METEX -TW07). Providing oral and written feedback to students on their work (High-leverage practice #16).

Revised: 10/1/2013