Measuring Quality Technology Integration in Science Classrooms

Caroline E. Parker¹ · Cathlyn D. Stylinski² · Christina R. Bonney³ · Jacqueline DeLisi¹ · Joseph Wong¹ · Cassie Doty²

Published online: 16 July 2019

© The Author(s) 2019

Abstract

Researchers, evaluators, and practitioners need tools to distinguish between applications of technology that support, enhance, and transform classroom instruction and those that are ineffective or even deleterious. Here we present a new classroom observation protocol designed specifically to capture the quality of technology use to support science inquiry in high school science classrooms. We iteratively developed and piloted the Technology Observation Protocol for Science (TOP-Science), building directly on our published framework for quality technology integration. Development included determining content and face validity, as well as the calculation of reliability estimates across a broad range of high school science classrooms. The resulting protocol focuses on the integration of technology in classrooms on three dimensions: science and engineering practices, student-centered teaching, and contextualization. It uses both quantitative coding and written descriptions of evidence for each code; both are synthesized in a multi-dimensional measure of quality. Data are collected and evaluated from the perspective of the teacher's intentions, actions, and reflections on these actions. This protocol fills an important gap in understanding technology's role in teaching and learning by doing more than monitoring technology's presence or absence, and considering its integration in the context of the Next Generation Science Standards science and engineering practices. Its applications include research, evaluation, and potentially peer evaluation.

Keywords Classroom observation protocol · Technology integration · High school science · Science inquiry · Evaluation

Introduction

Today's science teachers are challenged with immersing their students in the practices of science inquiry (DiBiase and McDonald 2015; Herrington et al. 2016). This includes having students ask testable questions, design related studies, collect and interpret data, engage in argumentation from evidence, connect findings to other research, and more. The importance of these practices is readily apparent in the Next Generation Science Standards (NGSS), which address eight science and engineering practices (SEPs) in the context of disciplinary core ideas and crosscutting concepts (National Research Council 2012).

At the same time, digital technology has become a commonplace learning and organizational tool in many schools. For example, according to a 2017 report on connectivity in education, over 40 governors invested resources in improving technology

infrastructure in their states, resulting in over 30 million students in 70,000 schools gaining broadband access between 2013 and 2016 (Education Super Highway 2017). Indeed, technology applications have the potential to support and extend student application and learning of science practices. This support or extension can occur with the use of everyday applications, such as Internet searches and word processing, as well as with applications designed specifically for learning environments, such as classroom management systems (Herrington and Kervin 2007). It can even take advantage of the same applications used by science professionals, such as delineating watershed boundaries, mapping land use, and measuring water quality with GIS to address questions about factors impacting stream health (e.g., Stylinski & Doty 2013). These tools have the potential to support, enhance, and even transform the learning experience if used effectively. However, they can also be used in ineffective ways that confuse, distract, and interfere with learning.

Researchers and evaluators need instruments that will enable them to describe the effective integration of technology into science classrooms to understand this changing classroom landscape and better support teachers. Observation protocols provide researchers and evaluators with critical tools for examining pedagogical approaches and curricular resources, including technology hardware and software, and their relationship to student learning (e.g., Bielefeldt 2012). A recent review identified 35

✉ Caroline E. Parker
cparker@edc.org

¹ Education Development Center, Waltham MA, USA

² University of Maryland Center for Environmental Science, Frostburg, MD, USA

³ Brown University, Providence, RI, USA

observation protocols aimed at describing teachers' instructional practice for science classrooms; however, less than a third accounted for technology use (Stylinski et al. 2018). Of these, most only documented the presence of technology. Fewer than half of those that included technology described the quality of technology use by looking at the integration of technology with pedagogical practices. None comprehensively captured the quality of technology use in the specific context of instruction focused on science disciplinary practices.

In this paper, we present the Technology Observation Protocol for Science (TOP-Science), a new classroom observation protocol designed specifically to capture the quality of technology use, to support science inquiry in high school science classrooms. We describe how we determined face validity and construct validity through (a) a careful alignment of the Protocol with our published framework of technology integration in science classrooms and (b) iterative pilot testing that included expert review and inter-rater reliability.

Theoretical Foundation of the Protocol

The Protocol builds on a published theoretical framework that consists of four dimensions: technology type, science and engineering practices (SEPs), student-centered teaching, and relevance to students' lives (Parker, Stylinski, Bonney, Schillaci & McAuliffe 2015). These dimensions are critical to high-quality technology integration, and their inclusion in our observation protocol supports the instrument's content validity. In this section, we describe each of these dimensions and their connection to the Protocol presented in this paper.

Technology Type

Digital technology in classrooms can be defined as computer hardware, software, data, and media that use processors to store or transmit information in digital format (Pullen et al. 2009). In our framework, we expanded on this definition by describing three types based on the use of technology. Specifically, some technology applications are primarily used in *instructional* settings, such as assessment tools (e.g., clicker software) and online course management (e.g., PowerSchool, Google Classroom). Other technology applications are most commonly used in *STEM workplace* settings (e.g., computer modeling, image data analysis). Finally, *ubiquitous* technology applications are used in many settings, including in the classroom, workplace, and elsewhere in daily life (e.g., word processing, Internet search engines). The designation of a technology application to the instructional, ubiquitous, or STEM workplace category may change over time as technologies evolve and are adapted to new settings. For example, Microsoft Excel was originally developed for businesses, universities, and other workplace environments; but it now appears regularly in many different settings such as K-12 schools and homes. In the Protocol, we

treated technology type differently from the other three dimensions. As described later, we consider the integration of technology with each of the other three dimensions, and we simply had observers record technology type as a way to capture the diversity of technology applications in the classroom setting.

Authentic Science and Engineering Practices

Real-world technology applications offer the potential to help teachers align with Next Generation Science Standards (NGSS), which weave eight science and engineering practices with cross-cutting concepts and core ideas (NRC, 2012). These eight practices provide a thoroughly vetted description of science practices relevant for the classroom, and thus, they form one of three central dimensions of our observation protocol, which, as noted, are considered in the light of technology integration. Specifically, users consider the integration of technology with this description of authentic science inquiry.

Technology, regardless of type, can scaffold students' understanding and learning during science-inquiry-based activities (e.g., Devolder et al. 2012; Rutten et al. 2012). Classroom-based technology applications, both those designed specifically for instruction and those drawing from science workplace settings, can help teachers engage students in real-world science and engineering pursuits—deepening their understandings of science inquiry. This linkage between science and technology expands options for classroom science inquiry such as allowing students to observe phenomena otherwise not possible in a classroom setting (D'Angelo et al., 2014; Smetana and Bell 2011). For example, computer simulations, such as those offered by *SimQuest*, scaffold students' understanding by asking them to conduct multiple trials of an experiment to explore scientific principles and the effects of changing certain variables on outcomes of interest (De Jong 2006).

Student-Centered Teaching

Student-centered teaching involves instructional practices that allow students to have a more self-directed experience with their learning. The Nellie Mae Education Foundation (NMEF) defines student-centered teaching as encompassing personalized instruction, student autonomy or ownership of the learning process, and competency-based instruction (NMEF and Parthenon-EY 2015; Reif et al. 2016).¹ Our Protocol breaks student-centered teaching into three aspects closely aligned with the Nellie Mae definition: personalized instruction, autonomy, and competency-based teaching. *Personalized instruction* occurs when students' individual skills, interests, and developmental needs are taken into account for

¹ Nellie Mae's definition includes a fourth component of student-centered teaching—"learning can happen anytime, anywhere"—however, because this component addresses learning that happens outside the classroom, it extends beyond the scope of the Protocol.

instruction. Students have *autonomy* over their own learning when they have frequent opportunities to take ownership of their learning process and have a degree of choice. *Competency-based* teaching is defined as teaching that progresses based on students' mastery of skills and knowledge, rather than being dictated by how much time has been spent on an activity or topic (NMEF and Parthenon-EY 2015).

Many teachers have capitalized on the increased availability of technology by using digital devices and technology resources to encourage student-centered learning through increased personalization, individual pacing, and student ownership (Reif et al. 2016). Maddux and Johnson (2006) note that technology implementation strategies are most successful when they are more student-centered and move away from teacher-directed instruction. The use of technology in student-centered instruction can allow students to learn in ways that would not otherwise be possible without the technology. For example, online modules may facilitate personalized learning by allowing students to learn at their own pace, using various tools and resources to support learning and providing an appropriate level of scaffolding (Patrick et al. 2013). In addition, a learning management system such as Blackboard or Moodle might facilitate competency-based learning by helping to document students' progress toward demonstrating competencies, which might otherwise be too burdensome without technology (Sturgis and Patrick 2010). Technology might also support student-centered teaching such that students may have autonomy over what kind of technology to choose for an activity, allowing technology to be a "potential enabler" of student-centered learning (NMEF and Parthenon-EY 2015, p. 6).

Contextualizing Learning for Relevance

The use of technology in classrooms enables teachers to make learning more relevant to the world in which students live (Brophy et al. 2008; Hayden et al. 2011; Miller et al. 2011). Since the 1990s, there has been an emphasis on trying to make science more relevant to students (Fensham 2004). The proponents of making science content more relevant to students are generally those who argue that such instructional practices lead to student-centered learning and increase student motivation for learning science (Moeller and Reitzes 2011; Ravitz et al. 2000; Stearns et al. 2012).

One of the more common ways to increase the relevance of science to students includes making connections between the content and the real world in which students live. This idea of increasing relevance is also sometimes referred to as authentic learning (Lombardi 2007; Reeves et al. 2002). Grounding lessons in the local geographic context, making connections to youth culture, or matching classroom tasks and activities with those of professionals in practice are all ways in which learning can be made more relevant or authentic to students. Technology can support and facilitate this type of authentic learning via such tools as web resources and search engines, online discussions or emails, online simulated environments, digital photography, or

voice recorders (Herrington and Kervin 2007). Technology can facilitate authentic learning by using animations, sounds, and images, allowing students to engage more easily with the material on a conceptual level. For example, high school students in Australia used web resources in conjunction with a field trip to the Sydney Olympic Park, to research a problem the park was having with mosquitoes and rats attracted by food scraps, standing water, and trash left behind by visitors. The "geography challenge" was presented to students using animated scenarios involving the mosquitoes and rats affecting visitors' experiences at the park, and invited them, as "geography consultants," to investigate and suggest recommendations on ways to mitigate the problems and "restore an ecological balance to the area" (Brickell and Herrington 2006, p. 536). Students conducted web research before visiting the site, and presented a report with their findings to "park authorities."

Summary

While the three central dimensions of the Protocol—SEPs, student-centered teaching, and contextualization—provide a useful frame for describing the integration of technology with key pedagogical practices, they are not mutually exclusive. The SEPs encourage making science learning authentic; that is, contextualizing learning, through reflecting the real work of scientists and including questions of interest to the field of science. The SEPs also encourage autonomy, a key facet of student-centered teaching, particularly through opportunities to define researchable questions.

Development of the Protocol

The development and testing process of the TOP-Science Protocol reflects that undertaken by similar observation tools (e.g., Dringenberg et al. 2012; Sawada et al., 2002). As noted, we built the initial draft of the Protocol on the theoretical framework previously described and after review of similar observation protocols. We refined this first draft through preliminary piloting and based on feedback from several experts in instrument development. We then conducted four rounds of iterative pilot testing of the Protocol (Table 1), each followed by feedback-driven revisions as well as additional expert review.

In the first two rounds, individual observers completed feedback sheets that were then analyzed by the team and used to revise the Protocol. In the third and fourth rounds, observer pairs completed observations, compared results, and provided feedback based on the comparisons. Our research team, along with additional trained observers, conducted 66 observations of diverse science classes in 26 high schools across seven states. The seven states included in our study represent a broad range of areas from suburban, urban, and rural locations, as well as from the west coast, New England, and Mid-Atlantic regions of the country. The classrooms observed ranged from ninth to

twelfth grades (including mixed grades). They occurred in a diversity of school settings including low-resourced and high-resourced schools, public schools, and charter schools, and included some that self-identify as STEM schools.

The final Protocol consists of four parts (Appendix A): (1) pre-observation teacher questions to understand teachers' intention for technology integration in their class (Fig. A1), (2) observation sheets where codes and field notes are recorded for the four framework dimensions (Fig. A2) in 10-min intervals, (3) post-observation teacher questions to gather the teacher's reflection on their lesson (Fig. A3), and (4) a code summary sheet that produces a multi-dimensional measure of technology integration quality (Fig. A4).

A detailed Reference Guide walks the observer through each step in the Protocol process, and provides all instructions needed to implement the Protocol. For each of the three categories, there is a detailed description of what observers should look for under each, with descriptors and examples for each category and technology level. Table 2 provides sample technology codes and evidence text for each category, demonstrating different levels of technology integration for similar lessons.

The pre- and post-observation teacher questions offer insight into teacher technology integration goals and reflections on implementation. The observation sheet provides a quantitative metric and qualitative description of technology integration focused on teacher practices. Observations occur in 10-min intervals; within each interval, observers complete codes for each dimension and include field notes providing evidence for each of the codes. The code summary sheet includes a template to synthesize results across the 10-min intervals, also incorporating teacher responses to the pre- and post-questions. Four key features emerged from the protocol development process. Each is described below.

Multi-Level Coding System Coupled with Descriptive Text

To address the challenging process of developing meaningful and user-friendly codes that could describe the presence of each of our framework dimensions, including the extent to which technology was integrated, we tested a range of options, beginning with the most basic check for presence or absence of the category and technology. As we had already found in our literature review, however, the presence of technology is not the same as quality integration of that technology. To address this, the next iteration applied the terms absent/low/medium/high for each category. Ultimately, across the four rounds of pilot testing, we found that observers found it easiest to describe each dimension as either not present, incidentally present, or embedded in the lesson, and the level of technology integration in each of the dimensions as minimally integrated, partially integrated, or fully integrated.

Our initial draft of the Protocol utilized tallies to indicate the frequency of observed technology use and relevant instructional practices. After the first round of pilot testing, we realized that

tallies of codes did not paint a rich enough picture of the quality of technology integration. Therefore, we introduced field notes to complement these tallies (as written text alone would present too onerous a task for our 10-min observation intervals), and piloted this structure in subsequent rounds. Specifically, observers supply written text describing evidence for each chosen code. Later, they combine these data to describe the quality of technology integration in the classroom (see “Synthesis of Data into a Quality Measure” for more information). Thus, the Protocol does not limit observers to a checklist or set of codes, but instead couples quantitative and qualitative data. This pairing provides the necessary flexibility to capture multiple and diverse ways that the same technology can be used for different purposes. For example, after selecting the appropriate technology-practice integration that involves GIS (a workplace technology application), the observer could describe evidence of how the teacher used GIS to help students brainstorm environmental science research questions, collect geospatially based data, present geospatially explicit findings, develop maps to share with community members, and more.

Integration of Technology with the Three Dimensions

In our initial attempts to operationalize the four dimensions from our Framework into an observation protocol, we recorded the technology use category separately from the three other categories (SEPs, student-centered teaching, and contextualization). We quickly realized, however, that in order to truly assess the quality of technology use, we had to more clearly define technology integration within the three other dimensions. Coding technology use separately from the dimensions and then combining the codes post hoc was artificial and did not accurately reflect a true measure of integration. To address this challenge, as described earlier, we replaced it with a system of coding the degree of technology *integration* with each of the components, not just technology use. For example, in one pilot classroom, students studied chemical reactions by choosing a reaction of interest and designing a demonstration that would illustrate this reaction during an open house event. With Chromebook laptop computers, they researched chemical reactions and potential risks and identified required supplies. Some students also used Google Docs to record this information and plan their demonstration in collaboration with other students. With our coding system, we were able to document the extent to which the activity was student-centered (i.e., *embedded* student centeredness, due to the level of autonomy students had over their proposal and research methods), and to what extent technology use was integrated in those student-centered practices (i.e., *partially integrated*, because students had options about the kind of technology to use in the activity, and the technology enhanced the student-centeredness of their demo designs). For this particular example, it would not be considered *fully integrated* because, while the use of technology enhanced the student-centeredness of the activity, the level of student centeredness was not necessarily contingent on the use of technology. Students would still have been able to choose

Table 1 Protocol pilot testing

Protocol round	Date range	No. of teachers (no. of observations)	No. of observers	No. of high schools	No. of states
1	Oct–Dec 2015	8 (8)	6	6	3
2	Apr–May 2016	7 (7)	5	6	4
3	Oct–Dec 2016	10 (10)	6	8	3
4	Apr–Jun 2017	24 (41)	8	17	6

Table 2 An example from the Reference Guide that illustrates three levels of technology integration for the three central dimensions of the protocol

Embedded SEP (planning and carrying out an investigation)

Context: Students develop hypotheses, design, and carry out an investigation about the relationship between certain variables of the rates of a reaction. When present, technology enables increasing measurement sophistication and supports interpretation and analysis.

- 0: no tech Students create hypotheses about which variables will affect the timing of chemical reactions (baking soda and vinegar) and then design and implement a study to test their hypothesis. Chemical reaction time is visually determined when bubbles stop forming, and measured using the classroom clock. *No tech: students observe changes in reaction time visually and use classroom clock to measure time.*
- 1: minimally integrated Students use smartphone to record time more accurately. *Minimal: students observe changes in reaction time visually and use phones to provide more accurate time measurement.*
- 2: partially integrated Chemical reaction time is measured using a Vernier probe to assess change in pressure in a test tube and document differences in reaction times. *Partial: students measure reaction time using Vernier probes, use phones to provide accurate time measurement.*
- 3: fully integrated Vernier probes are connected to computers that record the data to an excel file and allow students to graph rates of change by dependent variable (e.g., amount of baking soda, temperature of vinegar). *Full: Probes are connected to laptops, digital data are recorded and used to calculate rates of change.*

Embedded personalized instruction

Context: Students learn about kinetic energy using kinetic graphs. When present, technology promotes personalized, individualized instruction and learning opportunities.

- 0: no tech Students complete worksheets asking them to compare and contrast various kinetic graphs (showing relationships between time and velocity). Teacher differentiates assignment by providing graphs of varying levels of complexity for different groups of students, and worksheets are completed in pairs or small groups that may collaborate with each other. *No tech: printed material*
- 1: minimally integrated Students complete a paper-based lab in which they use motion detectors to replicate kinetic graphs by moving their own bodies. However, the software is not flexible enough to effectively accommodate different levels of complexity. *Minimal: Students use motion detectors, but the lack of functionalities of the technology does not allow the teacher to provide effective student centered teaching.*
- 2: partially integrated The teacher differentiates the assignment by providing electronic data, allowing students to manipulate the data at varied levels of complexity. *Partial: The use of the motion detectors and digital data and graphs allow the students and the teacher to adjust some of the lab components for student-centered teaching.*
- 3: fully integrated Students complete and submit the lab through a learning management system (LMS), which allows student groups to collaborate and discuss various aspects of their graphs online. Teacher also grades the labs and provides feedback through the LMS. *Full: The use of motion detectors and the LMS to collaborate and complete the assignment enable a level of personalized learning that would otherwise be very difficult to accomplishment without the technology.*

Embedded contextualization

Context: Water conservation lesson using local watershed organization's website. When present, technology provides students access to more real-life data and context.

- 0: no tech Teacher provides printed reading material from local watershed organization's website. Students can see the local watershed data on the printed material and look for patterns in the data. *No tech: printed material.*
- 1: minimally integrated Teacher provides access to local watershed organization's website. Students can see the watershed data online and look for patterns in the data. *Minimal: students read teacher-provided information online which is similar to analog experience of reading on paper.*
- 2: partially integrated Students download data for stream health and land cover in upstream watershed for their local stream site. They create scatter plots of the data to examine the relationship between stream health and land cover (e.g., percent impervious surface). *Partial: downloading, manipulating, and analyzing local stream site data allows students to better understand the information and increase interest in watershed health and needed conservation efforts.*
- 3: fully integrated Students examine their site compared to others along the trend line. Students also gather information on cleanup efforts from the EPA web site. *Full: Students are able to obtain information not only of their local sites, but also wider geographical area so they can expand their analyses to better learn and understand how they local efforts compare with those in other regions.*

reactions to study, design their demonstration, and collaborate with other students without technology.

Focus on Teacher Intentions, Actions, and Reflections

The Protocol explicitly focuses on observing the teacher and his/her actions and statements during the class period, and does not include codes for student actions. While relevant student actions and statements can be recorded in the evidence section, the observer does not attempt to capture students' responses, as it is difficult to follow all of the students and know how they are responding without talking with them. In addition, their responses may be varied and confounded with factors unrelated to the teacher's actions or intentions.

This focus on the teacher requires that the Protocol also include an opportunity for the teacher to describe their intentions and goals before class and their reflections after class, which provides context for the observation. Pairing these insights with the teacher's actions during the lesson offers a more encompassing perspective for the observer to consider when evaluating the quality of technology integration.

Synthesis of Data into a Quality Measure

At the conclusion of the observation, the observer has a collection of 10-min "interval sheets" with codes for the three categories and evidence (written text) of each code. In order to aggregate these data into a meaningful picture of the quality of technology integration across the full lesson, we developed a post-observation summary that combines the quantitative and qualitative information from each interval sheet, as well as information from the pre- and post-observation teacher questions. This summary consists of an average and the highest-level technology codes for each category (science and engineering practices, student-centered teaching, and contextualized teaching) that is supported by relevant qualitative evidence from field notes.

Examining Validity and Reliability

We examined the Protocol's validity throughout the development process, and inter-rater reliability of the quantitative codes was calculated using results from the fourth round of iterative piloting.

Instrument Validity

We addressed content validity of the Protocol by aligning it with an existing published theoretical framework and its literature base, and comparing the Protocol with similar protocols. The development of the Protocol consisted of an iterative process of collecting data from a variety of classrooms, reviewing interpretations and definitions of items and constructs, and reflecting on the literature and previously developed protocols. In this way, we established categories and indicators that reflect the range of practices of

technology integration (American Educational Research Association 2014). We established face validity of the Protocol with an expert panel who reviewed and submitted written feedback on three drafts of the Protocol and supporting materials. In particular, the expert review confirmed the categories and ratings of technology integration and their feedback resulted in definitions that were more closely aligned with understandings of technology across a broader range of classrooms. We also made refinements to the reference guide and other support materials after each review.

Inter-Rater Reliability

We calculated inter-rater reliability using results of our fourth implementation (see Table 1). We observed 24 teachers in 17 schools in six states (MA, RI, CA, MD, WV, and OR). Grade levels ranged from 8th to 12th grade; schools were in urban, suburban, and rural areas, with a range of demographic characteristics.

Eight observers completed a total of 41 classroom observations, with two to three observers in each classroom. Each observer was observed from three to 10 total class periods. To the extent possible (given geographic limitations), different sets of partners observed different classrooms. After each observation, individuals completed the Protocol and recorded their results online. The pairs then compared results using a worksheet that helped to reconcile codes and identify feedback for further revisions. Inter-rater reliability was conducted on the 10-min intervals within each observed lesson, resulting in a total of 215 paired intervals. Krippendorff's alpha was used to determine reliability; Krippendorff's alpha is appropriate for this analysis because it "generalizes across scales of measurement; can be used with any number of observers, with or without missing data" (Hayes and Krippendorff 2007, p. 78). Alpha values ranged from 0.40 (student engagement) to 1.0 (STEM workplace technology, which had 100% agreement) (Table 3).

Given these results, we made a series of adjustments to the Protocol and to the definitions and guidance included in the Reference Guide. Specifically, the code for student engagement had the lowest alpha estimate (0.40). That code was removed rather than revised, in order to retain the focus on teacher activities rather than student activities. Student-level observations, including engagement as well as academic achievement, were not within the scope of this Protocol.

The four codes for technology type (instructional, ubiquitous, STEM instructional, and STEM workplace) had alpha estimates ranging from 0.46 to 1.0. Given this range and elimination of technology type from the post-observation code summary, the four codes were removed and replaced with a space for a qualitative description of the technology type. The Reference Guide provided additional support on classifying technology types into each category.

We retained six codes in the final Protocol with Krippendorff's alpha estimates from 0.51 to 0.78. Four codes (SEP level, SEP technology, student-centered level, student-centered technology)

had acceptable values above 0.68, while contextualization level and contextualization technology were below 0.65. Given that the values were lower than desired and that the post-observation checks between observers resulted in recommendations for clarifications to the Reference Guide, we revised the Guide to address the recommendations, including refining descriptions for each code.

Discussion and Conclusion

The TOP-Science Protocol successfully addresses the gap around observing and documenting the quality of technology integration in high school classrooms (Bielefeldt 2012; Styliniski et al. 2018). In particular, it offers a tool that captures how teachers utilize technology to support teaching and learning of scientific pursuits and practices. The Protocol's focus on integration of technology in three key dimensions expands other observation protocols by moving beyond simply registering technology use, and by targeting key categories that are specifically relevant for science classrooms. Its use of a limited number of descriptors, a straightforward three-level coding system, and the integration of descriptive text justifying each code choice provides a manageable frame for each 10-min observation interval. Together, these strategies direct the observer's attention on teacher intentions, actions, and reflections without overwhelming them with trying to interpret student actions and perspectives. The final step of summarizing interval data by category offers an overall quality measure of technology use, while still providing the option to substitute any other synthesis calculation.

The Protocol has some limitations that should be considered in any application and could be addressed in future revisions. First, its structure suggests that full integration of technology in all three categories points to the highest quality. However, such extensive integration across an entire class period is likely not realistic or even desirable, as technology

is not appropriate for all learning outcomes and activities. Our Protocol does not account for this variation in appropriate use of technology. That is, it is effective at identifying high-quality technology use, as well as incidental applications of technology, but does not elucidate when technology might or might not be the most appropriate tool to support learning. Thus, further work is needed to explore how to determine technology use relative to technology benefits and affordances. Second, the inter-rater reliability scores in some areas were lower than desired; revisions were made to both the coding structure and the Reference Guide to address the issues we identified, and additional inter-rater reliability using the updated Protocol and Reference Guide is needed to examine additional evidence of the Protocol's reliability. Finally, the Protocol does not explicitly examine the impact of observed teacher practices on student outcomes. Our Protocol allows a view into teachers' intention and actions; however, the final arbiter of high-quality instruction must always be student outcomes. A separate evaluative tool is needed to identify or measure these outcomes—ideally in ways that reflect the instructional dimensions relevant for science classrooms (i.e., science practices, student-centered teaching, and contextualization).

Despite these limitations, our Protocol successfully frames the quality of technology integration from the perspective of the teacher and in the context of today's science classrooms and is being used in research (Marcum-Dietrich et al. 2019). As described above, the Protocol was developed with care to determine both face validity and construct validity, and adequate reliability has been demonstrated. Additional studies of inter-rater reliability using the final version of the Protocol will serve to further assess its strength as a measure of the quality of technology integration. The inclusion of a post-observation code summary adds an important component not included in many other protocols, by helping researchers to synthesize the observations of what is observed in the classrooms using the Protocol's theoretical framework.

Table 3 Krippendorff's alpha values for Protocol codes

Code	K-alpha
Instructional technology	0.56
Ubiquitous technology	0.46
STEM instructional technology	0.77
STEM technology	1.0
Student engagement	0.40
SEP level	0.69
SEP technology	0.78
SCT level	0.73
SCT technology	0.76
Contextualization level	0.51
Contextualization technology	0.57

8 observers, 215 observations

Acknowledgments We thank all of our study teachers who donated their time and opened their classrooms in support of this research.

Funding This material is based upon work supported by the National Science Foundation under Grants #1438396 and 1438368. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Compliance with Ethical Standards

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Conflict of Interest The authors declare that they have no conflict of interest.

Appendix A. Technology Observation Protocol for Science

Observer and Class Information	
Observation code	
Observer name	
Teacher name	
School name	
Course name	
Number of students	
Student grade level	
Observation date(s)	
Clock time for observation	
Pre-Observation Teacher Questions	
1. What are the big ideas (e.g., content, skills, competencies) that students will understand from this lesson or unit?	
2. Please provide a description of how you will use technology to support understanding these big ideas.	
3. What specific technology hardware and software will you or your students be using (e.g., learning management systems, Vernier probes, Excel, GPS/GIS)?	
4. If this lesson is part of a larger unit or project-based learning, briefly describe what you have done previously and what you plan to do later in terms of science content, science and engineering practices, and technology use.	
5. Why have you chosen to use technology with this lesson?	
6. To what extent is this lesson personalized or differentiated for students based on their developmental needs, skills, and interests?	
7. Please provide any materials you can before the class (e.g., attachments of handouts you will be using, web links), or provide the observer with the handouts when they get to the class.	
8. Is there anything else we should know about the selected lesson/unit or your students?	

Protocol citation: Stylinski, C., Bonney, C., Parker, C.E., Wong, J., Delisi, J., & Doty, C. (2018). Technology Observation Protocol (TOP-Science). University of Maryland Center for Environmental Science; Frostburg, MD and Education Development Center; Waltham, MA. <http://topscience.edc.org/index.php/resources/>

Fig. A1 Part 1. Pre-observation information (*complete one for each observation*)

Interval Description		Observation code:	Interval No.
Time period covered by this sheet (10 minute increments)	Time started: _____ Time finished: _____		
Activities Observed e.g., lecture, discussion, individual seatwork, hands-on activity, group work, assessment, test preparation, homework review			
Hardware Types Used e.g., desktop/laptop computers, mobile devices, smartboard, probes, graphing calculator			
Software or Application Used e.g., word processing, web browser			
Technology Application Types	☐ INSTRUCTIONAL ☐ UBIQUITOUS ☐ STEM WORKPLACE	Detailed Description of Technology Application	
Interval Evidence and Coding			
CATEGORY (may check more than one)	LEVEL ^a (select one)	TECH INTEGRATION ^b (select one)	EVIDENCE FOR CATEGORY, LEVEL, & TECH INTEGRATION CODES
Science and Engineering Practices ☐ Asking questions & defining problems ☐ Developing & using models ☐ Planning & carrying out investigations ☐ Analyzing & interpreting data ☐ Using mathematics & computational thinking ☐ Constructing explanations & designing solutions ☐ Engaging in argument from evidence ☐ Obtaining, evaluating, & communicating information	☐ 0 ☐ 1 ☐ 2	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ N/A	
Student Centered Teaching ☐ Autonomous ☐ Personalized ☐ Competency-based	☐ 0 ☐ 1 ☐ 2	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ N/A	
Contextualized Teaching ☐ Youth Experience ☐ Science careers or work ☐ Local/geographic context	☐ 0 ☐ 1 ☐ 2	☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ N/A	
Observer Reflection			

^aLEVEL CODES: 0 = No evidence • 1 = Incidental • 2 = Embedded

^bTECH INTEGRATION CODES: 0 = No Tech • 1 = Minimally integrated • 2 = Partially integrated • 3 = Fully integrated • Not Applicable (if category code is N or None)

Protocol citation: Styliniski, C., Bonney, C., Parker, C.E., Wong, J., Delisi, J., & Doty, C. (2018). Technology Observation Protocol (TOP-Science). University of Maryland Center for Environmental Science; Frostburg, MD and Education Development Center; Waltham, MA. <http://topscience.edc.org/index.php/resources/>

Fig. A2 Part 2. Interval sheet (complete one for every 10-min interval during the observation)

Post-Observation Teacher Questions	
1. How representative of your usual technology use was the observed lesson/unit?	
2. In what ways did the technology observed help students understand the big ideas? Did the lesson go as planned in terms of students' understanding these ideas and engaging with the technology? If the lesson did not go as planned, what changes would you make?	
3. Is there anything you would have done differently with regard to technology use?	
4. How would the lesson be different if you taught it without technology? (If you have previously taught this or a similar lesson without technology, please describe the differences.)	
5. Is there anything else we should know about the observed lesson/unit?	
6. Other follow up questions. (Observers can ask any additional clarifying questions).	

Protocol citation: Styliński, C., Bonney, C., Parker, C.E., Wong, J., Delisi, J., & Doty, C. (2018). Technology Observation Protocol (TOP-Science). University of Maryland Center for Environmental Science; Frostburg, MD and Education Development Center; Waltham, MA. <http://topscience.edc.org/index.php/resources/>

Fig. A3 Part 3. Interval sheet (complete one for every 10-min interval during the observation)

© 2018 EDC All Rights Reserved

Observers can use the post-observation summary to synthesize the observations from the activity sheets. Much of the information can be copied directly from the interval sheets (green cells), and some needs to be synthesized by the observer from multiple data sources (blue cells). These cells require the observer to summarize from specific data sources and should be completed as soon after the classroom observation as possible. Note: while there are numeric values provided for each of the category levels and tech levels, the qualitative evidence will generally be an essential component. That is, the protocol authors do not recommend using the quantitative values alone in calculating and/or describing the quality of technology integration in the classroom.

Notes:

Green=cell that should be copied directly from interval sheets

Blue = cell that observer needs to complete synthesizing from multiple data sources

Content	Data source	1-3 sentence description
Name of course and grades of students	Pre-questions	
Big ideas students are working on and how teacher has designed technology to support those ideas	Pre- and post-questions	
Specific technology use in the lesson (e.g., hardware, software, technology application types, activities students engaged in)	Pre-questions and interval sheets	
How this lesson fits in with other technology use; coherence of content and technology use	Pre-questions and interval sheets	
Challenges in lesson and how teacher addressed them (e.g., technology glitch, classroom management)	Interval sheets and post-questions	

	SEP Level	SEP Tech	How technology contributed to student use and/or understanding of SEPs (Key evidence and 1-3 examples; note specific SEP as relevant)
Interval sheet 1			
Interval sheet 2			
Interval sheet 3...			
Average science and engineering practices score			

	Student-centered teaching (SCT) Level	SCT Tech	How technology contributed to student-centered teaching (Key evidence and 1-3 examples; note autonomy, personalization, competency-based as relevant)
Interval sheet 1			
Interval sheet 2			
Interval sheet 3...			
Average student-centered teaching score			

	Context Level	Context Tech	How technology contributed to contextualization (Key evidence and 1-3 examples; note youth-focused, science-focused; time/place relevance as relevant)
Interval sheet 1			
Interval sheet 2			
Interval sheet 3...			
Average contextualization score			

Highest Interval sheet # ____	Level	Tech	How technology contributed to each (Key evidence and 1-3 examples)
SEP			
SCT			
Contextualization			
Summary			

© 2018 EDC All Rights Reserved

Protocol citation: Stylinski, C., Bonney, C., Parker, C.E., Wong, J., Delisi, J., & Doty, C. (2018). Technology Observation Protocol (TOP-Science). University of Maryland Center for Environmental Science; Frostburg, MD and Education Development Center; Waltham, MA. <http://topscience.edc.org/index.php/resources/>

Fig. A4 Part 4. Post-observation summary

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

American Educational Research Association, American Psychological Association, National Council on Measurement in Education, &

Joint Committee on Standards for Educational and Psychological Testing. (2014). *Standards for Educational and Psychological Testing*. Washington, DC: AERA.

Bielefeldt, T. (2012). Guidance for technology decisions from classroom observation. *Journal of Research on Technology in Education (International Society for Technology in Education)*, 44(3), 205–223.

Brickell, G., & Herrington, J. (2006). Scaffolding learners in authentic, problem based e-learning environments: The geography challenge. *Australas J Educ Technol*, 22(4), 531–547.

Brophy, S., Klein, S., Portsmore, M., & Rogers, C. (2008). Advancing engineering education in P-12 classrooms. *J Eng Educ*, 97(3), 369–387.

D'Angelo, C., Rutstein, D., Harris, C., Haertel, G., Bernard, R., & Borokhovski, E. (2014, March). Simulations for STEM Learning:

- Systematic Review and Meta-Analysis Report Overview. Menlo Park, CA: SRI Education.
- De Jong, T. (2006). Technological advances in inquiry learning. *Science*, 312(5773), 532–533.
- Devolder, A., van Braak, J., & Tondeur, J. (2012). Supporting self-regulated learning in computer-based learning environments: Systematic review of effects of scaffolding in the domain of science education. *J Comput Assist Learn*, 28(6), 557–573.
- DiBiase, W., & McDonald, J. R. (2015). Science teacher attitudes toward inquiry-based teaching and learning. *The Clearing House*, 88(2), 29–38.
- Dringenberg, E., Wertz, R. E. H., Purzer, S., & Strobel, J. (2012). *Development of the science and engineering classroom learning observation protocol*. Presented at the 2012 American Society for Engineering Education National Conference. Retrieved August 8, 2015 from <http://www.asee.org/public/conferences/8/papers/3324/view>.
- Education Super Highway. (2017). *The 2016 state of the states*. Retrieved July 30, 2018, https://s3-us-west-1.amazonaws.com/esh-sots-pdfs/2016_national_report_K12_broadband.pdf.
- Fensham, P. J. (2004). Increasing the relevance of science and technology education for all students in the 21st century. *Sci Educ Int*, 15(1), 7–26.
- Hayden, K., Ouyang, Y., Scinski, L., Olszewski, B., & Bielefeldt, T. (2011). Increasing student interest and attitudes in STEM: Professional development and activities to engage and inspire learners. *Contemporary Issues in Technology and Teacher Education*, 11(1), 47–69.
- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Commun Methods Meas*, 1(1), 77–89.
- Herrington, J., & Kervin, L. (2007). Authentic learning supported by technology: 10 suggestions and cases of integration in classrooms. *Educational Media International*, 44(3), 219–236.
- Herrington, D. G., Bancroft, S. F., Edwards, M. M., & Schairer, C. J. (2016). I want to be the inquiry guy! How research experiences for teachers change beliefs, attitudes, and values about teaching science as inquiry. *J Sci Teach Educ*, 27(2), 183–204.
- Lombardi, M. M. (2007). Authentic learning for the 21st century: An overview. In D. G. Oblinger (Ed.) *Educause Learning Initiative*, ELI Paper 1: 2007. Retrieved October 19, 2017 from: https://www.researchgate.net/profile/Marilyn_Lombardi/publication/220040581_Authentic_Learning_for_the_21st_Century_An_Overview/links/0f317531744eedf4d1000000.pdf.
- Maddux, C. D., & Johnson, D. L. (2006). Information technology, type II classroom integration, and the limited infrastructure in schools. *Comput Sch*, 22(3–4), 1–5.
- Marcum-Dietrich, N., Bruozas, M., & Staudt, C. (2019). *Embedding computational thinking into a middle school science meteorology curriculum*. Baltimore, MD: Interactive poster presented at the annual meeting of the National Association of Research in Science Teaching (NARST).
- Miller, L. M., Chang, C.-I., Wang, S., Beier, M. E., & Klisch, Y. (2011). Learning and motivational impacts of multimedia science game. *Computers and Education*, 57(1), 1425–1433.
- Moeller, B., & Reitzes, T. (2011). *Integrating technology with student-centered learning*. Quincy, MA: Education Development Center, Inc., Nellie Mae Education Foundation.
- National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.
- Nellie Mae Education Foundation, & Parthenon-EY. (2015). *The connected classroom: Understanding the landscape of technology supports for student-centered learning*.
- Parker, C. E., Stylinski, C. D., Bonney, C. R., Schillaci, R., & McAuliffe, C. (2015). Examining the quality of technology implementation in STEM classrooms: Demonstration of an evaluative framework. *J Res Technol Educ*, 47(2), 105–121.
- Patrick, S., Kennedy, K., & Powell, A. (2013). *Mean what you say: Defining and integrating personalized, blended and competency education*. Vienna, VA: International Association for K-12 Online Learning.
- Pullen, D. L., Baguley, M., & Marsden, A. (2009). Back to basics: Electronic collaboration in the education sector. In J. Salmons & L. Wilson (Eds.), *Handbook of research on electronic collaboration and organization synergy* (pp. 205–222). Hershey, PA: IGI Global.
- Ravitz, J., Wong, Y., & Becker, H. (2000). *Constructivist-compatible and practices among US teachers*. Irvine, CA: Center for Research on Information Technology and Organizations.
- Reeves, T. C., Herrington, J., & Oliver, R. (2002). *Authentic activities and online learning. Annual Conference Proceedings of Higher Education Research and Development Society of Australasia*. Perth: Retrieved October 19, 2017, from http://researchrepository.murdoch.edu.au/id/eprint/7034/1/authentic_activities_online_HERDSA_2002.pdf.
- Reif, G., Schultz, G., & Ellis, S. (2016). *A qualitative study of student-centered learning practices in New England high schools*. Boston: Nellie Mae Education Foundation and University of Massachusetts Donahue Institute.
- Rutten, N., van Joolingen, W. R., & van der Veen, J. T. (2012). The learning effects of computer simulations in science education. *Comput Educ*, 58(1), 136–153.
- Stylinski, C.D. & Doty, C. (2013). The inquiring with GIS (iGIS) project: helping teachers create and lead local GIS-based investigations. In: Pages 161-190. In J. G. MaKinster, N. M. Trautman, & M. Barnett (Eds.), *Teaching science and investigating environmental issues with geospatial technology: Designing effective professional development for teachers*. Springer Publishing Co.
- Stylinski, C.D., DeLisi, J., Wong, J., Bonney, C., Parker, C.E. & Doty, C. (2018). *Mind the Gap: Reviewing measures of quality and technology use in classroom observation protocols*. Paper presented at NARST International Conference, Atlanta GA.
- Sawada, D., Piburn, M. D., Judson, E., Turley, J., Falconer, K., Benford, R., Bloom, I. (2002). Measuring Reform Practices in Science and Mathematics Classrooms: The Reformed Teaching Observation Protocol. *Sch Sci Math*, 102(6):245–253
- Smetana, L. K., & Bell, R. L. (2011). Computer simulations to support science instruction and learning: A critical review of the literature. *Int J Sci Educ*, 34(9), 1337–1370.
- Stearns, L. M., Morgan, J., Capraro, M. M., & Capraro, R. M. (2012). A teacher observation instrument for PBL classroom instruction. *Journal of STEM Education: Innovations and Research*, 13(2), 7–16.
- Sturgis, C., & Patrick, S. (2010). *When success is the only option: Designing competency-based pathways for next generation learning*. Vienna, VA: International Association for K-12 Online Learning.